

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea in Fisica

Il filtro di Kalman: analisi di serie storiche e applicazione ai dati finanziari

Relatrice:
Prof.ssa Elisa Ercolessi

Presentata da:
Gabriele Carosi

Correlatore:
Dott. Claudio Sanavio

[...] carpe diem, quam minimum credula postero.

Indice

Abstract	1
Introduzione	2
1 Filtro di Kalman	4
1.1 Dagli space state model ai filtri	4
1.2 Previsione e aggiornamento	8
1.3 Kalman gain	16
1.4 Trattazione bayesiana	20
2 Simulazioni	24
2.1 Moto rettilineo uniforme	24
2.2 Moto uniformemente accelerato	34
2.3 Modello approssimativo	39
2.4 Oscillatore armonico	43
3 Serie storiche di rendimenti	48
3.1 L'ipotesi di random walk	48
3.2 L'esperimento di Malkiel	51
3.3 Random Walk Asset Pricing	53
3.4 Il filtro di Kalman e i titoli di borsa	57
3.4.1 Metodologia	57
3.4.2 Risultati	60
3.4.3 Approfondimento	66
Conclusione	69
Appendice	70
Moto rettilineo uniforme	70
Moto uniformemente accelerato	73
Modello approssimativo	76
Oscillatore armonico	78
Serie storiche di titoli di borsa	81
Bibliografia	85
Ringraziamenti	86

Abstract

Historia magistra vitae, scriveva Cicerone nel *De Oratore*; il passato deve insegnare a comprendere meglio il futuro. Un concetto che a primo acchito può sembrare confinato nell'ambito della filosofia e della letteratura, ma che ha invece applicazioni matematiche e fisiche di estrema importanza. Esistono delle tecniche che permettono, conoscendo il passato, di effettuare delle migliori stime del futuro? Esistono dei metodi che permettono, conoscendo il presente, di aggiornare le stime effettuate nel passato?

Nel presente elaborato viene illustrato come argomento centrale il filtro di Kalman, un algoritmo ricorsivo che, dato un set di misure di una certa grandezza fino al tempo t , permette di calcolare il valore atteso di tale grandezza al tempo $t + 1$, oltre alla varianza della relativa distribuzione prevista; permette poi, una volta effettuata la $(t + 1)$ -esima misura, di aggiornare di conseguenza valore atteso e varianza della distribuzione dei valori della grandezza in esame.

Si è quindi applicato questo algoritmo, testandone l'efficacia, prima a dei casi fisici, quali il moto rettilineo uniforme, il moto uniformemente accelerato, l'approssimazione delle leggi orarie del moto e l'oscillatore armonico; poi, introducendo la teoria di Kendall conosciuta come ipotesi di random walk e costruendo un modello di asset pricing basato sui processi di Wiener, si è applicato il filtro di Kalman a delle serie storiche di rendimenti di strumenti di borsa per osservare se questi si muovessero effettivamente secondo un modello di random walk e per prevedere il valore al tempo finale dei titoli.

Introduzione

La realtà è costituita nel suo complesso da moltissime grandezze che evolvono nel tempo, delle quali è possibile conoscere i valori che hanno assunto nel passato, ma non i valori futuri. A essere sinceri, spesso di queste grandezze non si possono conoscere neanche i valori veri, dal momento che tutte le misure sono affette da errore. Esistono tuttavia dei metodi che permettono, conoscendo il valore misurato di una grandezza al tempo presente, di stimarne il valore vero nell'istante successivo e di aggiornare la stima una volta effettuata la misura al tempo corrispondente; tali algoritmi vengono chiamati *filtri*. In questo elaborato ci si vuole concentrare su un particolare filtro noto come *filtro di Kalman*, costituito da quattro formule che definiscono in maniera univoca la distribuzione prevista del valore vero al tempo successivo e la distribuzione aggiornata del medesimo una volta effettuata la misura. Nella dimostrazione di queste formule, che costituisce il corpo principale del primo capitolo insieme a un'analisi di carattere squisitamente teorico delle varie grandezze coinvolte, si sono seguiti principalmente i lavori di [Durbin 2012] e [Reid 2001], con alcuni approfondimenti prettamente matematici tratti da [Pascucci 2009]. Si mostrerà, sempre nel corso del primo capitolo, come per altro le sopra citate formule siano dimostrabili in più di un modo, a partire da ipotesi di lavoro diverse; in particolare si proporrà una dimostrazione alternativa a partire dal Teorema di Bayes.

Nel secondo capitolo, invece, ci si concentrerà sull'applicazione del filtro di Kalman a quattro sistemi fisici di crescente difficoltà, per testarne il corretto funzionamento. Si partirà da un moto rettilineo uniforme in cui la traiettoria è disturbata da un rumore browniano, per proseguire poi con un moto uniformemente accelerato, prima assumendo un modello evolutivo conforme alle leggi del moto secondo le quali si muove il sistema in esame, poi sfruttando invece un modello approssimativo, preparando il terreno per l'ultimo esempio, ovvero l'applicazione del filtro di Kalman a un punto che si muove di moto armonico, ipotizzando un modello evolutivo ad accelerazione costante. Si vedrà e si analizzerà il comportamento delle principali grandezze in gioco e la loro evoluzione temporale. Per i casi presentati si è tratta ispirazione da [Reid 2001], introducendo tuttavia varianti interessanti, proprio come l'utilizzo di modelli non corrispondenti all'evoluzione reale dei sistemi.

Nel terzo capitolo invece, dopo un'introduzione generale dell'ipotesi di random walk e della teoria del mercato efficiente nel corso della quale si sono richiamate opere pionieristiche in tale ambito come [Fama 1965] e [Kendall 1953], si è dimostrata, seguendo il lavoro di [Requião da Cunha 2022] e partendo dai processi di Wiener, una classe molto particolare di processi stocastici, una formula che descrive l'andamento dei rendimenti dei titoli a tempo discreto con un moto browniano geometrico. Dunque si è ricondotto lo studio dell'evoluzione temporale del rendimento dei titoli di borsa a un moto rettilineo uniforme disturbato da un rumore browniano. Applicando il filtro di Kalman e ottenendo

per tre diversi titoli le serie dei rendimenti previsti, si sono da queste stimati, sulla base di banali strategie di mercato, i valori previsti dei titoli al tempo finale, confrontati poi con quelli reali: in tutti e tre i casi il valore finale viene sottostimato con una discrepanza che però, per uno dei titoli, risulta problematica in quanto prevedrebbe una decrescita del valore, quando nella realtà si ha una crescita.

Le applicazioni del filtro di Kalman, comunque, sono davvero molteplici. Esistono alcuni lavori, come quello svolto da [Mohamed 1999], che mostrano per esempio l'impiego di questo algoritmo nei GPS per prevedere e aggiornare le traiettorie. Lo studio condotto invece da [Kanieski 2013] dimostra un possibile impiego del filtro di Kalman in sistemi di controllo di carattere energetico. Ma anche nella stessa finanza gli impieghi del filtro di Kalman, a partire dal corpo teorico riguardo i processi stocastici costruito da Markov, sono ben più sostanziosi di quelli a cui ci si limiterà nel corso di questa trattazione, come dimostra per esempio lo studio condotto da [Manoliu 2010], nel quale si utilizza l'algoritmo in esame per prevedere l'evoluzione del prezzo di alcuni derivati sull'energia.

Non sorprende, comunque, che uno strumento che fornisce una chiave per stimare il futuro e aggiornare le previsioni sia così potente e diffusamente utilizzato. Di fatto, per quanto sia scarsamente conosciuto come argomento, è più presente nella vita quotidiana di quanto quasi tutte le persone possano immaginare. Per di più le premesse da cui partire per svilupparne il corpo teorico sono alquanto semplici e si procede quindi nel seguito a illustrarle.

Capitolo 1

Filtro di Kalman

In questo primo capitolo verranno dimostrate le formule di previsione e di aggiornamento che nel complesso costituiscono il filtro di Kalman, prestando poi particolare attenzione ad alcune grandezze estremamente emblematiche ed esplicative del comportamento e del funzionamento del filtro stesso, come l'innovazione e il Kalman gain. Imponendo determinate ipotesi di lavoro, che saranno dettagliatamente illustrate nella prima sezione, si otterranno risultati che verranno poi confermati anche tramite una metodologia differente, ossia dimostrati a partire dal teorema di Bayes.

A livello formale, a partire dal presente capitolo e per tutto il resto della trattazione, poiché si utilizzeranno elementi che a seconda delle circostanze possono essere sia vettori sia matrici, in conformità alla notazione adottata da [Reid 2001] si indicherà con \mathbf{A} un qualsiasi termine vettoriale o matriciale, denotando conseguentemente con $\mathbf{A} \cdot \mathbf{B}$ il prodotto righe per colonne tra \mathbf{A} e \mathbf{B} , o equivalentemente, qualora tale operazione sia il prodotto di un vettore colonna per un vettore riga, il prodotto tensoriale tra \mathbf{A} e \mathbf{B} .

1.1 Dagli space state model ai filtri

In questa prima sezione vengono introdotti, seguendo la trattazione svolta in [Durbin 2012], gli *space state model*, ossia modelli matematici che permettono di trattare con una unica metodologia di base numerosi problemi di analisi delle serie temporali e grazie ai quali si pongono le basi necessarie per la trattazione del filtro di Kalman.

Si consideri per il momento una grandezza x che varia discretamente nel tempo e si consideri l'insieme dei valori che tale grandezza ha effettivamente assunto in ciascuno degli istanti discreti di tempo $t = 0, \dots, t = n$, ossia la *serie storica* dei valori veri x_1, \dots, x_n ; si supponga inoltre di aver effettuato in ciascuno degli istanti di tempo una misura della grandezza, ottenendo la serie storica delle misure y_1, \dots, y_n . Come anche per tutti gli altri osservabili della fisica, non è possibile conoscere il valore vero che x assume in un determinato istante di tempo, del quale si può al limite effettuare una stima tramite alcuni strumenti della probabilità e della statistica; lo *space state model* è uno di questi e consente di individuare una relazione tra i valori veri e quelli misurati e che consente, in particolare, di dedurre le proprietà degli x_t a partire dagli y_t . Il modello base per la rappresentazione di una serie storica è il cosiddetto modello additivo, basato sulla seguente relazione funzionale:

$$y_t = \varphi_t + \gamma_t + \varepsilon_t \tag{1.1}$$

dove φ_t è detto trend ed è una variabile la cui derivata temporale è in modulo piccola, γ_t è detto seasonal ed è una componente periodica, ε_t è un termine irregolare detto errore. Per il momento si assume che tutti i termini siano scalari, senza perdita di generalità rispetto al caso vettoriale.

Una classe molto importante di serie storiche sono quelle descrivibili tramite un modello di *random walk*, per le quali vale cioè la seguente relazione:

$$x_{t+1} = x_t + \eta_t \quad (1.2)$$

dove le η_t sono tutte variabili aleatorie distribuite secondo una distribuzione normale con media nulla e medesima varianza finita, ossia indicando con

$$x \sim N(a; b) \quad (1.3)$$

una variabile aleatoria distribuita secondo una distribuzione normale con media a e varianza b , si ha che

$$\eta_t \sim N(0; \sigma_\eta^2) \quad \forall t \quad (1.4)$$

Sistemi il cui comportamento è descrivibile tramite il *local level model* costituiscono un esempio di space state model in cui il trend è un random walk, non è presente seasonal e tutte le variabili aleatorie sono distribuite secondo una distribuzione normale. Il sistema tramite cui si definisce il local level model è il seguente:

$$\begin{cases} y_t = x_t + \varepsilon_t \\ x_{t+1} = x_t + \eta_t \end{cases} \quad (1.5)$$

dove ε_t è l'errore sulla misura e η_t è il rumore del random walk; sono ambedue errori distribuiti gaussianamente con media nulla e con una varianza data, ossia $\varepsilon_t \sim N(0; \sigma_\varepsilon^2)$ e $\eta_t \sim N(0; \sigma_\eta^2)$. Questo modello, seppur semplice, è estremamente importante, perché costituisce un mezzo efficace per studiare numerosi fenomeni, tra cui, come si vedrà verso la fine della trattazione, l'evoluzione temporale dei titoli di borsa. Presenta inoltre tutte le caratteristiche tipiche di uno space state model: gli x_t sono gli stati, ovvero quel set di valori veri che non possono essere conosciuti, mentre le y_t sono le osservazioni. In questo framework, come poc'anzi anticipato, gli ε_t rappresentano l'errore dello strumento di misura, che impedisce una perfetta stima dello stato x_t , mentre gli η_t denotano il rumore del sistema. Questo è il tipico modello che descrive una particella soggetta a moto browniano senza drift aggiuntivi: in altre parole, la particella rimarrebbe ferma se non fosse affetta dal rumore gaussiano. Tale modello, giacché le variabili sono funzioni del tempo, sarà un modello non stazionario (le grandezze in esame non sono cioè costanti nel tempo); si sa inoltre che sia ε_t , sia η_t non dipendono da x_1 , cioè gli errori sono indipendenti dal valore vero iniziale della grandezza in esame. A questo punto, avendo analizzato la relazione sussistente tra valori veri e misurati per i sistemi descrivibili tramite un local level model, potrebbe sorgere una domanda: qual è il valore di aspettazione¹ di x_t data la serie storica delle misure?

Come già ampiamente esplicitato, non è possibile conoscere con certezza i valori di x_t , ragion per cui questi seguono delle distribuzioni di probabilità: in altre parole, una x_t

¹Il valore di aspettazione si calcola come valor medio.

può assumere una gamma di possibili valori, alcuni più probabili e altri meno probabili, distribuiti secondo una certa densità di probabilità. Si assuma quindi che la distribuzione dei possibili valori di x_1 sia una distribuzione normale con media μ_1 e con una certa varianza, sia essa P_1 . Si immagini poi di effettuare un certo numero t di misure della grandezza che possono essere rappresentate in un vettore \mathbf{Y}_n . Per poter rispondere al quesito posto occorre quindi saper descrivere tale vettore e ciò richiede, dal momento che si deve considerare una successione di variabili aleatorie, l'introduzione di alcuni elementi, come le matrici di covarianza, che rientrano nella teoria dell'analisi multivariata: questa consiste in una generalizzazione del caso unidimensionale al caso di dimensioni superiori, ossia con più variabili. Quanto si ricava dalle relazioni introdotte per sistemi descrivibili tramite il local level model è che \mathbf{Y}_t è un vettore t -variato, ossia che una qualsiasi combinazione lineare delle sue t componenti sia una distribuzione normale unidimensionale. Infatti la seconda relazione dell'Equazione (1.5), si può riscrivere nella maniera di seguito mostrata:

$$x_t = x_1 + \sum_{j=1}^{j=t-1} \eta_j \quad (1.6)$$

Sostituendo tale uguaglianza nella prima relazione dell'Equazione (1.5) si ottiene la seguente espressione per la t -esima misura:

$$y_t = x_1 + \sum_{j=1}^{j=t-1} \eta_j + \varepsilon_t \quad (1.7)$$

Dal momento che i termini η_t e ε_t sono variabili aleatorie distribuite normalmente con media nulla, mentre la media della distribuzione di x_1 è μ_1 , la media dei possibili valori che si possono ottenere dalla misura al tempo t è la costante μ_1 . Ragionando sulla varianza e trascurando per il momento P_1 , si osserva che si hanno due contributi, ossia σ_η^2 e σ_ε^2 , a partire dai quali si può costruire la matrice di covarianza, i cui elementi sono definiti nel seguente modo: ogni termine sulla diagonale principale, che si può indicare con σ_{ii}^2 , è la varianza della distribuzione della i -esima componente del vettore cui la matrice di covarianza fa riferimento e per questo è sempre un valore non negativo; ogni elemento σ_{ij}^2 con $i \neq j$ è invece la covarianza tra le due componenti di indice corrispondente nel vettore cui la matrice fa riferimento e nel caso in cui assuma valore positivo, significa che al crescere di una delle due componenti, cresce anche l'altra, se assume valore negativo accade il contrario, se è nulla allora le due componenti sono indipendenti. Il contributo dato dalle ε_t alla varianza deve comparire solamente sulla diagonale principale: esso infatti quantifica l'incertezza che si ha esclusivamente sulla t -esima misura e il valore che assume è indipendente dalle precedenti misure, per cui ε_t è indipendente da \mathbf{Y}_{t-1} e quindi non può comparire nei termini di correlazione tra le diverse componenti di quest'ultimo vettore. Occorre quindi identificare la correlazione tra le misure: supponendo di avere, per esempio, un punto materiale che si muove secondo una determinata legge, la posizione a un determinato istante t dipende dalla posizione all'istante $t - 1$. Per approfondire questo passaggio, si supponga di non considerare le η_j ; allora, facendo riferimento all'Equazione (1.7), gli elementi del vettore \mathbf{Y}_t costituirebbero una popolazione di variabili aleatorie indipendenti identicamente distribuite, che rappresentano tutte un possibile valore della misura all'istante iniziale. Risulta chiaro quindi che una eventuale correlazione tra i diversi elementi di \mathbf{Y}_t debba essere data dalle η_j . Più nel dettaglio,

per ogni t , nell'espressione di y_t compaiono $t - 1$ termini del tipo η_j , ciascuno dei quali è distribuito normalmente con varianza σ_η^2 . Quindi il termine y_t avrà una correlazione nulla con y_{t-1} ²; avrà una correlazione pari a σ_η^2 con y_{t-2} , avrà una correlazione pari a $2\sigma_\eta^2$ con y_{t-3} e così via fino ad arrivare a y_1 . Tuttavia, si deve tener presente che i termini del tipo η_j , comparando nell'espressione di y_t , contribuiscono anche alla varianza, oltre che alla covarianza: essendo tali termini in numero $t - 1$ nell'espressione di y_t , si aggiungerà alla varianza data da ε_t precedentemente computata un valore pari a $(t - 1)\sigma_\eta^2$. Infine occorre sommare ai termini nella diagonale principale la varianza della distribuzione di x_1 , in quanto essa contribuisce, in accordo con l'Equazione (1.7), alla varianza di ciascuna delle misure, ma non ne costituisce elemento di correlazione. Si ottiene di conseguenza che il vettore \mathbf{Y}_t è un vettore distribuito secondo la seguente distribuzione normale:

$$\mathbf{Y}_t \sim N \left(\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_1 \end{pmatrix}, P_1 \mathbf{I} + \boldsymbol{\Sigma} \right) \quad (1.8)$$

dove con \mathbf{I} si denota la matrice identità e per quanto detto sinora la matrice $\boldsymbol{\Sigma}$ è costruita come segue.

$$(\boldsymbol{\Sigma})_{ij} = \begin{cases} (i - 1)\sigma_\eta^2 & \text{per } i < j \\ \sigma_\varepsilon^2 + (i - 1)\sigma_\eta^2 & \text{per } i = j \\ (j - 1)\sigma_\eta^2 & \text{per } i > j \end{cases} \quad (1.9)$$

Avendo esplicitato i parametri che definiscono univocamente la distribuzione del vettore \mathbf{Y}_t , si può tornare alla domanda dalla quale si è partiti: qual è il valore di aspettazione di x_t data la serie storica delle misure? Prima di procedere si precisa il significato di una ulteriore notazione: nel seguito della trattazione di indicherà con $p(x)$ la distribuzione di probabilità di una qualsiasi variabile aleatoria x . Si è già posta come ipotesi che la distribuzione di probabilità del primo valore vero della grandezza in esame sia

$$p(x_1) = N(\mu_1; P_1) \quad (1.10)$$

Siano poi date t misure di tale grandezza; come ipotesi di lavoro, in risposta al quesito posto, si assume che la distribuzione di probabilità di x_t date le misure fino al tempo t sia

$$p(x_t | \mathbf{Y}_t) = N(\mu_t; P_t) \quad (1.11)$$

dove, in termini generali, con $p(x|a)$ si indica la distribuzione di probabilità della variabile aleatoria x conoscendo il valore dell'elemento a , che può essere sia uno scalare, sia un vettore come in questo caso. Si immagini ora di voler effettuare una previsione. Senza conoscere la $(t + 1)$ -esima misura, quale sarà la distribuzione di x_{t+1} ? Ossia:

$$p(x_{t+1} | \mathbf{Y}_t)$$

che sarà nel seguito chiamata *distribuzione prevista*, che forma assume? Si supponga poi di effettuare la $(t + 1)$ -esima misura: come si può aggiustare la previsione effettuata? Ossia:

$$p(x_{t+1} | \mathbf{Y}_{t+1})$$

²Questo accade perché la sommatoria delle η_j si estende fino a $t - 1$ e non fino a t .

che sarà chiamata *distribuzione aggiornata*, che forma assume? Teoricamente nulla impedirebbe di utilizzare la conoscenza acquisita riguardo il vettore \mathbf{Y}_t per poter calcolare, sfruttando i risultati dell'analisi multivariata, medie, varianze e covarianze condizionate. Tuttavia, come facilmente si può immaginare, siccome le y_t sono concatenate a formare una serie storica, questo approccio risulterebbe estremamente dispendioso e difficile in termini di calcolo. Esistono però degli strumenti estremamente potenti che permettono di rispondere al quesito posto, ottenendo gli stessi risultati che si otterrebbero attraverso l'applicazione pedissequa delle regole dell'analisi multivariata: si tratta dei *filtri*, ossia algoritmi ricorsivi che, note le condizioni al tempo t , consentono di aggiornare la distribuzione dei valori veri al tempo t e di prevedere la distribuzione dei valori veri al tempo $t + 1$. Nel prosieguo ci si concentrerà nell'analisi di un particolare filtro, ovvero il *filtro di Kalman*.

1.2 Previsione e aggiornamento

Come già accennato, supponendo di aver effettuato t misure della grandezza \mathbf{x} , che verrà d'ora in avanti trattata come un vettore per una trattazione più esaustiva dell'argomento, il filtro di Kalman è uno strumento che permette di effettuare una previsione sulla distribuzione del valore vero della grandezza in esame al tempo $t + 1$ e di aggiornare tale distribuzione una volta eseguita la $(t + 1)$ -esima misura. In questo paragrafo l'obiettivo è quindi quello di trovare, facendo riferimento in generale alla trattazione svolta da [Reid 2001], dapprima le espressioni della media e della varianza della distribuzione prevista del valore vero, ossia quella che si può stimare di avere al tempo $t + 1$ date le misure fino al tempo t ; poi si troveranno le espressioni della media e della varianza della distribuzione aggiornata del valore vero, ossia quella che viene individuata per il tempo $t + 1$ una volta effettuata la $(t + 1)$ -esima misura. Per poter rappresentare tali distribuzioni occorre quindi introdurre innanzitutto una appropriata notazione per indicare ai valori previsti e aggiornati. Nel seguito si scriverà

$$a_{t+1|t}$$

per riferirsi alla stima del valore del generico elemento a al tempo $t + 1$ sapendo che sono state effettuate le misure fino al tempo t ; in altre parole, si *condiziona* la stima del valore di a al tempo $t + 1$ alle misure effettuate fino al tempo t . Dunque la notazione $a_{t+1|t}$ viene utilizzata per riferirsi ai valori previsti. Per quanto riguarda gli aggiornamenti, siccome in tal caso si deve avere a disposizione la conoscenza delle misure fino al tempo a cui la stima del valore è riferita, si scriverà

$$a_{t+1|t+1}$$

per riferirsi al valore aggiornato di a . Sfruttando questa notazione, si può già fornire, grazie all'ipotesi di lavoro presentata nel precedente paragrafo, una rappresentazione della forma assunta dalle distribuzioni prevista e aggiornata:

$$p(\mathbf{x}_{t+1}|\mathbf{Y}_t) = N(\boldsymbol{\mu}_{t+1|t}; \mathbf{P}_{t+1|t}) \quad (1.12)$$

$$p(\mathbf{x}_{t+1}|\mathbf{Y}_{t+1}) = N(\boldsymbol{\mu}_{t+1|t+1}; \mathbf{P}_{t+1|t+1}) \quad (1.13)$$

dove le $\boldsymbol{\mu}$ indicano le medie delle distribuzioni e le \mathbf{P} le varianze, che, come si vedrà nel seguito, sono in realtà matrici di covarianza.

Come già anticipato, si supponga di trattare \mathbf{x}_t come un vettore, in modo da illustrare l'argomento in maniera più approfondita. Nel seguito ci si riferirà alle \mathbf{x}_t chiamando tali elementi semplicemente *vettori di stato*, sapendo che si tratta, per la precisione, dei valori veri dei vettori di stato. In questo caso, come insegna la meccanica analitica, esisterà una matrice che determina l'evoluzione temporale del sistema, detta *matrice della transizione di stato*, che sarà nel seguito indicata con \mathbf{F} e che si assume essere costante, per non complicare eccessivamente la trattazione. Esisterà allora una relazione del tipo

$$\mathbf{x}_{t+1} = \mathbf{F} \cdot \mathbf{x}_t \quad (1.14)$$

In realtà questa relazione risulta incompleta. Infatti, come già introdotto nella precedente sezione, il valore vero è soggetto a un errore, che potrebbe sia essere dato da una intrinseca variazione statistica, qualora si consideri un campione di valori, sia essere dato, nel caso in cui si abbiano vettori di stato soggetti a leggi fisiche, da elementi esterni di disturbo, come il vento in una traiettoria. Questo errore è un rumore browniano e viene indicato con il vettore \mathbf{w}_t ; esso deve avere chiaramente la stessa dimensione del vettore di stato: la i -esima componente di \mathbf{w}_t corrisponde infatti al rumore browniano sulla i -esima componente di \mathbf{x}_t . Considerando la seconda relazione dell'Equazione (1.5) e generalizzandola al caso in cui le grandezze in esame non siano scalari ma vettori, risulta ragionevole definire come legge dell'evoluzione temporale dei vettori di stato la seguente uguaglianza:

$$\mathbf{x}_{t+1} = \mathbf{F} \cdot \mathbf{x}_t + \mathbf{w}_t \quad (1.15)$$

Siccome le \mathbf{x}_t sono trattate come vettori, di conseguenza anche le misure saranno dei vettori e saranno indicate con \mathbf{y}_t . A tale proposito, risulta necessario introdurre una matrice, che si indica con \mathbf{H} ed è chiamata *matrice di transizione in uscita*, che trasformi i valori veri in valori misurati. Infatti il vettore di stato \mathbf{x}_t può avere anche dimensione differente rispetto al vettore della misura \mathbf{y}_t , qualora non tutte le componenti del vettore di stato vengano misurate e lo scopo di \mathbf{H} è rendere confrontabili tali grandezze, ossia trasformare la dimensione del vettore di stato nella dimensione del vettore della misura. Per esempio, se si ha un vettore di stato di dimensione n e si decide di misurarne solamente le prime k componenti, \mathbf{H} sarà una matrice $k \times n$ tale che

$$(\mathbf{H})_{ij} = \begin{cases} 1 & \text{per } i = j \\ 0 & \text{per } i \neq j \end{cases} \quad (1.16)$$

Per ottenere l'espressione della t -esima misura, al prodotto tra \mathbf{H} e \mathbf{x}_t occorre sommare un altro vettore, indicato con \mathbf{z}_t , che definisca l'errore sulla misura al tempo t e avente quindi stessa dimensione di \mathbf{y}_t . Di conseguenza si ha la seguente relazione:

$$\mathbf{y}(t) = \mathbf{H} \cdot \mathbf{x}(t) + \mathbf{z}(t) \quad (1.17)$$

Per poter procedere a dimostrare la forma che le distribuzioni aggiornata e prevista assumono, occorre introdurre la funzione

$$E : X \rightarrow \mathbb{R}^d$$

che rappresenta il valore atteso: esso viene calcolato su una grandezza, in generale di dimensione d , appartenente allo spazio X e distribuita secondo una certa distribuzione di probabilità e restituisce il valore medio di tale distribuzione, che dovrà chiaramente avere la stessa dimensione della variabile di cui si calcola il valore atteso e quindi appartenere a \mathbb{R}^d . Il valore atteso è una funzione lineare e può essere condizionato, restituendo come valore in tal caso la media della distribuzione di probabilità $\mathbf{x} \in X$ condizionata a una certa informazione. Più avanti si approfondirà dal punto di vista matematico quest'ultimo aspetto.

Iniziando con lo studio della distribuzione prevista, si può sfruttare la notazione precedentemente introdotta per poter scrivere nel seguente modo la media di tale distribuzione:

$$\boldsymbol{\mu}_{t+1|t} = E[\mathbf{x}_{t+1}|\mathbf{Y}_t] \quad (1.18)$$

Poiché la matrice F è costante e la funzione E è lineare, si ha la seguente uguaglianza:

$$\boldsymbol{\mu}_{t+1|t} = \mathbf{F} \cdot E[\mathbf{x}_t|\mathbf{Y}_t] \quad (1.19)$$

dove il termine a destra corrisponde proprio alla definizione di $\boldsymbol{\mu}_t$, per cui

$$\boldsymbol{\mu}_{t+1|t} = \mathbf{F} \cdot \boldsymbol{\mu}_t \quad (1.20)$$

Si è quindi mostrata la procedura per trovare la media della distribuzione del vettore di stato al tempo $t + 1$ una volta nota la media della distribuzione al tempo t .

Individuata questa prima legge per la media, si pone ora il problema di capire come si comportano le varianze. Occorre innanzitutto notare che, non essendo più il valore vero uno scalare, bensì un vettore, la varianza sarà espressa tramite una matrice di covarianza, costruita nella maniera mostrata di seguito. Data una popolazione statistica di n elementi su cui sono rilevati k caratteri quantitativi indicati con \mathbf{A} , ogni termine \mathbf{A}_i , con $i = 1, \dots, k$, è un vettore di n elementi, ciascuno indicato con a_{ji} : in particolare, si dice che a_{ji} è la modalità della j -esima unità statistica dell' i -esimo carattere. Date queste premesse, secondo la trattazione dell'analisi multivariata effettuata da Harold Hotelling, matematico e statistico statunitense pioniere in tale ambito, l'elemento generico della matrice di covarianza è così definito³:

$$\sigma_{ik}^2 = \frac{1}{n} \sum_{j=1}^n (a_{ji} - \mu_i)(a_{jk} - \mu_k) \quad (1.21)$$

dove con μ_i si denota il valor medio dell' i -esimo carattere. Risulta quindi ragionevole, alla luce della spiegazione riportata, definire la matrice di covarianza per il problema in esame nel seguente modo:

$$\mathbf{P}_t = (\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot (\mathbf{x}_t - \boldsymbol{\mu}_t)^T \quad (1.22)$$

dove con l'apice T si denota il trasposto di un vettore o di una matrice. Al fine di avere una notazione comoda per eseguire i passaggi successivi, è possibile riscrivere la matrice \mathbf{P} come mostrato nella seguente relazione:

$$\mathbf{P}_{t|t} = E[(\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot (\mathbf{x}_t - \boldsymbol{\mu}_t)^T | \mathbf{Y}_t] \quad (1.23)$$

³La definizione adottata è quella mostrata in [Fornasini 2008].

nella quale si condiziona appunto la varianza alle misure effettuate fino al tempo t e si deve perciò considerare il valore atteso dello scarto medio, proprio perché si sta eseguendo un condizionamento. Rispetto al caso precedente, si è semplicemente esplicitato quanto si era solamente supposto, cioè di aver effettuato le misure fino al tempo t . D'ora in poi si assume come matrice di covarianza di riferimento la matrice dell'Equazione (1.23). In maniera coerente con la definizione riportata e con la notazione introdotta, si può scrivere

$$\mathbf{P}_{t+1|t} = E[(\mathbf{x}_{t+1|t} - \boldsymbol{\mu}_{t+1|t}) \cdot (\mathbf{x}_{t+1|t} - \boldsymbol{\mu}_{t+1|t})^T | \mathbf{Y}_t] \quad (1.24)$$

Esplicitando l'espressione dei vettori di stato e della media della distribuzione prevista, si può riscrivere la precedente uguaglianza nella seguente maniera:

$$\mathbf{P}_{t+1|t} = E[(\mathbf{F} \cdot \mathbf{x}_t + \mathbf{w}_t - \mathbf{F} \cdot \boldsymbol{\mu}_t) \cdot (\mathbf{F} \cdot \mathbf{x}_t + \mathbf{w}_t - \mathbf{F} \cdot \boldsymbol{\mu}_t)^T | \mathbf{Y}_t] \quad (1.25)$$

Applicando poi la proprietà distributiva si ottiene

$$\begin{aligned} \mathbf{P}_{t+1|t} = E[(\mathbf{F} \cdot \mathbf{x}_t - \mathbf{F} \cdot \boldsymbol{\mu}_t) \cdot (\mathbf{F} \cdot \mathbf{x}_t - \mathbf{F} \cdot \boldsymbol{\mu}_t)^T + (\mathbf{F} \cdot \mathbf{x}_t - \mathbf{F} \cdot \boldsymbol{\mu}_t) \cdot \mathbf{w}_t^T + \\ + \mathbf{w}_t \cdot (\mathbf{F} \cdot \mathbf{x}_t - \mathbf{F} \cdot \boldsymbol{\mu}_t)^T + \mathbf{w}_t \cdot \mathbf{w}_t^T | \mathbf{Y}_t] \end{aligned} \quad (1.26)$$

Ricordando che \mathbf{F} è costante e che la funzione E è lineare si ha

$$\begin{aligned} \mathbf{P}_{t+1|t} = \mathbf{F} \cdot E[(\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot (\mathbf{x}_t - \boldsymbol{\mu}_t)^T | \mathbf{Y}_t] \cdot \mathbf{F}^T + E[\mathbf{w}_t \cdot \mathbf{w}_t^T | \mathbf{Y}_t] + \\ + \mathbf{F} \cdot E[(\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot \mathbf{w}_t^T | \mathbf{Y}_t] + E[\mathbf{w}_t \cdot (\mathbf{x}_t - \boldsymbol{\mu}_t)^T | \mathbf{Y}_t] \cdot \mathbf{F}^T \end{aligned} \quad (1.27)$$

Per poter procedere nella dimostrazione occorre però soffermarsi su alcuni elementi di teoria della misura, tratti da [Pascucci 2009]. Siano dati due spazi misurabili (A, \mathcal{F}) e (B, \mathcal{G}) , dove ogni spazio misurabile è costituito dalla coppia di un insieme e di una σ -algebra su tale insieme. Si consideri una funzione

$$f : A \rightarrow B$$

Tale funzione è detta misurabile rispetto alla σ -algebra \mathcal{F} se e solo se per definizione

$$f^{-1}(G) \in \mathcal{F} \quad \forall G \in \mathcal{G}$$

Sia dato ora uno spazio di probabilità $(\Omega, \mathcal{H}, \mathbb{P})$, dove Ω è lo spazio campionario, \mathcal{H} è la σ -algebra degli eventi probabilizzati e \mathbb{P} è la misura di probabilità. Una funzione

$$X : \Omega \rightarrow \mathbb{R}^d$$

è una variabile aleatoria se e solo se per definizione

$$X \in m(\mathcal{H})$$

dove con $m(\mathcal{H})$ si indica l'insieme delle funzioni misurabili rispetto alla σ -algebra \mathcal{H} ; in altre parole, si ha una variabile aleatoria se e solo se per definizione tale funzione attribuisce un valore reale a tutti e i soli gli eventi probabilizzati. Risulta chiaro, osservando l'Equazione (1.7), che le componenti del vettore \mathbf{Y}_t sono delle variabili aleatorie; in particolare, sinora si è per semplicità sostituito alla variabile aleatoria il valore che essa assume in corrispondenza di un determinato evento. Per di più, la teoria della probabilità

afferma che una σ -algebra può essere generata anche da un vettore aleatorio e questa è la σ -algebra generata dall'insieme dei valori che possono assumere le componenti del vettore aleatorio. D'ora in avanti si assumerà di indicare con \mathbf{Y}_t la σ -algebra generata da tale vettore. Infine si può anche condizionare il valore atteso di una variabile aleatoria rispetto a una σ -algebra. Esso corrisponde al valore atteso della variabile aleatoria in esame limitatamente all'insieme degli eventi a cui corrispondono, tramite la stessa variabile aleatoria, valori nell'insieme che genera la σ -algebra rispetto alla quale si sta condizionando. Indicando con \mathcal{B} la σ -algebra dei boreliani, ossia la σ -algebra generata da tutti gli aperti di \mathbb{R} e con $\sigma(J)$ la σ -algebra generata dalla variabile aleatoria J , ossia la σ -algebra generata dall'insieme dei valori che tale variabile aleatoria assume, si può citare un importante risultato nella teoria della misura, noto con il nome di *Lemma di Freezing*, secondo cui, equipaggiato lo spazio campionario con la σ -algebra degli eventi probabilizzati \mathcal{H} , data una funzione f a valori scalari tale che $f \in m(\mathcal{B})$, una variabile aleatoria $K \in m(\mathcal{H})$ e una variabile aleatoria J tale che $\sigma(J)$ è indipendente da \mathcal{H} , allora si ha che

$$E[f(J, K)|\mathcal{H}] = E[f(J, k)]|_{k=K} \quad (1.28)$$

cioè si calcola il valore atteso della funzione "congelando" la sua componente misurabile rispetto alla σ -algebra degli eventi. Si è ritenuto opportuno svolgere questa breve digressione anche per attribuire un maggiore spessore matematico agli argomenti trattati.

Nel caso in esame, tornando al filtro di Kalman, si può porre:

$$\begin{aligned} \mathcal{H} &= \sigma(\mathbf{Y}_t) = \mathbf{Y}_t \\ J &= \mathbf{w}_t \\ K &= (\mathbf{x}_t - \boldsymbol{\mu}_t) \end{aligned} \quad (1.29)$$

Infatti nella prima uguaglianza si è semplicemente sfruttata la notazione introdotta; inoltre il rumore browniano è indipendente dalle misure, o meglio i valori che assume l'errore sistematico sono indipendenti dai valori che possono assumere le misure effettuate fino al tempo t ; infine il vettore $\mathbf{x}_t - \boldsymbol{\mu}_t$ è misurabile rispetto alla σ -algebra generata dal vettore delle misure: infatti, in termini qualitativi, a ogni valore che assume \mathbf{y}_t e di conseguenza il vettore \mathbf{Y}_t , esiste un possibile valore di $\mathbf{x}_t - \boldsymbol{\mu}_t$. Infine si pone $f(\mathbf{J}, \mathbf{K}) = \mathbf{J} \cdot \mathbf{K}$. Applicando il Lemma di Freezing si ottiene

$$E[(\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot \mathbf{w}^T | \mathbf{Y}_t] = E[\mathbf{s} \cdot \mathbf{w}_t] |_{\mathbf{s}=(\mathbf{x}_t - \boldsymbol{\mu}_t)} \quad (1.30)$$

Per la linearità del valore atteso si ha

$$E[(\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot \mathbf{w}^T | \mathbf{Y}_t] = \mathbf{s} |_{\mathbf{s}=(\mathbf{x}_t - \boldsymbol{\mu}_t)} \cdot E[\mathbf{w}_t^T] \quad (1.31)$$

Ma poiché il rumore browniano sui vettori di stato è distribuito secondo una distribuzione normale con media nulla e dunque il suo valore atteso è nullo, si ha che

$$E[(\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot \mathbf{w}^T | \mathbf{Y}_t] = 0 \quad (1.32)$$

Analogamente si dimostra che

$$E[\mathbf{w}_t \cdot (\mathbf{x}_t - \boldsymbol{\mu}_t)^T | \mathbf{Y}_t] = 0 \quad (1.33)$$

Dunque si perviene alla seguente uguaglianza:

$$\mathbf{P}_{t+1|t} = \mathbf{F} \cdot E[(\mathbf{x}_t - \boldsymbol{\mu}_t) \cdot (\mathbf{x}_t - \boldsymbol{\mu}_t)^T | \mathbf{Y}_t] \cdot \mathbf{F}^T + E[\mathbf{w}_t \cdot \mathbf{w}_t^T | \mathbf{Y}_t] \quad (1.34)$$

Ricordando l'espressione per la matrice di covarianza mostrata nell'Equazione (1.23), si può allora scrivere

$$\mathbf{P}_{t+1|t} = \mathbf{F} \cdot \mathbf{P}_{t|t} \cdot \mathbf{F}^T + E[\mathbf{w}_t \cdot \mathbf{w}_t^T | \mathbf{Y}_t] \quad (1.35)$$

Poiché il rumore browniano è indipendente dalle misure, il condizionamento non ha effetto su di esso e di conseguenza

$$\mathbf{P}_{t+1|t} = \mathbf{F} \cdot \mathbf{P}_{t|t} \cdot \mathbf{F}^T + E[\mathbf{w}_t \cdot \mathbf{w}_t^T] \quad (1.36)$$

Si introduce ora la matrice \mathbf{Q} , ossia la matrice di covarianza della distribuzione del rumore browniano; \mathbf{Q}_t è quindi la matrice di covarianza del rumore \mathbf{w}_t al tempo t ed è definita nel seguente modo:

$$\mathbf{Q}_t = E[\mathbf{w}_t \cdot \mathbf{w}_t^T] \quad (1.37)$$

In particolare, come verrà mostrato contestualmente alle simulazioni, esistono dei vincoli sulla costruzione della matrice \mathbf{Q} , che dipendono dalla natura del problema in esame. Per di più in tale sede verrà utilizzata una matrice \mathbf{Q} costante, per non rendere eccessivamente difficoltosa la realizzazione delle simulazioni medesime. Grazie a questa matrice si può scrivere l'espressione della matrice di covarianza prevista:

$$\mathbf{P}_{t+1|t} = \mathbf{F} \cdot \mathbf{P}_{t|t} \cdot \mathbf{F}^T + \mathbf{Q}_t \quad (1.38)$$

Si sono quindi ottenute le relazioni di previsione della media e della varianza, i due termini che definiscono univocamente la distribuzione dei vettori di stato previsti. Occorre quindi trovare le relazioni per l'aggiornamento. Chiaramente la media della distribuzione aggiornata dovrà tener conto sia della previsione sia della misura effettuata, per cui si inizia dal supporre per la media una relazione del tipo

$$\boldsymbol{\mu}_{t+1|t+1} = \mathbf{K}'_{t+1} \cdot \boldsymbol{\mu}_{t+1|t} + \mathbf{K}_{t+1} \cdot \mathbf{y}_{t+1} \quad (1.39)$$

dove \mathbf{K}'_{t+1} e \mathbf{K}_{t+1} sono le cosiddette *matrici di guadagno*, che in un certo senso dovranno quantificare il peso, nell'elaborazione dell'aggiornamento, attribuito alla previsione e alla misura; nel seguito si darà un sostegno matematico a questa impostazione qualitativa, ma per il momento le si assuma come due generiche matrici, sulle quali viene imposta tuttavia una condizione essenziale per ottenere la migliore stima per l'aggiornamento: si devono trovare le matrici di guadagno che minimizzino l'errore. Inoltre è importante specificare che queste matrici in generale non hanno la stessa dimensione: infatti, detta n la dimensione del vettore di stato e m la dimensione della misura, allora \mathbf{K}'_{t+1} è una matrice di dimensione $n \times n$, mentre \mathbf{K}_{t+1} è una matrice di dimensione $n \times m$. Tenendo conto dell'espressione delle misure, ossia dell'Equazione (1.17), si può riscrivere l'espressione precedente nel seguente modo, sfruttando la linearità del valore atteso:

$$E[\boldsymbol{\mu}_{t+1|t+1}] = \mathbf{K}'_{t+1} \cdot E[\boldsymbol{\mu}_{t+1|t}] + \mathbf{K}_{t+1} \cdot \mathbf{H} \cdot E[\mathbf{x}_{t+1}] + \mathbf{K}_{t+1} \cdot E[\mathbf{z}_{t+1}] \quad (1.40)$$

Si impone ora la condizione che equivale all'assenza di bias, ossia che

$$E[\boldsymbol{\mu}_{t+1|t+1}] = E[\mathbf{x}_{t+1}] \quad (1.41)$$

Ciò equivale a dire che, una volta effettuata la $(t + 1)$ -esima misura, il valore atteso della media della distribuzione del vettore di stato equivale al valore atteso del vettore di stato stesso al tempo $t + 1$. Per migliorare l'Equazione (1.40), si può tener conto dell'Equazione (1.18) e scriverla come argomento del valore atteso:

$$E[\boldsymbol{\mu}_{t+1|t}] = E[\mathbf{F} \cdot \boldsymbol{\mu}_{t|t}] \quad (1.42)$$

Essendo poi \mathbf{F} costante la si può estrarre dalla funzione E ; si può poi riapplicare la condizione di assenza di bias e si ottiene

$$E[\boldsymbol{\mu}_{t+1|t}] = \mathbf{F} \cdot E[\boldsymbol{x}_t] \quad (1.43)$$

Se poi si applica l'Equazione (1.14) si arriva ad una espressione molto interessante, in quanto si stabilisce una uguaglianza tra i valori attesi della distribuzione al tempo $t + 1$ e la previsione:

$$E[\boldsymbol{\mu}_{t+1|t}] = E[\boldsymbol{x}_{t+1}] \quad (1.44)$$

Si può, a questo punto, sostituire tale relazione nell'Equazione (1.40) ed eseguire i dovuti raccoglimenti, ottenendo

$$E[\boldsymbol{\mu}_{t+1|t+1}] = (\mathbf{K}'_{t+1} + \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot E[\boldsymbol{x}_{t+1}] + \mathbf{K}_{t+1} \cdot E[\boldsymbol{z}_{t+1}] \quad (1.45)$$

Se ora si impone la condizione che l'errore sulla misura sia distribuito secondo una distribuzione normale con media nulla, l'ultimo termine nella precedente espressione è nullo e si giunge quindi all'equazione seguente:

$$E[\boldsymbol{\mu}_{t+1|t+1}] = (\mathbf{K}'_{t+1} + \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot E[\boldsymbol{x}_{t+1}] \quad (1.46)$$

Imponendo nuovamente la condizione di assenza di bias si ha che necessariamente deve essere

$$\mathbf{K}'_{t+1} + \mathbf{K}_{t+1} \cdot \mathbf{H} = \mathbf{I} \quad (1.47)$$

dove con \mathbf{I} si indica la matrice identità. Richiamando ora l'Equazione (1.39) e sostituendo in essa la relazione appena mostrata, si può scrivere

$$\boldsymbol{\mu}_{t+1|t+1} = (\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot \boldsymbol{\mu}_{t+1|t} + \mathbf{K}_{t+1} \cdot \boldsymbol{y}_{t+1} \quad (1.48)$$

e riordinando i termini

$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + \mathbf{K}_{t+1} \cdot (\boldsymbol{y}_{t+1} - \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t}) \quad (1.49)$$

Il termine \mathbf{K} è detto **Kalman gain** e in un qualche modo dovrà quantificare contemporaneamente il peso attribuito sia alla previsione sia alla misura nell'ottenimento dell'aggiornamento, in quanto l'Equazione (1.47) permette di determinare univocamente una matrice di guadagno una volta individuata l'altra. Grazie al Kalman gain si è ottenuta l'espressione che, a partire dalla conoscenza della t -esima misura, esplicita come, una volta effettuata la $(t + 1)$ -esima misura, si riadatti la media della distribuzione dei valori veri del vettore di stato al tempo $t + 1$; in altre parole si è trovata la legge di aggiornamento della media. Rimane ora da individuare la legge di aggiornamento della varianza.

Per poter ottenere una espressione che sia esplicativa, occorre innanzitutto introdurre una nuova quantità, ossia lo scarto, indicato con il vettore \mathbf{u} e definito nel seguente modo:

$$\mathbf{u}_{t+1|t} = \mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t} \quad (1.50)$$

Ricordando come si era definita la matrice di covarianza, per l'aggiornamento è possibile scrivere:

$$\mathbf{P}_{t+1|t+1} = E[(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t+1}) \cdot (\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t+1})^T | \mathbf{Y}_t] \quad (1.51)$$

In realtà questa formula è molto grezza. Si può migliorarla notando innanzitutto che in questo momento della trattazione, essendo l'attenzione concentrata sull'aggiornamento, si presume di aver effettuato la $(t + 1)$ -esima misura. Quindi è ridondante precisare che il valore atteso è condizionato all'insieme delle misure fino al tempo t , perché se si è effettuata la $(t + 1)$ -esima misura sicuramente si conoscono i valori di tutte le misure precedenti; è dunque possibile più semplicemente scrivere

$$\mathbf{P}_{t+1|t+1} = E[(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t+1}) \cdot (\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1|t+1})^T] \quad (1.52)$$

Sfruttando l'introduzione dello scarto ed eseguendo i dovuti calcoli⁴, non riportati per non appesantire la trattazione, la precedente espressione diventa

$$\begin{aligned} \mathbf{P}_{t+1|t+1} &= (\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot E[\mathbf{u}_{t+1|t} \cdot \mathbf{u}_{t+1|t}^T] \cdot (\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H})^T + \\ &+ \mathbf{K}_{t+1} \cdot E[\mathbf{z}_{t+1} \cdot \mathbf{z}_{t+1}^T] \cdot \mathbf{K}_{t+1}^T + 2(\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot E[\mathbf{u}_{t+1|t} \cdot \mathbf{z}_{t+1}^T] \cdot \mathbf{K}_{t+1}^T \end{aligned} \quad (1.53)$$

Dal momento che, come già esposto in precedenza, si considera un errore distribuito secondo una distribuzione normale con media nulla, sfruttando il fatto che il valore atteso di due variabili aleatorie indipendenti equivale al prodotto dei valori attesi delle due variabili aleatorie ed essendo lo scarto indipendente dall'errore sulla misura, si conviene che

$$E[\mathbf{u}_{t+1|t} \cdot \mathbf{z}_{t+1}^T] = E[\mathbf{u}_{t+1|t}] \cdot E[\mathbf{z}_{t+1}^T] = 0 \quad (1.54)$$

Per riscrivere la precedente relazione in forma più compatta, si può notare innanzitutto che, per come è definita la matrice di covarianza, è lecito scrivere

$$E[\mathbf{u}_{t+1|t} \cdot \mathbf{u}_{t+1|t}^T] = \mathbf{P}_{t+1|t} \quad (1.55)$$

Risulta utile introdurre a questo punto una nuova matrice indicata con \mathbf{R} , che rappresenta la matrice dell'errore sulla misura ed è definita come segue:

$$\mathbf{R}_{t+1} = E[\mathbf{z}_{t+1} \cdot \mathbf{z}_{t+1}^T] \quad (1.56)$$

Si ottiene in questo modo la relazione che mostra l'aggiornamento della covarianza della distribuzione dei valori veri del vettore di stato:

$$\mathbf{P}_{t+1|t+1} = (\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot \mathbf{P}_{t+1|t} \cdot (\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H})^T + \mathbf{K}_{t+1} \cdot \mathbf{R}_{t+1} \cdot \mathbf{K}_{t+1}^T \quad (1.57)$$

Si è quindi esaurita la trattazione riguardo la previsione e l'aggiornamento, che si può riassumere nelle seguenti quattro relazioni chiave:

$$\begin{aligned} \boldsymbol{\mu}_{t+1|t} &= \mathbf{F} \cdot \boldsymbol{\mu}_t \\ \mathbf{P}_{t+1|t} &= \mathbf{F} \cdot \mathbf{P}_{t|t} \cdot \mathbf{F}^T + \mathbf{Q}_t \\ \boldsymbol{\mu}_{t+1|t+1} &= \boldsymbol{\mu}_{t+1|t} + \mathbf{K}_{t+1} \cdot (\mathbf{y}_{t+1} - \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t}) \\ \mathbf{P}_{t+1|t+1} &= (\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot \mathbf{P}_{t+1|t} \cdot (\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H})^T + \mathbf{K}_{t+1} \cdot \mathbf{R}_{t+1} \cdot \mathbf{K}_{t+1}^T \end{aligned} \quad (1.58)$$

⁴Si faccia riferimento a [Reid 2001].

dove le prime due uguaglianze definiscono la previsione rispettivamente della media e della matrice di covarianza della distribuzione di probabilità del $(t+1)$ -esimo vettore di stato quando sono state effettuate t misure, mentre le seconde due definiscono l'aggiornamento rispettivamente della media e della matrice di covarianza della distribuzione di probabilità del $(t+1)$ -esimo vettore di stato una volta effettuata la $(t+1)$ -esima misura. Quello che emerge evidentemente da queste relazioni è che, per il funzionamento del filtro di Kalman, in realtà, non è necessario conoscere tutte le misure prima del tempo t , ma è sufficiente conoscere esclusivamente la misura al tempo corrente e anche in questo si sostanzia la straordinaria potenza dell'algoritmo in esame. Conclusa la trattazione analitica dell'algoritmo, nel prossimo paragrafo ci si concentrerà sull'interpretazione e su una maggiore e più ampia spiegazione del significato del Kalman gain.

1.3 Kalman gain

In questo paragrafo si vuole ottenere una espressione matematica che espliciti il significato qualitativo accennato del Kalman gain; per poter ottenere tale formula, occorre tuttavia prima concentrarsi su una finalità pratica della grandezza in esame, ossia quella di minimizzare gli errori delle stime: l'idea è che con l'aggiornamento si cerca di stimare nel migliore dei modi possibili il valore vero, ossia si tiene conto sia della misura sia della previsione nel tentativo di ottenere un valore quanto più simile possibile a quello vero. In precedenza è stato introdotto il termine di scarto \mathbf{u} , per la cui definizione si rimanda all'Equazione (1.50). Poiché lo scarto consiste proprio nella differenza tra il valore vero della grandezza in esame e la relativa previsione, per ottenere un aggiornamento che sia quanto più simile al valore vero, occorre minimizzare lo scarto. Si può allora introdurre il termine \mathbf{L} , che sarà, tra tutti i possibili termini \mathbf{K}_{t+1} , quello che determina lo scarto minimo e sarà di conseguenza definito nella seguente maniera:

$$\mathbf{L} = \min_{\mathbf{K}_{t+1}} E[\mathbf{u}_{t+1|t+1} \cdot \mathbf{u}_{t+1|t+1}^T | \mathbf{Y}_{t+1}] \quad (1.59)$$

Siccome il prodotto $\mathbf{u}_{t+1|t+1} \cdot \mathbf{u}_{t+1|t+1}^T$ fornisce come risultato una matrice diagonale, essendo il prodotto tensoriale di un vettore colonna per un vettore riga, allora il termine \mathbf{K} che minimizza il valore atteso di tale prodotto sarà anche quello che minimizza la traccia del valore atteso della matrice, dove si indica la funzione che restituisce la traccia di una matrice \mathbf{A} con $tr(\mathbf{A})$. Inoltre si deve tener conto del fatto che le operazioni di traccia e di valore atteso sono tra loro commutative⁵. Allora \mathbf{L} può essere riscritto nella seguente maniera:

$$\mathbf{L} = \min_{\mathbf{K}_{t+1}} tr(E[\mathbf{u}_{t+1|t+1} \cdot \mathbf{u}_{t+1|t+1}^T | \mathbf{Y}_{t+1}]) \quad (1.60)$$

Ricordando la relazione tra la matrice di covarianza e lo scarto, riportata nell'Equazione (1.55), si nota facilmente che vale quanto segue:

$$\mathbf{L} = \min_{\mathbf{K}_{t+1}} tr(\mathbf{P}_{t+1|t+1}) \quad (1.61)$$

Per poter trovare il minimo si deve calcolare la derivata parziale di \mathbf{L} rispetto a \mathbf{K} e porre tale derivata nulla, in modo da trovare il termine \mathbf{L} che effettivamente minimizza

⁵Per una giustificazione di questo fatto si rimanda a [Reid 2001].

lo scarto. Per poter scrivere il risultato, essendo $\mathbf{P}_{t+1|t+1}$ un termine dato dalla somma di due prodotti tra tre matrici, come esplicitato nella quarta relazione dell'Equazione (1.58), si può applicare una regola generale dell'algebra, secondo la quale, data una matrice \mathbf{A} e una matrice simmetrica \mathbf{B} , vale la seguente uguaglianza:

$$\frac{\partial \text{tr}(\mathbf{A} \cdot \mathbf{B} \cdot \mathbf{A}^T)}{\partial \mathbf{A}} = 2\mathbf{A} \cdot \mathbf{B} \quad (1.62)$$

Tenendo in considerazione tale regola, tenendo conto della quarta relazione dell'Equazione (1.58) e ricordando che la derivata parziale di \mathbf{L} rispetto a \mathbf{K} va posta uguale a zero per trovare il minimo, attraverso semplici calcoli per i quali si rimanda a [Reid 2001], si ottiene:

$$\begin{aligned} \frac{\partial \mathbf{L}}{\partial \mathbf{K}_{t+1}} &= \frac{\partial \text{tr}(\mathbf{P}_{t+1|t+1})}{\partial \mathbf{K}_{t+1}} = \\ &= -2(\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}) \cdot \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T + 2\mathbf{K}_{t+1} \cdot \mathbf{R}_{t+1} = 0 \end{aligned} \quad (1.63)$$

Isolando ora il termine \mathbf{K}_{t+1} si perviene alla seguente uguaglianza:

$$\mathbf{K}_{t+1} = \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T \cdot (\mathbf{H} \cdot \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T + \mathbf{R}_{t+1})^{-1} \quad (1.64)$$

Questa è l'espressione del Kalman gain che minimizza il termine di scarto.

L'Equazione (1.64) risulta molto utile per adempiere alle finalità esplicative di questo paragrafo. Per procedere in tal senso, occorre tuttavia prima introdurre un nuovo termine chiamato *innovazione*, indicato con il vettore $\boldsymbol{\nu}$ e definito come la differenza tra la misura al tempo $t + 1$ e la previsione della misura in base alla conoscenza di tutte le misure fino al tempo t . Dunque si tratta di un indicatore che quantifica quanta nuova informazione viene apportata dall'effettuazione di una nuova misura. Denotando con $\tilde{\mathbf{z}}_{t+1|t}$ la previsione della misura al tempo $t + 1$ conoscendo tutte le misure fino al tempo t , si può scrivere

$$\tilde{\mathbf{z}}_{t+1|t} = E[\mathbf{y}_{t+1} | \mathbf{Y}_t] \quad (1.65)$$

e ricordando ora la definizione della misura, in base all'Equazione (1.17), si può migliorare la precedente relazione nel seguente modo:

$$\tilde{\mathbf{z}}_{t+1|t} = E[\mathbf{H} \cdot \mathbf{x}_{t+1} + \mathbf{z}_{t+1} | \mathbf{Y}_t] \quad (1.66)$$

Essendo l'errore distribuito secondo una distribuzione normale con media nulla, si ha

$$\tilde{\mathbf{z}}_{t+1|t} = \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t} \quad (1.67)$$

Dunque si può scrivere l'innovazione nella seguente maniera, che assumiamo come definizione, che si è visto essere coerente con la teoria precedentemente introdotta:

$$\boldsymbol{\nu}_{t+1} = \mathbf{y}_{t+1} - \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t} \quad (1.68)$$

Il vantaggio di aver introdotto l'innovazione è che mentre la sequenza delle misure è formata da termini che possono avere correlazione tra loro, la sequenza delle innovazioni è costituita da termini tra loro non correlati e quindi ortogonali; dal punto di vista statistico le due sequenze sono tuttavia assolutamente equivalenti. Una proprietà importante

dell'innovazione è che essa ha media nulla: lo si può dimostrare partendo dalla sua definizione matematica e applicando il valore atteso condizionato rispetto alle misure fino al tempo t ad ambo i membri.

$$E[\boldsymbol{\nu}_{k+1}|\mathbf{Y}_t] = E[\mathbf{y}_{t+1} - \tilde{\mathbf{z}}_{t+1|t}|\mathbf{Y}_t] \quad (1.69)$$

Grazie alla linearità del valore atteso si ha che

$$E[\boldsymbol{\nu}_{k+1}|\mathbf{Y}_t] = E[\mathbf{y}_{t+1}|\mathbf{Y}_t] - E[\tilde{\mathbf{z}}_{t+1|t}|\mathbf{Y}_t] \quad (1.70)$$

Per quanto sia possibile completare la dimostrazione con passaggi più semplici, si preferisce procedere nel seguente modo, sia per fornire alla trattazione una maggiore consistenza matematica, sia per una più profonda comprensione del risultato che si vuole dimostrare. Sfruttando gli elementi di teoria della misura precedentemente introdotti, si cita una importante proposizione, tratta da [Pascucci 2009], nota come *Lemma della Torre*, secondo la quale, data una σ -algebra \mathcal{F} e una variabile aleatoria $X \in m(\mathcal{F})$, allora

$$E[E[X|\mathcal{F}]|\mathcal{F}] = E[X|\mathcal{F}] \quad (1.71)$$

Nel procedimento per dimostrare che la media dell'innovazione è nulla, poiché, per definizione, il vettore $\tilde{\mathbf{z}}_{t+1|t}$ è già un valore atteso condizionato rispetto a \mathbf{Y}_t , applicando il Lemma della Torre si ottiene che

$$\begin{aligned} E[\tilde{\mathbf{z}}_{t+1|t}|\mathbf{Y}_t] &= E[E[\mathbf{H} \cdot \mathbf{x}_{t+1} + \mathbf{z}_{t+1}|\mathbf{Y}_t]|\mathbf{Y}_t] = \\ &= E[\mathbf{H} \cdot \mathbf{x}_{t+1} + \mathbf{z}_{t+1}|\mathbf{Y}_t] = \tilde{\mathbf{z}}_{t+1|t} \end{aligned} \quad (1.72)$$

Di conseguenza si ha che

$$\begin{aligned} E[\boldsymbol{\nu}_{k+1}|\mathbf{Y}_t] &= E[\mathbf{y}_{t+1}|\mathbf{Y}_t] - E[\tilde{\mathbf{z}}_{t+1|t}|\mathbf{Y}_t] = \\ &= E[\mathbf{y}_{t+1}|\mathbf{Y}_t] - E[\mathbf{H} \cdot \mathbf{x}_{t+1} + \mathbf{z}_{t+1}|\mathbf{Y}_t] = \\ &= E[\mathbf{y}_{t+1}|\mathbf{Y}_t] - E[\mathbf{y}_{t+1}|\mathbf{Y}_t] = 0 \end{aligned} \quad (1.73)$$

dove nell'ultimo passaggio, si è sfruttata la definizione della misura. Si è quindi dimostrata questa proprietà dell'innovazione, che sarà confermata anche dalle simulazioni riportate nel prossimo capitolo. Questa proprietà è ragionevole se si considera che la misura equivale al vettore di stato a meno di un errore distribuito gaussianamente con media nulla; poiché dalla condizione di assenza di bias si ottiene che il valore atteso del vettore di stato coincide con la media della distribuzione prevista e poiché questa coincide con il proprio valore atteso, essendo una costante, allora l'innovazione deve necessariamente essere mediamente nulla. Ciò implica che mediamente il valore della grandezza misurato coincide con quello previsto, cioè mediamente la misura non aggiunge informazione rispetto alla previsione. Si sottolinea che ciò vale *in media* e non sui singoli valori e il Kalman gain serve proprio a quantificare, vedendo il tema da una prospettiva diversa, l'informazione aggiunta dalla misura su ogni singolo punto.

Per una migliore comprensione dell'argomento in esame, si vuole introdurre ora una nuova grandezza, ossia la matrice di covarianza dell'innovazione. In accordo con la trattazione sinora svolta, essa sarà definita nel seguente modo:

$$\mathbf{S}_{t+1} = E[\boldsymbol{\nu}_{t+1} \cdot \boldsymbol{\nu}_{t+1}^T] \quad (1.74)$$

Sostituendo all'innovazione la sua definizione si può scrivere in maniera più esplicita

$$\mathbf{S}_{t+1} = E[(\mathbf{y}_{t+1} - \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t}) \cdot (\mathbf{y}_{t+1} - \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t})^T] \quad (1.75)$$

Esplicitando la definizione della misura e sfruttando l'introduzione del termine $\mathbf{u}_{t+1|t}$ si ottiene

$$\mathbf{S}_{t+1} = E[(\mathbf{z}_{t+1} + \mathbf{H} \cdot \mathbf{u}_{t+1|t}) \cdot (\mathbf{z}_{t+1} - \mathbf{H} \cdot \mathbf{z}_{t+1|t})^T] \quad (1.76)$$

Eseguendo esplicitamente il prodotto tensoriale si ottengono due termini, entrambi scritti, sfruttando la linearità del valore atteso, nella seguente forma a meno di trasposizioni:

$$\mathbf{H} \cdot E[\mathbf{u}_{t+1|t} \cdot \mathbf{z}_{t+1}^T]$$

Per quanto già detto in precedenza

$$\mathbf{H} \cdot E[\mathbf{u}_{t+1|t} \cdot \mathbf{z}_{t+1}^T] = \mathbf{H} \cdot E[\mathbf{u}_{t+1|t}] \cdot E[\mathbf{z}_{t+1}^T] = 0 \quad (1.77)$$

dal momento che l'errore sulla misura lo si suppone distribuito con media nulla. Per cui si ottiene che

$$\mathbf{S}_{t+1} = E[\mathbf{z}_{t+1} \cdot \mathbf{z}_{t+1}^T] + \mathbf{H} \cdot E[\mathbf{u}_{t+1|t} \cdot \mathbf{u}_{t+1|t}^T] \cdot \mathbf{H}^T \quad (1.78)$$

da cui, sfruttando le definizioni riportate nell'Equazione (1.55) e nell'Equazione (1.56), si ha l'espressione della matrice di covarianza dell'innovazione:

$$\mathbf{S}_{t+1} = \mathbf{R}_{t+1} + \mathbf{H} \cdot \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T \quad (1.79)$$

Sostituendo questa uguaglianza nell'Equazione (1.64) si ottiene una forma compatta per indicare il Kalman gain:

$$\mathbf{K}_{t+1} = \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T \cdot \mathbf{S}_{t+1}^{-1} \quad (1.80)$$

Per poter comprendere questa espressione, si deve riprendere in considerazione l'innovazione; l'introduzione di tale grandezza è molto utile perché permette di riscrivere, attraverso delle semplici sostituzioni, l'aggiornamento della media della distribuzione dei vettori di stato nella seguente maniera:

$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + \mathbf{K}_{t+1} \cdot \boldsymbol{\nu}_{t+1} \quad (1.81)$$

Si può osservare come l'espressione dell'aggiornamento della media sia data dal contributo di due termini, ossia la media prevista e il prodotto tra il Kalman gain e l'innovazione. Ciò avviene perché prima viene eseguita al tempo t la previsione della media relativa al tempo $t + 1$, poi si esegue la misura al tempo $t + 1$, con la relativa innovazione che ne quantifica il discostamento dalla previsione; quindi si aggiorna la distribuzione dei valori dei vettori di stato. Il Kalman gain quantifica, in tale aggiornamento, quale sia il contributo della misura effettuata al tempo $t + 1$; più precisamente, costituisce il peso dato alle misurazioni rispetto alla stima dello stato effettuata nell'istante di tempo precedente: se il Kalman gain è elevato, il filtro attribuisce nell'aggiornamento un peso maggiore alle misurazioni effettuate e quindi l'aggiornamento si conforma a esse in modo più reattivo; se invece il Kalman gain è piccolo, il valore aggiornato sarà più prossimo a quello previsto, conformandosi maggiormente alle previsioni del modello. Di conseguenza, si può affermare, il Kalman gain misura la prestazione del filtro in relazione al modello in termini di guadagno percentuale. La ragione per cui il filtro prevede

che si effettui una sorta di media pesata tra la stima e la misura è attribuire maggiore rilevanza ai valori migliori, ossia quelli con una minore incertezza, poiché questi sono da ritenersi più affidabili. Questa media pesata, computata proprio grazie al Kalman gain, nella quale i pesi sono calcolati tramite la covarianza, restituisce dunque come risultato la stima aggiornata dello stato del sistema, che si colloca tra lo stato previsto e quello misurato, con una incertezza minimizzata.

1.4 Trattazione bayesiana

Completata la trattazione analitica anche del Kalman gain, si mostrerà in questo paragrafo, facendo riferimento a [Durbin 2012], come sia possibile ottenere gli stessi risultati ottenuti in precedenza per l'aggiornamento con un approccio sostanzialmente diverso: questa procedura rispecchia una trattazione bayesiana dell'argomento, cioè una trattazione nella quale si presume la conoscenza del *Teorema di Bayes* e si assume che le distribuzioni degli errori e delle previsioni siano distribuzioni normali. Per prima cosa, occorre dunque introdurre il Teorema di Bayes. Siano dati due eventi, A e B . Allora la probabilità che si verifichi A sapendo che si è verificato B è data dalla seguente formula:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \quad (1.82)$$

Alla luce di questa importante legge, si può chiaramente intendere per quale ragione la metodologia con la quale si vuole procedere sia particolarmente utile per l'aggiornamento, al quale ambito del filtro di Kalman si limita la trattazione: il Teorema di Bayes, infatti, permette proprio di sapere come varia la distribuzione di probabilità di una variabile aleatoria sapendo che si è verificato un certo evento. Siano dati ora tre eventi, A , B e C . Estendendo la formula precedente, si può dire che la probabilità che si verifichi l'evento A sapendo che si sono verificati gli eventi B e C è data dalla seguente formula:

$$p(A|B, C) = \frac{p(B|A, C)p(A|C)}{p(B|C)} \quad (1.83)$$

Si supponga a questo punto di conoscere le misure fino al tempo $t - 1$. Allora è possibile scrivere la seguente relazione:

$$p(\mathbf{x}_t|\mathbf{y}_t, \mathbf{Y}_{t-1}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{Y}_{t-1})p(\mathbf{x}_t|\mathbf{Y}_{t-1})}{p(\mathbf{y}_t|\mathbf{Y}_{t-1})} \quad (1.84)$$

Notando che

$$p(\mathbf{x}_t|\mathbf{y}_t, \mathbf{Y}_{t-1}) = p(\mathbf{x}_t|\mathbf{Y}_t) \quad (1.85)$$

l'idea è quella di scrivere la distribuzione dell'aggiornamento, ossia esplicitare la forma di $p(\mathbf{x}_t|\mathbf{Y}_t)$ a partire dalle tre distribuzioni in funzione delle quali quest'ultima è stata scritta grazie al Teorema di Bayes, sfruttando l'ipotesi di lavoro, cioè che le previsioni abbiano una distribuzione normale. Seguendo l'Equazione (1.17) e considerando, coerentemente alla trattazione sinora svolta, il vettore \mathbf{z}_t come errore della previsione sulla misura \mathbf{y}_t , si ha

$$E[\mathbf{y}_t|\mathbf{Y}_{t-1}] = \mathbf{H} \cdot \boldsymbol{\mu}_{t|t-1} \quad (1.86)$$

Si consideri ora, per semplificare la trattazione, che non si abbia un vettore di stato, ma una semplice grandezza scalare, in modo da poter sfruttare le relazioni valide per un local level model. Al termine della spiegazione ci si ricondurrà al caso generale. La precedente relazione la si può allora più semplicemente riscrivere come

$$E[y_t|\mathbf{Y}_{t-1}] = \mu_{t|t-1} \quad (1.87)$$

Denotando con V la funzione che restituisce la varianza e sostituendo alla misura la relativa espressione mostrata nell'Equazione (1.5), si ha che

$$V[y_t|\mathbf{Y}_{t-1}] = V[\varepsilon_t + x_t|\mathbf{Y}_{t-1}] \quad (1.88)$$

Sfruttando l'indipendenza dell'errore sulla misura dal valore vero, si può dimostrare⁶ che

$$V[y_t|\mathbf{Y}_{t-1}] = V[\varepsilon_t|\mathbf{Y}_{t-1}] + V[x_t|\mathbf{Y}_{t-1}] \quad (1.89)$$

Poiché l'errore sulla t -esima misura è indipendente dalle misure effettuate precedentemente, allora il condizionamento su di esso non ha effetto, per cui si ha

$$V[y_t|\mathbf{Y}_{t-1}] = V[\varepsilon_t] + V[x_t|\mathbf{Y}_{t-1}] \quad (1.90)$$

e ricorrendo alle notazioni precedentemente introdotte si può scrivere

$$V[y_t|\mathbf{Y}_{t-1}] = \sigma_\varepsilon^2 + P_{t|t-1} \quad (1.91)$$

Di conseguenza

$$p(y_t|\mathbf{Y}_{t-1}) = N(\mu_{t|t-1}; P_{t|t-1} + \sigma_\varepsilon^2) \quad (1.92)$$

Per quanto riguarda $p(y_t|x_t, \mathbf{Y}_{t-1})$, i ragionamenti sono analoghi. Si ha infatti

$$E[y_t|x_t, \mathbf{Y}_{t-1}] = \mu_{t|t-1} \quad (1.93)$$

proprio perché l'errore sulla misura ha una media nulla e $E[x_t|\mathbf{Y}_{t-1}] = \mu_{t|t-1}$. Inoltre, qualora si conosca il valore vero x_t , la varianza della distribuzione $p(y_t|x_t, \mathbf{Y}_{t-1})$ è la varianza della sola misura e quindi si ha che

$$V[y_t|x_t, \mathbf{Y}_{t-1}] = \sigma_\varepsilon^2 \quad (1.94)$$

Dunque si può concludere che

$$p(y_t|x_t, \mathbf{Y}_{t-1}) = N(\mu_{t|t-1}; \sigma_\varepsilon^2) \quad (1.95)$$

Rimane infine da considerare l'ultima distribuzione presente nella relazione descritta nell'Equazione (1.84), ossia $p(x_t|\mathbf{Y}_{t-1})$. Per quanto detto in precedenza,

$$E[x_t|\mathbf{Y}_{t-1}] = \mu_{t|t-1} \quad (1.96)$$

e

$$V[x_t|\mathbf{Y}_{t-1}] = P_{t|t-1} \quad (1.97)$$

⁶Per una dimostrazione precisa si rimanda a [Durbin 2012].

Dunque

$$p(x_t|\mathbf{Y}_{t-1}) = N(\mu_{t|t-1}; P_{t|t-1}) \quad (1.98)$$

L'Equazione (1.84) può quindi essere riscritta nella seguente maniera:

$$p(x_t|y_t, \mathbf{Y}_{t-1}) = k_0 \exp\left(-\frac{(x_t - \mu_{t|t-1})^2}{2P_{t|t-1}}\right) \exp\left(-\frac{(x_t - \mu_{t|t-1})^2}{2\sigma_\varepsilon^2}\right) \exp\left(\frac{(y_t - \mu_{t|t-1})^2}{2(P_{t|t-1} + \sigma_\varepsilon^2)}\right) \quad (1.99)$$

dove k_0 è la costante di normalizzazione. In forma più leggera, definendo

$$B = \frac{(x_t - \mu_{t|t-1})^2}{P_{t|t-1}} + \frac{(x_t - \mu_{t|t-1})^2}{\sigma_\varepsilon^2} - \frac{(y_t - \mu_{t|t-1})^2}{P_{t|t-1} + \sigma_\varepsilon^2} \quad (1.100)$$

si ha

$$p(x_t|y_t, \mathbf{Y}_{t-1}) = k_0 \cdot e^{-\frac{B}{2}} \quad (1.101)$$

Per ottenere i risultati desiderati si possono riordinare i vari termini nell'espressione di B :

$$B = \frac{P_{t|t-1} + \sigma_\varepsilon^2}{\sigma_\varepsilon^2 P_{t|t-1}} (x_t - \mu_{t|t-1} - \frac{P_{t|t-1}}{P_{t|t-1} + \sigma_\varepsilon^2} (y_t - \mu_{t|t-1})) \quad (1.102)$$

Se ora, secondo una notazione già adottata in precedenza, si indica, in riferimento alla distribuzione $p(x_t|y_t, \mathbf{Y}_{t-1})$, con $\mu_{t|t}$ la media e con $P_{t|t}$ la varianza di tale distribuzione, si ha che, alla luce della precedente relazione,

$$\begin{aligned} \mu_{t|t} &= \mu_{t|t-1} + \frac{P_{t|t-1}}{P_{t|t-1} + \sigma_\varepsilon^2} (y_t - \mu_{t|t-1}) \\ P_{t|t} &= \frac{P_{t|t-1} \sigma_\varepsilon^2}{P_{t|t-1} + \sigma_\varepsilon^2} \end{aligned} \quad (1.103)$$

Per ricondursi al caso generale si inizierà dalla prima relazione dell'Equazione (1.103). Apportando le seguenti sostituzioni:

$$\begin{aligned} \mu_{t|t} &\rightarrow \boldsymbol{\mu}_{t|t} \\ \mu_{t|t-1} &\rightarrow \boldsymbol{\mu}_{t|t-1} \\ \frac{P_{t|t-1}}{P_{t|t-1} + \sigma_\varepsilon^2} &\rightarrow \mathbf{K}_t \\ (y_t - \mu_{t|t-1}) &\rightarrow (\mathbf{y}_t - \mathbf{H} \cdot \boldsymbol{\mu}_{t|t-1}) \end{aligned} \quad (1.104)$$

si ottiene una relazione equivalente alla terza riportata nell'Equazione (1.58). Le prime due sostituzioni dell'Equazione (1.104) sono banali, in quanto semplicemente si estende il valore atteso della distribuzione dal caso scalare al caso vettoriale. Per quanto riguarda la quarta sostituzione, l'unica precisazione che occorre fare è che in questo caso si è inserito anche il termine \mathbf{H} , giacché si sta confrontando la media della distribuzione dei vettori di stato con il vettore delle misure, che potrebbero avere dimensioni diverse e la matrice \mathbf{H} serve per l'appunto a rendere confrontabili queste due grandezze potenzialmente diverse. Per quanto riguarda la terza sostituzione, occorre far riferimento all'Equazione (1.64): la varianza $P_{t|t-1}$, nel caso in cui si consideri la distribuzione di un vettore anziché di uno scalare, sarà sostituita dalla matrice di covarianza $\mathbf{P}_{t|t-1}$; si dovrà poi considerare

la matrice \mathbf{R}_t anziché la varianza dell'errore scalare sulla misura σ_ε^2 ; inserendo nell'espressione opportunamente le matrici \mathbf{H} per rendere confrontabili le diverse matrici che rappresentano la varianza della distribuzione di grandezze potenzialmente diverse, ossia rispettivamente la matrice di covarianza prevista $\mathbf{P}_{t|t-1}$ e la matrice \mathbf{R} dell'errore sulla misura, ci si riconduce alla forma desiderata. Ripetendo operazioni analoghe ci si riconduce agevolmente anche dalla seconda relazione dell'Equazione (1.103) alla quarta dell'Equazione (1.58) e dunque al caso generale.

Si è quindi visto come sia possibile ottenere le equazioni del filtro di Kalman attraverso un approccio bayesiano, partendo dal teorema più importante nell'ambito della probabilità condizionata. I passaggi esposti in questa sezione sono molto interessanti, in quanto permettono di capire più a fondo l'argomento in esame. In particolare permettono di giustificare un importante risultato che verrà meglio mostrato nel prossimo capitolo, ossia che le matrici di covarianza prevista e aggiornata tendono a stabilizzarsi su valori limite e ciò accade proprio per la natura bayesiana del filtro: quando vengono effettuate le misure tramite un processo di inferenza bayesiana si ristabilizza l'errore sulla previsione e di conseguenza anche sull'aggiornamento, dal momento che, qualora l'errore sulla misura fosse molto grande, al limite si attribuirebbe grazie al Kalman gain maggiore peso alle previsioni; questo nell'algoritmo viene svolto dal termine $\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}$, che svolge, come nel seguito si mostrerà, una funzione di reset dell'errore sulla previsione. Conclusa quindi la prima parte di introduzione teorica, si procede nella trattazione mostrando i risultati ottenuti dalla realizzazione di simulazioni in cui sono state applicate le formule sinora ricavate.

Capitolo 2

Simulazioni

Si mostreranno ora i risultati di alcune simulazioni eseguite al fine di testare l'utilità, l'importanza e l'efficacia del filtro di Kalman. Considerando quindi un punto materiale che si muove secondo determinate leggi del moto, lo scopo di questo capitolo è mostrare come applicare il filtro di Kalman per l'aggiornamento e la previsione dei valori veri della posizione del punto materiale in seguito a delle misure, prendendo spunto dagli esempi riportati in [Reid 2001].

Si anticipano due premesse importanti di carattere tecnico, che varranno per tutte le simulazioni. Poiché si suppone di conoscere con certezza le condizioni iniziali, per cui è come se fossero distribuite secondo una delta di Dirac, cioè secondo una distribuzione di probabilità con varianza nulla e si assume che esse non abbiano correlazione, cioè siano totalmente indipendenti le une dalle altre, si pone la matrice di covarianza iniziale come matrice nulla, che evolve secondo le leggi riportate nell'Equazione (1.58). La seconda premessa riguarda la matrice \mathbf{R} , ovvero la matrice dell'errore sulla misura. Nelle simulazioni essa verrà supposta costante e assume la seguente espressione, in accordo con la definizione riportata nell'Equazione (1.56):

$$(\mathbf{R})_{ij} = \begin{cases} \sigma_{\varepsilon_i}^2 & \text{per } i = j \\ 0 & \text{per } i \neq j \end{cases} \quad (2.1)$$

dove $\sigma_{\varepsilon_i}^2$ è la distribuzione dell'errore, che si ricorda essere distribuito gaussianamente con media nulla, della i -esima componente misurata. D'altro canto si suppone che le misure siano indipendenti le une dalle altre, per cui la correlazione tra le varie misure è nulla, dunque l'errore sulla misura è dato dall'errore sulla singola misura, per cui gli unici termini non nulli sono quelli sulla diagonale principale.

2.1 Moto rettilineo uniforme

In questa prima sezione si tratterà il caso più semplice, ovvero il caso di un punto che si muove a velocità costante, secondo la seguente legge, dove con \mathbf{q} si indicano le coordinate spaziali e con \mathbf{v} la velocità:

$$\mathbf{q}(t) = \mathbf{q}(0) + \mathbf{v}t, \quad (2.2)$$

Il vettore di stato \mathbf{x} è un vettore costituito dalla posizione e dalla velocità, quindi si può scrivere tale vettore come segue:

$$\mathbf{x} = \begin{pmatrix} \mathbf{q}(t) \\ \mathbf{v} \end{pmatrix} \quad (2.3)$$

Per semplicità si considererà nel seguito il caso di un moto unidimensionale. Per descrivere l'evoluzione del vettore di stato nel tempo si adopera una matrice chiamata *flusso di fase* e denotata con \mathbf{A} . In particolare, essendo la velocità costante e dovendosi trasformare la posizione secondo l'Equazione (2.2), si avrà che

$$\begin{pmatrix} \frac{dq(t)}{dt} \\ \frac{dv(t)}{dt} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} q(t) \\ v(t) \end{pmatrix} \quad (2.4)$$

Si supponga di considerare, conformemente alla teoria esposta nel precedente capitolo, anche un rumore browniano sul valore vero della grandezza in esame, che nel caso fisico potrebbe essere per esempio costituito dal vento; si indichi tale termine, come già in precedenza, con il vettore $\mathbf{w}(t)$, il cui primo elemento sarà il rumore sulla posizione e il secondo il rumore sulla velocità. Essendo, sulla base di quanto detto,

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (2.5)$$

si ottiene la seguente relazione:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A} \cdot \mathbf{x}(t) + \mathbf{w}(t) \quad (2.6)$$

Per quanto riguarda le misure, tenendo conto della relativa espressione, di seguito riportata

$$\mathbf{y}(t) = \mathbf{H} \cdot \mathbf{x}(t) + \mathbf{z}(t) \quad (2.7)$$

occorre identificare la matrice \mathbf{H} . In questa simulazione, come in quelle seguenti, si suppone di effettuare solo misure relative alla posizione; quindi y_t è uno scalare con le dimensioni di una lunghezza e di conseguenza si avrà che

$$\mathbf{H} = (1, \ 0) \quad (2.8)$$

Chiaramente, alla luce di quanto detto, anche $z(t)$ sarà non un vettore, bensì uno scalare. Inoltre, poiché viene misurata solamente la posizione, la matrice \mathbf{R} è in realtà uno scalare coincidente con la varianza della distribuzione dell'errore sulla misura della posizione, proprio perché quest'ultimo è uno scalare. Per verificare di aver definito correttamente gli elementi sinora introdotti, si può dapprima osservare l'Equazione (1.79). Essendo la matrice \mathbf{H} un vettore riga bidimensionale, si ottiene che la matrice di covarianza dell'innovazione è uno scalare: moltiplicando infatti a destra una matrice 2×2 per il trasposto di un vettore riga bidimensionale si ottiene un vettore colonna bidimensionale, che moltiplicato a sinistra per un vettore riga bidimensionale restituisce uno scalare. Guardando poi all'Equazione (1.80), il prodotto tra una matrice 2×2 e il trasposto di un vettore riga bidimensionale è un vettore colonna bidimensionale, per cui il Kalman gain è una matrice 2×1 , in accordo con quanto affermato nel precedente capitolo.

Risulta ora necessario discretizzare le formule relative al modello in esame, poiché le misure vengono prese in intervalli discreti Δt . L'Equazione (2.7) non causa problemi, giacché misura, errore sulla misura e vettore di stato sono tutti riferiti allo stesso istante di tempo, quello in cui la misura viene effettuata. Maggiori complicazioni si hanno per l'Equazione (2.6). Si trascuri per il momento il rumore sul valore vero della grandezza. Discretizzando la relazione si ottiene

$$\frac{\Delta \mathbf{x}}{\Delta t} = \mathbf{A} \cdot \mathbf{x} \quad (2.9)$$

e moltiplicando ambo i membri per Δt

$$\Delta \mathbf{x} = (\mathbf{A}\Delta t) \cdot \mathbf{x} \quad (2.10)$$

che è equivalente a

$$\mathbf{x}_{k+1} = \mathbf{x}_k + (\mathbf{A}\Delta t) \cdot \mathbf{x}_k \quad (2.11)$$

da cui, ponendo $\mathbf{F} = \mathbf{I} + \mathbf{A}\Delta t$, ossia

$$\mathbf{F} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \quad (2.12)$$

si ottiene

$$\mathbf{x}_{k+1} = \mathbf{F} \cdot \mathbf{x}_k \quad (2.13)$$

La matrice \mathbf{F} , in accordo con tutta la teoria introdotta precedentemente, si configura proprio come l'operatore che determina la relazione tra uno stato e il successivo. Considerando che il rumore sul valore vero è riferito al k -esimo istante, si ottiene la coppia di relazioni che definisce l'evoluzione nel tempo del sistema e delle misure:

$$\begin{cases} \mathbf{x}_{k+1} = \begin{pmatrix} 1 & \Delta t \\ 0 & 1 \end{pmatrix} \cdot \mathbf{x}_k + \mathbf{w}_k \\ y_k = (1, 0) \cdot \mathbf{x}_k + z_k \end{cases} \quad (2.14)$$

Le leggi riassunte nell'Equazione (1.58) relative al filtro di Kalman per la previsione e l'aggiornamento della posizione sono state quindi applicate a serie di valori generati secondo il sistema presentato nell'Equazione (2.14), tenendo conto degli accorgimenti già ampiamente esplicitati. Tuttavia, per procedere a discutere della simulazione vera e propria, occorre prima trattare gli errori. A tal proposito, come anticipato nelle premesse al presente capitolo, la matrice di covarianza iniziale della distribuzione dei vettori di stato è nulla, per cui rimane da trovare solo la distribuzione del rumore browniano sui vettori di stato. Nel precedente capitolo si è definito \mathbf{w}_t come un vettore distribuito secondo una distribuzione gaussiana con media nulla e con matrice di covarianza \mathbf{Q}_t , per cui

$$\mathbf{w}_t \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{Q}_t \right) \quad (2.15)$$

Per trovare l'espressione di \mathbf{Q}_t occorre partire dalla definizione:

$$\mathbf{Q}_t = E[\mathbf{w}_t \cdot \mathbf{w}_t^T] \quad (2.16)$$

Grazie alla teoria relativa alle trasformate di Laplace, come dimostrato in [Reid 2001], data una perturbazione scalare $w(t)$ avente le dimensioni di una lunghezza, continua rispetto al tempo e distribuita gaussianamente con media nulla, è possibile dare la seguente rappresentazione del rumore browniano:

$$\mathbf{w}(t) = \int_0^{\Delta t} \begin{pmatrix} \Delta t - t' \\ 1 \end{pmatrix} w(t+t') dt' \quad (2.17)$$

dove si tratta la variabile di integrazione dt' come una grandezza adimensionale, alla quale si associa tuttavia il significato fisico del tempo. Sostituendo la precedente espressione nell'Equazione (2.16) si ottiene

$$\begin{aligned} \mathbf{Q}_t = E[& \left(\int_0^{\Delta t} \begin{pmatrix} \Delta t - t' \\ 1 \end{pmatrix} w(t+t') dt' \right) \cdot \\ & \cdot \left(\int_0^{\Delta t} \begin{pmatrix} \Delta t - t'' \\ 1 \end{pmatrix} w(t+t'') dt'' \right)^T] \end{aligned} \quad (2.18)$$

Poiché dt' e dt'' sono variabili di integrazione tra loro indipendenti, allora il prodotto dei due integrali equivale all'integrale doppio, per cui si ottiene

$$\begin{aligned} \mathbf{Q}_t = E[& \int_0^{\Delta t} \int_0^{\Delta t} \left(\begin{pmatrix} \Delta t - t' \\ 1 \end{pmatrix} w(t+t') \right) \cdot \\ & \cdot \left(\begin{pmatrix} \Delta t - t'' \\ 1 \end{pmatrix} w(t+t'') \right)^T dt' dt''] \end{aligned} \quad (2.19)$$

Siccome le perturbazioni $w(t+t')$ e $w(t+t'')$ sono scalari, possono essere estratti dall'operazione di trasposizione e di prodotto tensoriale, per cui si ha

$$\mathbf{Q}_t = E \left[\int_0^{\Delta t} \int_0^{\Delta t} \begin{pmatrix} \Delta t - t' \\ 1 \end{pmatrix} \cdot (\Delta t - t'', 1) w(t+t') w(t+t'') dt' dt'' \right] \quad (2.20)$$

Considerando ora uno spazio di probabilità $(\Omega, \mathcal{H}, \mathbb{P})$, poiché il valore atteso è definito come un integrale sullo spazio campionario e poiché $w(t)$ è una variabile aleatoria continua rispetto al tempo, quindi misurabile rispetto alla σ -algebra prodotto $\mathcal{H} \times \mathcal{B}(t)$, dove con $\mathcal{B}(t)$ si indica la σ -algebra dei boreliani introdotta nell'insieme degli istanti di tempo, allora si può applicare il Teorema di Fubini, il quale afferma che, dati due spazi misurabili (X, \mathcal{F}) e (Y, \mathcal{G}) e data una funzione $f \in m(\mathcal{F} \times \mathcal{G})$, vale la seguente uguaglianza, ossia l'ordine di integrazione non è importante:

$$\int_X \left(\int_Y f(x, y) dy \right) dx = \int_Y \left(\int_X f(x, y) dx \right) dy \quad (2.21)$$

Dall'applicazione del Teorema di Fubini all'espressione della matrice di covarianza sul rumore browniano si ottiene che

$$\mathbf{Q}_t = \int_0^{\Delta t} \int_0^{\Delta t} E \left[\begin{pmatrix} \Delta t - t' \\ 1 \end{pmatrix} \cdot (\Delta t - t'', 1) w(t+t') w(t+t'') \right] dt' dt'' \quad (2.22)$$

Dal momento che i vettori che compaiono nell'espressione della matrice di covarianza del rumore browniano sono delle costanti rispetto al valore atteso, per la linearità di quest'ultimo si può scrivere

$$\mathbf{Q}_t = \int_0^{\Delta t} \int_0^{\Delta t} \begin{pmatrix} \Delta t - t' \\ 1 \end{pmatrix} \cdot (\Delta t - t'', 1) E[w(t+t')w(t+t'')] dt' dt'' \quad (2.23)$$

Per l'indipendenza delle variabili t' e t'' si può definire il valore atteso presente nell'espressione di \mathbf{Q}_t nella seguente maniera:

$$E[w(t+t')w(t+t'')] = \lambda \delta(t' - t'') \quad (2.24)$$

dove con δ si indica la distribuzione di Dirac e λ è una costante che quantifica la varianza della perturbazione. Sostituendo la rappresentazione data per il valore atteso nell'Equazione (2.23) e svolgendo il prodotto tensoriale tra i due vettori si ottiene la seguente uguaglianza:

$$\mathbf{Q}_t = \int_0^{\Delta t} \begin{pmatrix} (\Delta t - t')^2 & (\Delta t - t') \\ (\Delta t - t') & 1 \end{pmatrix} \lambda dt' \quad (2.25)$$

e risolvendo l'integrale si ottiene l'espressione dell'errore sulla matrice di covarianza utilizzata nella simulazione, dove si rimuove l'indicizzazione temporale in quanto la matrice risultante è una costante:

$$\mathbf{Q} = \lambda \begin{pmatrix} \Delta t^3/3 & \Delta t^2/2 \\ \Delta t^2/2 & \Delta t \end{pmatrix} \quad (2.26)$$

Di conseguenza, indicando con σ_q^2 la varianza del rumore browniano sulla posizione e con σ_v^2 la varianza del rumore browniano sulla velocità, si ha che

$$\sigma_q^2 = \frac{\lambda \Delta t^3}{3} \quad \sigma_v^2 = \lambda \Delta t \quad (2.27)$$

per cui, indicando con w_q il rumore browniano sulla posizione e con w_v il rumore sulla velocità, vale che

$$w_q \sim N\left(0; \frac{\lambda \Delta t^3}{3}\right) \quad w_v \sim N(0; \lambda \Delta t) \quad (2.28)$$

Sono stati dunque trattati analiticamente tutti gli elementi necessari per poter discutere dell'applicazione del filtro di Kalman al caso in esame. In Figura (2.1) si può osservare il grafico che mostra l'evoluzione temporale del sistema, ottenuto tramite una simulazione computazionale in Python, della quale si riporta in Appendice il codice. In particolare, i punti che costituiscono le posizioni reali sono stati ottenuti sommando ai valori corrispondenti a un moto rettilineo uniforme, determinati dalle relazioni riportate nell'Equazione (2.14), un rumore generato casualmente secondo una distribuzione conforme a quella riportata nell'Equazione (2.27), ponendo $\lambda = 30 \text{ m}^2$. Per ottenere i valori misurati si è sommato al valore reale un errore di misura generato casualmente secondo una distribuzione $N(0; \sigma_\varepsilon^2)$, dove si è posto come valore della deviazione standard $\sigma_\varepsilon = 2 \text{ m}$. Le posizioni previste e aggiornate sono state ottenute applicando le relazioni esplicitate nell'Equazione (1.58). Come valore della velocità si è posto $v = 5 \text{ m s}^{-1}$ e come intervallo temporale $\Delta t = 0.2$, dove si ricorda che gli intervalli temporali vengono per semplicità trattati come grandezze adimensionali, in modo che tutti i termini della matrice \mathbf{Q} abbiano la stessa unità di misura.

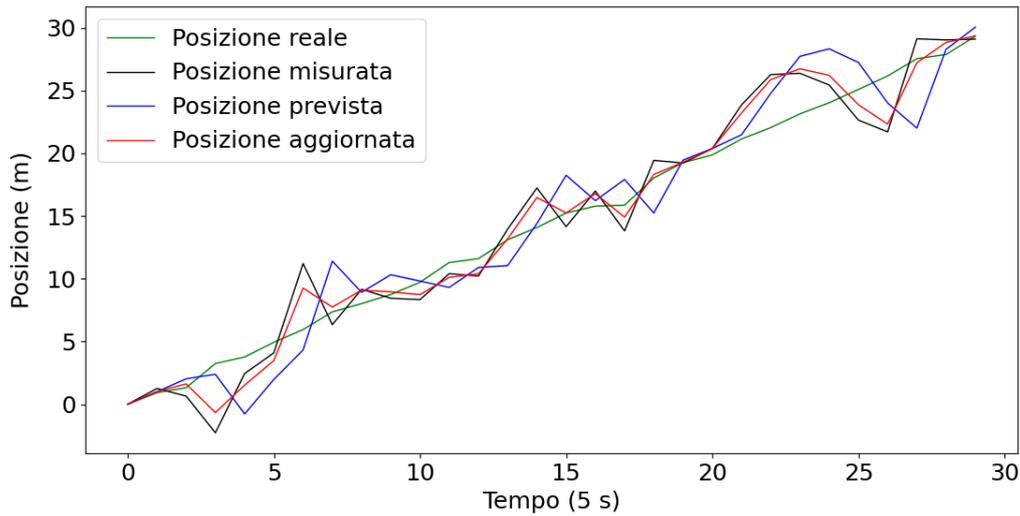


Figura (2.1): La figura rappresenta l'andamento nel tempo della posizione reale, misurata, prevista e aggiornata. Della simulazione, effettuata con 100 iterazioni, si sono mostrati in questo grafico i primi 30 punti di ogni serie per avere una buona visualizzazione di ciò che accade.

Dal grafico si constata come la posizione reale abbia un andamento al primo ordine lineare, esattamente come ci si attende, presentando delle leggere fluttuazioni per effetto del rumore browniano, anche se queste sono poco percettibili dal momento che $\sigma_q^2 \sim \Delta t^3$ e si è scelto per la simulazione un Δt piccolo. La posizione misurata si distacca dalla posizione reale sulla base del valore che assume l'errore sulla misura. La posizione prevista e la posizione aggiornata riportate nel grafico equivalgono ai valori attesi, o equivalentemente ai valori medi, delle relative distribuzioni. Si osserva come la posizione prevista si muova ipotizzando che il trend seguito in un intervallo temporale dalla posizione misurata prosegua anche nell'intervallo successivo, ricalibrando poi la traiettoria in base alle deviazioni che subisce la posizione aggiornata. Questo è esattamente quello che ci si aspetta osservando la prima relazione dell'Equazione (1.58): la posizione prevista è come se provasse a prolungare, secondo un modello di moto rettilineo uniforme, il tratto percorso dalla posizione aggiornata all'istante in cui la previsione viene effettuata. Per una approfondita comprensione di ciò che accade invece per la posizione aggiornata, occorre prima soffermarsi su alcuni importanti aspetti del Kalman gain, deducibili dall'immagine riportata in Figura (2.2).

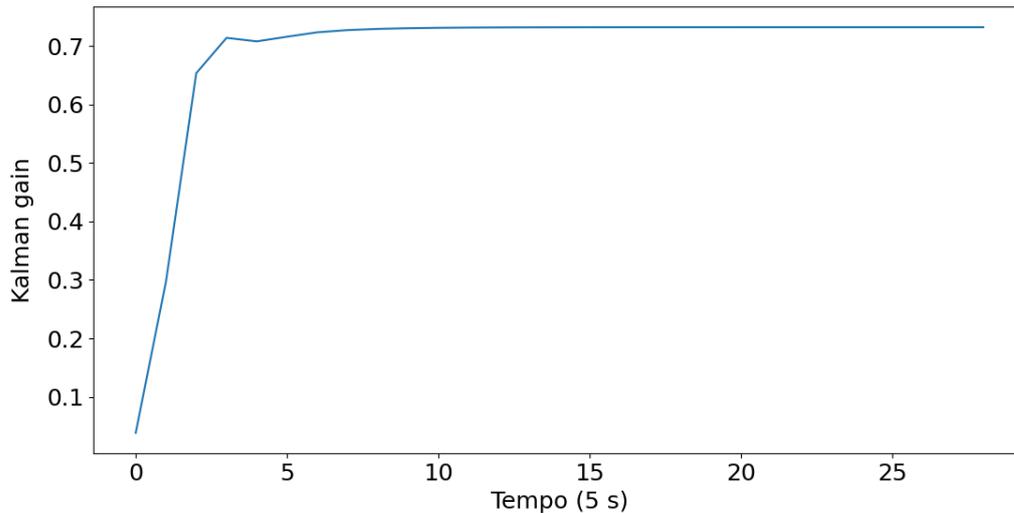


Figura (2.2): L'immagine rappresenta l'evoluzione temporale del Kalman gain. Si vede che tende molto velocemente a un valore limite. Della simulazione sono stati riportati nel grafico solamente i primi 30 punti della serie per osservare meglio il comportamento in tempi prossimi all'istante iniziale.

Per correttezza occorre innanzitutto precisare che nel grafico appena mostrato è riportato, essendo il Kalman gain un vettore bidimensionale, l'andamento temporale della sua prima componente, quella relativa alla posizione. Si osserva dapprima che il valore che essa assume all'inizio è 0, esattamente come ci si attende: infatti, essendo la matrice di covarianza iniziale nulla e l'errore sulla misura costante e in particolare sempre diverso da zero, inizialmente si attribuisce peso nullo alla misura; si vede anche che la prima componente del Kalman gain tende, già dopo poche ripetizioni, a un valore limite ≈ 0.7 . Che cosa vuol dire ciò? Indicando con $\mathbf{K}_{t+1}^{(1)}$ la prima componente del filtro di Kalman, si ha che, alla luce di quanto riportato nell'Equazione (1.79) e nell'Equazione (1.80),

$$\mathbf{K}_{t+1}^{(1)} = \frac{\mathbf{H} \cdot \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T}{\mathbf{H} \cdot \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T + \sigma_\varepsilon^2} \quad (2.29)$$

Quindi tale componente quantifica il rapporto tra l'errore sulla posizione dato dal rumore browniano e l'errore totale sulla posizione, dato dalla somma dell'errore browniano e dell'errore sulla misura. Se $\mathbf{K}_t^{(1)} \rightarrow (\approx 0.7)$, allora vuol dire che l'errore associato alla previsione costituisce circa il 70% dell'errore totale, ma vuol dire anche, sulla base di alcune considerazioni precedentemente riportate che esplicitano il significato dell'Equazione (1.48), che nell'elaborazione della posizione aggiornata si tiene conto per circa il 70% della misura e di conseguenza per circa il 30% della previsione: ciò implica che, punto per punto, la posizione aggiornata deve essere compresa tra la posizione aggiornata e la posizione misurata, più vicina a quest'ultima in quanto ritenuta più affidabile rispetto alla previsione, come per altro è visibile dall'Figura (2.2) e meglio dalla Figura (2.3), dove si riporta uno zoom del valore delle posizioni tra le cinquantesima e la sessantesima iterazione.

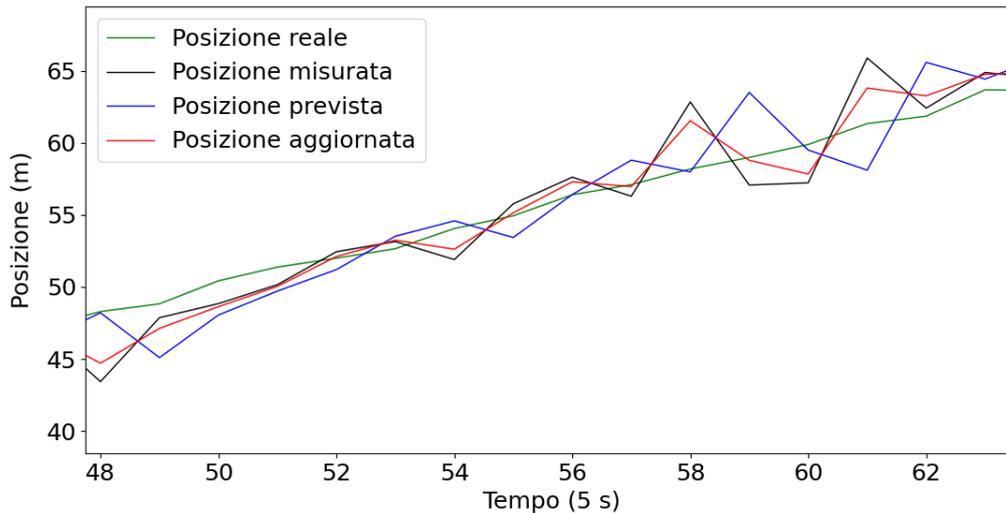


Figura (2.3): L'immagine permette di visualizzare più nel dettaglio come si comportano le varie serie: in particolare si osserva che il valore aggiornato si trova sempre tra quello previsto e quello misurato, nel tentativo di stimare il meglio possibile il valore vero.

Il comportamento che assume la posizione aggiornata, in generale, varia a seconda del valore limite assunto dal filtro di Kalman, il quale dipende da numerosi fattori. In primis è influenzato dal Δt : infatti, osservando l'espressione della matrice \mathbf{Q} data dall'Equazione (2.26), si nota chiaramente che all'aumentare del valore di Δt aumenta il valore dell'errore sulla matrice di covarianza e di conseguenza il valore della matrice di covarianza prevista, come si evince dalla seconda seconda relazione dell'Equazione (1.58); questo, alla luce dell'Equazione (2.29), comporta un aumento della prima componente del Kalman gain, che tende a 1 per valori molto elevati della matrice di covarianza prevista. Ciò ha anche una spiegazione più qualitativa: chiaramente il valore di \mathbf{Q} aumenta al crescere di Δt perché più ampi sono gli intervalli di tempo, più difficile diventa effettuare delle previsioni accurate; di conseguenza la distribuzione prevista viene delineata con un margine d'errore molto maggiore. Chiaramente, poiché la prima componente del Kalman gain quantifica il peso attribuito alla misura, come si evince dall'Equazione (1.48), in maniera ragionevole all'aumentare quindi di Δt diminuisce l'affidabilità attribuita dall' algoritmo alla previsione rispetto alla misura. D'altronde, in maniera più intuitiva e comprensibile direttamente dall'Equazione (2.29), la prima componente del Kalman gain aumenta all'aumentare della matrice di covarianza prevista, giacché in tal caso vuol dire che la posizione prevista ha una incertezza crescente, ritenendo quindi sempre più affidabile la misura, mentre diminuisce all'aumentare di σ_ϵ^2 , che corrisponde proprio alla varianza della distribuzione dell'errore sulla misura, la quale quindi sarebbe sempre meno affidabile. Ma perché il Kalman gain si stabilizza su un valore limite? Chiaramente all'inizio la matrice di covarianza prevista avrà valori crescenti, perché vi si somma sempre la matrice \mathbf{Q} e questo determina una crescita dei valori del Kalman gain. Tuttavia, osservando la quarta relazione dell'Equazione (1.58), si vede che compare il termine $\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}$, il quale svolge una funzione di "reset": di fatto si moltiplica la matrice di covarianza prevista per il peso attribuito alla previsione dal Kalman gain e si moltiplica la matrice \mathbf{R}_{t+1} per il peso attribuito dal Kalman gain alla misura; in

altre parole, la matrice di covarianza aggiornata è una sorta di media pesata degli errori sulla base dell'incertezza valutata tramite il Kalman gain. Al crescere del valore di quest'ultimo, diminuisce il peso della matrice di covarianza prevista e aumenta quello della matrice dell'errore sulla misura, che si sa però essere costante. A partire dalla matrice di covarianza aggiornata si calcola la matrice di covarianza prevista, della quale tuttavia il contributo alla matrice di covarianza aggiornata è sempre più ridotto dal termine $\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}$, che chiaramente decresce all'aumentare del Kalman gain: il risultato è che $\mathbf{P}_{t+1|t+1}$ si stabilizza su un valore limite e poiché da essa si calcola $\mathbf{P}_{t+1|t}$, anche quest'ultima si stabilizza su un valore limite, essendo le matrici \mathbf{F} e \mathbf{Q} costanti e ciò accade proprio quando l'incremento dato da \mathbf{Q} della matrice di covarianza prevista diventa uguale alla diminuzione del valore di quest'ultima nel computo della matrice di covarianza prevista per via del fattore $\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}$; quindi, in base all'Equazione (1.80), anche il Kalman gain si stabilizza su un valore costante. Vedendo la questione da una prospettiva diversa, se non si facessero misure il valore della matrice di covarianza prevista crescerebbe all'infinito, perché all'errore corrente sulla stima della posizione all'istante successivo andrebbero sommati anche tutti gli errori sulle stime precedenti; tuttavia, grazie al fatto che si eseguono misure, è come se ogni volta l'errore sulla previsione venisse resettato dal termine $\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}$, come già accennato, tramite un processo di inferenza bayesiana. Nella simulazione effettuata, in particolare, si ha che

$$\begin{aligned} P_{t+1|t+1} &\rightarrow \begin{pmatrix} 3.36 & 3.49 \\ 3.49 & 15.90 \end{pmatrix} m^2 \\ P_{t+1|t} &\rightarrow \begin{pmatrix} 5.47 & 7.27 \\ 7.27 & 21.90 \end{pmatrix} m^2 \end{aligned}$$

A questo punto potrebbe sorgere spontanea una domanda: perché, anche se non si effettuano misure della velocità, σ_v^2 si stabilizza su un valore limite? La grande potenza del filtro di Kalman è quella di sfruttare la covarianza tra la posizione e la velocità per aggiornare la distribuzione della velocità sulla base non solo della previsione della velocità, ma anche della previsione e della misura della posizione. Infatti, tramite dei semplici calcoli, a partire dall'Equazione (1.39) e imponendo la condizione espressa dall'Equazione (1.47), si perviene alla seguente espressione dell'aggiornamento della velocità:

$$v_{t+1|t+1} = -\mathbf{K}_{t+1}^{(2)} q_{t+1|t} + v_{t+1|t} + \mathbf{K}_{t+1}^{(2)} y_{t+1} \quad (2.30)$$

dove con $\mathbf{K}_{t+1}^{(2)}$ si è indicata la seconda componente del Kalman gain, con $q_{t+1|t}$ la previsione sulla posizione, con $v_{t+1|t}$ la previsione della velocità e con y_{t+1} la misura della posizione.

Come ulteriore approfondimento delle considerazioni riportate, ci si può chiedere quale sia in generale il tempo impiegato dal Kalman gain per stabilizzarsi sul valore limite. Esso chiaramente dovrà dipendere dal rapporto tra la varianza dell'errore sulla misura e la costante λ che identifica il rumore browniano sul valore vero: l'errore sulla misura è fisso, mentre la matrice di covarianza inizialmente cresce, almeno finché l'incremento dato da \mathbf{Q} nel calcolo della matrice di covarianza prevista non diventa pari al decremento nel valore di quest'ultima nel computo della matrice di covarianza aggiornata. Se l'errore sulla misura è elevato e λ è piccolo, l'apporto dato da \mathbf{Q} alla matrice di covarianza prevista è ridotto e di conseguenza, come mostrato in Figura (2.4), il Kalman gain rimane molto prossimo allo 0 per un elevato numero di intervalli di tempo. Quanto d'altro

canto si osserva dal grafico è che il Kalman gain subisce una crescita molto rapida tra il valore nullo iniziale e il valore limite; ciò accade perché, affinché il contributo dato da $\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}$ e quello di \mathbf{Q} si equivalgano, prima la matrice di covarianza prevista deve assumere valori che siano confrontabili con la matrice dell'errore sulla misura: finché quest'ultima condizione non viene verificata, infatti, il contributo di $\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H}$ rimane trascurabile rispetto a quello di \mathbf{Q} , per cui la matrice di covarianza prevista continua a crescere, mentre il Kalman gain rimane prossimo allo zero; quando il valore della matrice di covarianza prevista diviene confrontabile con l'errore sulla misura, allora inizia a crescere il Kalman gain fino a stabilizzarsi sul valore limite.

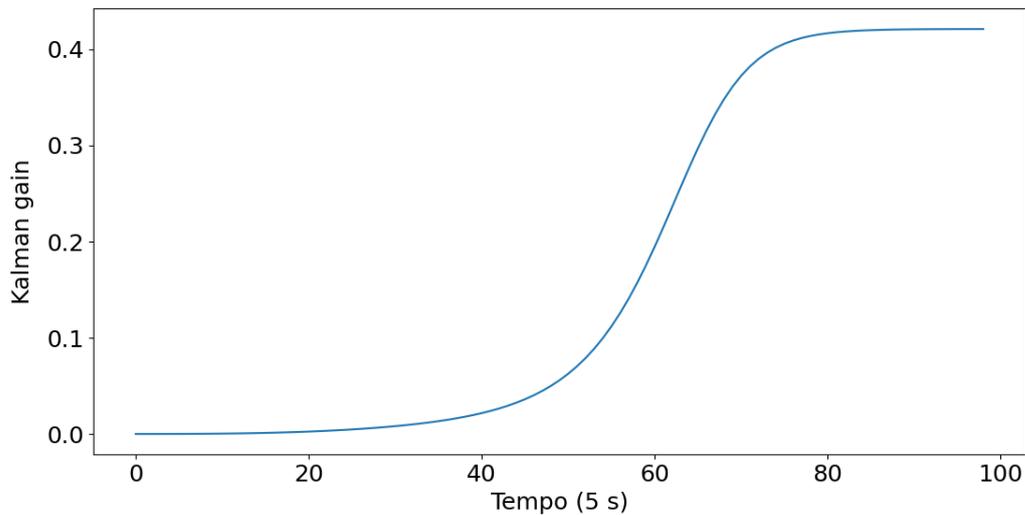


Figura (2.4): In questa immagine, in cui si sono utilizzati valori per l'errore assurdi a scopo puramente esplicativo, è rappresentato l'andamento temporale del Kalman gain adottando sempre il medesimo modello ma utilizzando $\lambda = 0.05 \text{ m}^2$ e $\sigma_\varepsilon = 100 \text{ m}$.

Nel caso limite in cui si ha $\lambda = 0$, il Kalman gain è nullo: infatti in tale situazione tramite il modello si possono eseguire previsioni perfette, per cui il peso attribuito alle misure è nullo. Analiticamente si spiega tale risultato perché risulterebbe nullo il vettore \mathbf{w}_t a ogni istante e sarebbe quindi nulla la matrice di covarianza prevista.

Trattando infine l'innovazione, se ne può osservare in Figura (2.5) l'andamento temporale.

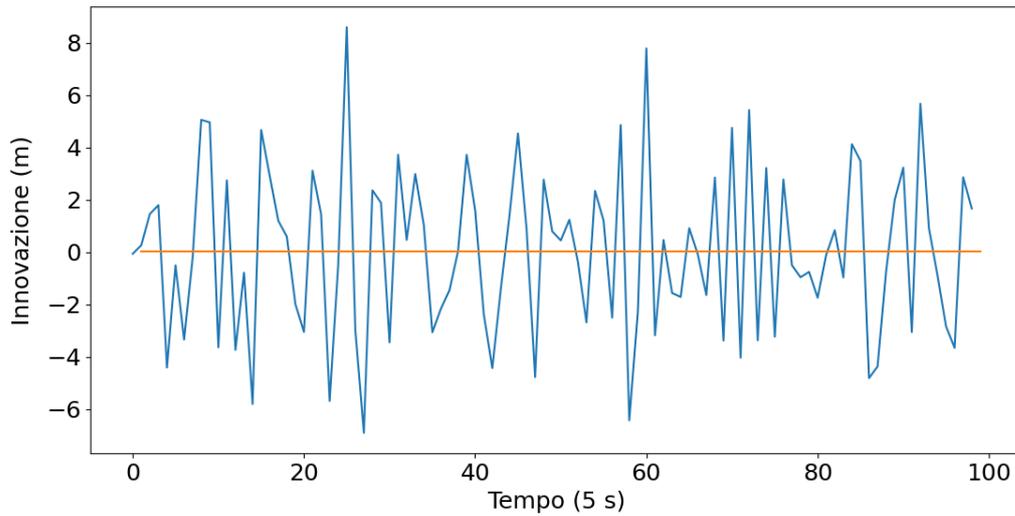


Figura (2.5): La linea blu nell'immagine rappresenta l'evoluzione temporale dell'innovazione ottenuta nella simulazione delle traiettorie rappresentate in Figura (2.1), mentre la linea arancione rappresenta la media dell'innovazione.

Nel grafico appena mostrato si vede come la media temporale della grandezza in esame sia molto prossima allo zero, che è esattamente quanto ci si attende. Dal grafico si osserva proprio che l'innovazione sia essa stessa distribuita secondo una distribuzione normale con media nulla. Una nota da aggiungere riguardo tale grandezza è che l'innovazione non la si calcola a partire dal tempo $t = 0$, ma dal tempo $t = 1$, giacché essa consiste proprio nella differenza tra la misura a un determinato tempo e la previsione effettuata nell'istante precedente e chiaramente prima dell'istante iniziale non è possibile effettuare previsioni; d'altra parte, poiché il Kalman gain quantifica il peso attribuito alla misura rispetto alla previsione effettuata nell'istante precedente, anch'esso si calcola a partire dal tempo $t = 1$.

2.2 Moto uniformemente accelerato

Qualora si abbia una accelerazione costante, limitando lo studio al caso unidimensionale, occorre innanzitutto definire un vettore di stato tridimensionale, nel quale la terza componente sia proprio l'accelerazione a , ottenendo

$$\mathbf{x} = \begin{pmatrix} q(t) \\ v(t) \\ a \end{pmatrix} \quad (2.31)$$

Di conseguenza la matrice del flusso di fase diventerà una matrice 3×3 e più nel dettaglio, siccome la posizione e la velocità hanno un'evoluzione temporale del tipo

$$\begin{cases} q(t) = q(0) + v(0)t + \frac{1}{2}at^2 \\ v(t) = v(0) + at \end{cases} \quad (2.32)$$

la matrice del flusso di fase sarà la seguente:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & t \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.33)$$

Alla relazione

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{A} \cdot \mathbf{x}(t) \quad (2.34)$$

occorre anche in questo caso aggiungere un rumore che, a differenza del caso mostrato nel precedente paragrafo, in accordo con la teoria che si sta costruendo, avrà tre componenti. Siccome si suppone nuovamente di prendere misure solamente della posizione, la matrice \mathbf{H} sarà

$$\mathbf{H} = (1, 0, 0) \quad (2.35)$$

Per riscrivere il sistema dell'Equazione (2.14) adattato a questo nuovo caso di studio si ripetono sostanzialmente gli stessi ragionamenti mostrati nel precedente paragrafo, ma con alcune differenze che richiedono accorgimenti particolari. Si trascuri dapprima il rumore e si consideri la relazione di partenza mostrata nell'Equazione (2.34), nella quale si moltiplicano quindi entrambi i membri per dt ottenendo

$$d\mathbf{x}(t) = (\mathbf{A}dt)\mathbf{x}(t) \quad (2.36)$$

Separando le variabili si ottiene, effettuando un abuso di notazione con il quale si semplifica la trattazione,

$$\frac{d\mathbf{x}(t)}{\mathbf{x}(t)} = \mathbf{A}dt = \begin{pmatrix} 0 & dt & tdt \\ 0 & 0 & dt \\ 0 & 0 & 0 \end{pmatrix} \quad (2.37)$$

Integrando ambo i membri si giunge alla seguente uguaglianza:

$$\log\left(\frac{\mathbf{x}_{k+1}}{\mathbf{x}_k}\right) = \begin{pmatrix} 0 & \Delta t & \frac{1}{2}\Delta t^2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.38)$$

Indicando quest'ultima matrice con \mathbf{G} , si può scrivere

$$\mathbf{x}_{k+1} = e^{\mathbf{G}} \cdot \mathbf{x}_k \quad (2.39)$$

Siccome la matrice \mathbf{G} contiene un termine quadratico in Δt e siccome anche le relazioni che determinano l'evoluzione temporale del sistema, ossia le relazioni riportate nell'Equazione (2.32), contengono termini che sono al massimo quadratici nel tempo, è opportuno arrestare lo sviluppo dell'esponenziale al primo ordine, ottenendo quindi

$$\mathbf{x}_{k+1} = (\mathbf{I} + \mathbf{G}) \cdot \mathbf{x}_k \quad (2.40)$$

Sostituendo a \mathbf{G} la relativa espressione si è dimostrato quanto segue:

$$\begin{cases} \mathbf{x}_{k+1} = \begin{pmatrix} 1 & \Delta t & \frac{1}{2}\Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{pmatrix} \cdot \mathbf{x}_k + \mathbf{w}_k \\ y_k = (1, 0, 0) \cdot \mathbf{x}_k + z_k \end{cases} \quad (2.41)$$

dove la seconda relazione la si ottiene semplicemente sostituendo a \mathbf{H} la sua espressione.

Per trovare l'errore sul vettore di stato si procede in maniera analoga a quanto eseguito nel caso del moto rettilineo uniforme. Con la trasformata di Laplace si può dimostrare¹ che

$$\mathbf{w}(t) = \int_0^{\Delta t} \begin{pmatrix} (\Delta t - t')^2/2 \\ (\Delta t - t') \\ 1 \end{pmatrix} w(t + t') dt' \quad (2.42)$$

Sostituendo tale rappresentazione per il rumore browniano sul vettore di stato nella definizione della matrice \mathbf{Q} , per la quale si rimanda all'Equazione (2.16), si ottiene la seguente uguaglianza:

$$\begin{aligned} \mathbf{Q}_t = E[& \left(\int_0^{\Delta t} \begin{pmatrix} (\Delta t - t')^2/2 \\ (\Delta t - t') \\ 1 \end{pmatrix} w(t + t') dt' \right) \cdot \\ & \cdot \left(\int_0^{\Delta t} \begin{pmatrix} (\Delta t - t'')^2/2 \\ (\Delta t - t'') \\ 1 \end{pmatrix} w(t + t'') dt'' \right)^T] \end{aligned} \quad (2.43)$$

da cui, ripetendo gli stessi passaggi svolti nel precedente paragrafo, si perviene alla seguente relazione:

$$\begin{aligned} \mathbf{Q}_t = \int_0^{\Delta t} \int_0^{\Delta t} & \begin{pmatrix} (\Delta t - t')^2/2 \\ (\Delta t - t') \\ 1 \end{pmatrix} \cdot \left((\Delta t - t'')^2/2, (\Delta t - t''), 1 \right) \times \\ & \times E[w(t + t')w(t + t'')] dt' dt'' \end{aligned} \quad (2.44)$$

Tenendo ancora conto che dt' e dt'' sono indipendenti e identicamente distribuiti, risolvendo il prodotto tra i due vettori e sostituendo al valore atteso l'espressione fornita dall'Equazione (2.24), si ottiene

$$\mathbf{Q}_t = \int_0^{\Delta t} \begin{pmatrix} (\Delta t - t')^4/4 & (\Delta t - t')^3/2 & (\Delta t - t')^2/2 \\ (\Delta t - t')^3/2 & (\Delta t - t')^2 & (\Delta t - t') \\ (\Delta t - t')^2/2 & (\Delta t - t') & 1 \end{pmatrix} \lambda dt' \quad (2.45)$$

Si ricava quindi, risolvendo l'integrale, l'espressione della matrice di covarianza del rumore browniano sul vettore di stato:

$$\mathbf{Q} = \begin{pmatrix} \Delta t^5/5 & \Delta t^4/8 & \Delta t^3/6 \\ \Delta t^4/8 & \Delta t^3/3 & \Delta t^2/2 \\ \Delta t^3/6 & \Delta t^2/2 & \Delta t \end{pmatrix} \quad (2.46)$$

Si è effettuata, tramite tutti gli elementi analizzati, una simulazione per mostrare l'andamento del sistema nel tempo e per testare il corretto funzionamento del filtro di Kalman. Il grafico ottenuto per la dinamica del sistema è riportato in Figura (2.6). Come valori delle costanti relative al moto si sono utilizzati $a = 2 \text{ ms}^{-2}$ e $v_0 = 3 \text{ ms}^{-1}$, dove con v_0 si indica la velocità iniziale, mentre per gli errori si è posto $\lambda = 50 \text{ m}^2$ e

¹Si rimanda a [Reid 2001].

$\sigma_\varepsilon = 2\text{ m}$; come intervallo temporale su cui effettuare stime, aggiornamenti e misure si è posto $\Delta t = 0.2$.

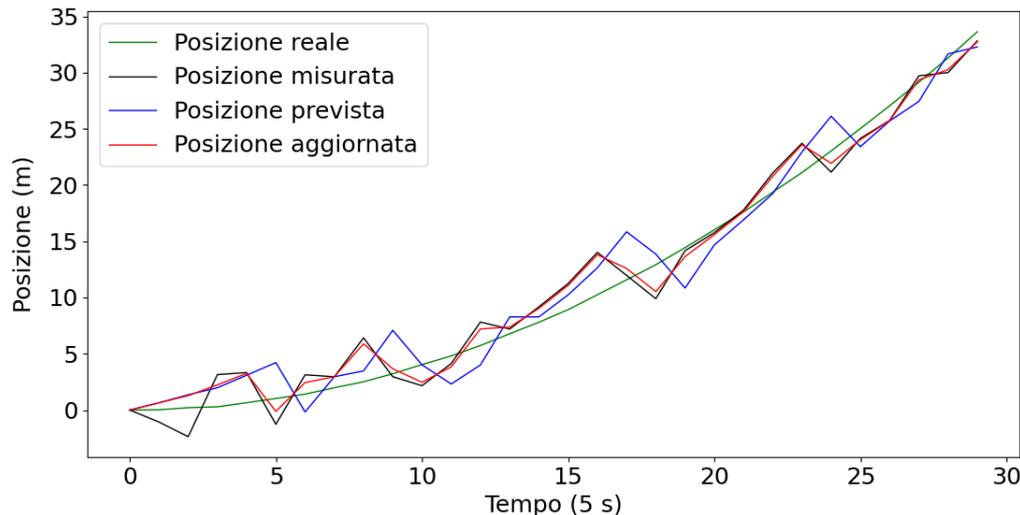


Figura (2.6): Nell'immagine è rappresentato l'andamento temporale della posizione reale, misurata, prevista e aggiornata.

La posizione reale segue un andamento parabolico con delle oscillazioni di fatto non osservabili dal grafico, in quanto, in questo caso, $\sigma_q^2 \sim \Delta t^5$. Inoltre, a differenza della simulazione presentata nel precedente paragrafo, si vede che la posizione prevista segue maggiormente la misura, in quanto il valore limite su cui si assesta la prima componente del Kalman gain è ≈ 0.8 , per cui si attribuisce maggior peso alle misure. Ciò accade perché in questo caso, nel computo della matrice di covarianza prevista, occorrono diversi elementi aggiuntivi, quali la varianza dell'accelerazione, la correlazione tra posizione e accelerazione e la correlazione tra velocità e accelerazione: il valore limite su cui si stabilizza $\mathbf{P}_{t+1|t}$ è dunque maggiore rispetto al caso del moto rettilineo uniforme e di conseguenza, nonostante il rumore browniano sia effettivamente molto minore, l'incertezza associata alle previsioni è comunque maggiore. Anche in questo caso si può osservare, sempre dal grafico in Figura (2.6), come la posizione prevista segua una traiettoria che consiste in prolungamenti, secondo un modello di evoluzione temporale parabolico, delle traiettorie seguite nei vari intervalli temporali dalla posizione aggiornata. Una cosa molto interessante che si può osservare, infine, è che inizialmente la posizione aggiornata segue molto fedelmente la previsione, discostandosi dal valore misurato: ciò accade perché, come già argomentato nel caso del moto rettilineo uniforme, la matrice di covarianza iniziale è nulla, per cui la prima componente del Kalman gain è inizialmente uguale a zero e non si tiene conto, nell'aggiornamento della posizione prevista, delle misure; poi la matrice di covarianza prevista cresce molto rapidamente, determinando una conseguente crescita del Kalman gain, per cui già dalla quinta iterazione la posizione aggiornata si riassetta molto più in prossimità della misura, in quanto ritenuta più affidabile della previsione.

Anche in questo caso l'innovazione si comporta in maniera compatibile con quanto atteso a livello teorico e il Kalman gain evolve in maniera molto simile a quanto osservato per il moto rettilineo uniforme, come mostrato nei due grafici riportati in Figura (2.7).

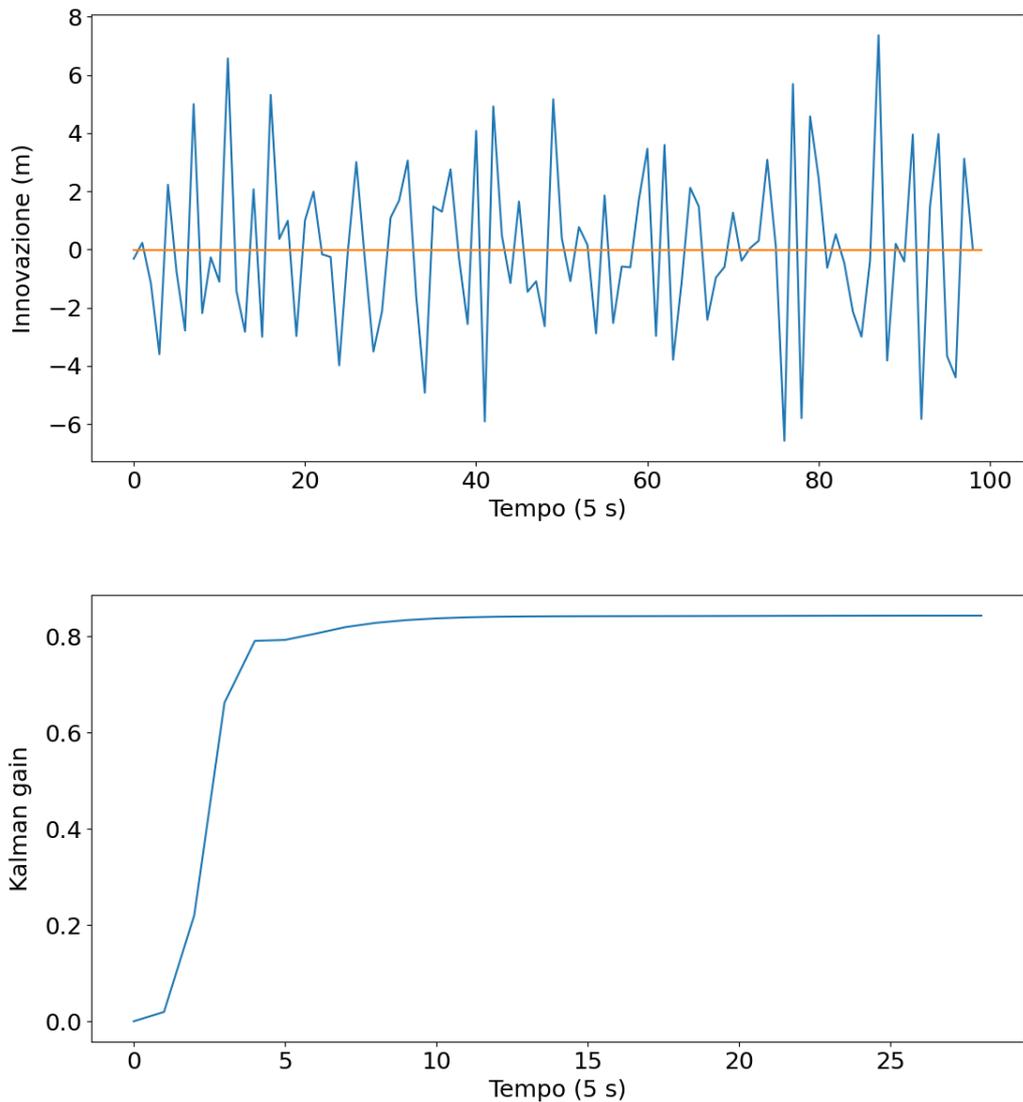


Figura (2.7): In alto è riportato in blu l'andamento temporale dell'innovazione, confrontato con la linea arancione che rappresenta la media effettuata su tutti i valori assunti dall'innovazione stessa, che, esattamente come ci si attende, assume un valore molto prossimo allo zero; in basso è riportato l'andamento temporale del Kalman gain, selezionando le prime 30 iterazioni sulle 100 effettuate per una migliore osservazione del comportamento del medesimo in prossimità dell'istante iniziale.

In questa parte ci si è limitati a riconfermare l'efficacia del filtro di Kalman, già mostrata nel caso del moto rettilineo uniforme, semplicemente adottando un modello diverso. Si vuole invece ora osservare il funzionamento dell'algoritmo qualora si adotti un modello diverso rispetto alle leggi orarie secondo le quali evolve il punto materiale.

2.3 Modello approssimativo

In questa sezione si vuole esaminare l'effetto del filtro di Kalman qualora il modello adoperato per la costruzione della matrice di transizione di stato sia diverso dal modello che determina il moto del punto materiale. Nella simulazione effettuata si è posto che il vettore di stato evolva, al netto del rumore, secondo un moto uniformemente accelerato, mentre come matrice \mathbf{F} si è adoperata quella costruita assumendo che il punto si muova di moto rettilineo uniforme. In buona sostanza si vuole valutare come funziona il filtro di Kalman qualora si approssimi il moto accelerato di un punto materiale con un moto a velocità costante. Dal momento che il sistema si muove con accelerazione costante, il vettore di stato ha tre componenti, per cui la matrice di transizione di stato va estesa al caso tridimensionale, considerando nulli gli elementi non previsti nella matrice costruita nel caso bidimensionale. Si avrà quindi

$$\mathbf{F} = \begin{pmatrix} 1 & \Delta t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (2.47)$$

Per le componenti del vettore di stato si sono invece ottenuti i valori tramite le leggi del moto uniformemente accelerato, nella seguente maniera:

$$\mathbf{x}(t) = \begin{pmatrix} q(t) \\ v(t) \\ a \end{pmatrix} = \begin{pmatrix} \frac{1}{2}at^2 \\ at \\ a \end{pmatrix} + \mathbf{w}(t) \quad (2.48)$$

Poiché il vettore di stato evolve secondo le leggi orarie del moto uniformemente accelerato e $\mathbf{w}(t)$ è riferito al vettore di stato e non alla matrice di transizione, l'errore sistematico sarà quello utilizzato nel modello dell'accelerazione costante; d'altronde la trasformata di Laplace per ottenere l'espressione di $\mathbf{w}(t)$ e di conseguenza della matrice \mathbf{Q}_t è applicata al vettore di stato e non alla matrice di transizione di stato. Quindi si avrà

$$\mathbf{Q} = \lambda \begin{pmatrix} \Delta t^5/20 & \Delta t^4/8 & \Delta t^3/6 \\ \Delta t^4/8 & \Delta t^3/3 & \Delta t^2/2 \\ \Delta t^3/6 & \Delta t^2/2 & \Delta t \end{pmatrix} \quad (2.49)$$

Poiché anche in questo caso si suppone di misurare solamente la posizione, la matrice di transizione in uscita è

$$\mathbf{H} = (1, 0, 0) \quad (2.50)$$

Il grafico che risulta dalla simulazione è riportato in Figura (2.8). Per quanto riguarda gli errori si è posto $\lambda = 50 m^2$ e $\sigma_\varepsilon = 2 m$, come valore per l'accelerazione si è utilizzato $a = 4 ms^{-2}$, per la velocità iniziale $v(0) = 0$ e come apiezza degli intervalli temporali si è posto $\Delta t = 0.2$. Rispetto alla simulazione trattata nel precedente paragrafo si è adoperato un valore per l'accelerazione maggiore perché fosse più evidente l'effetto della variazione della velocità ed emergesse un maggiore divario tra il modello utilizzato per la costruzione della matrice di transizione di stato e quello utilizzato per l'evoluzione temporale del vettore di stato.

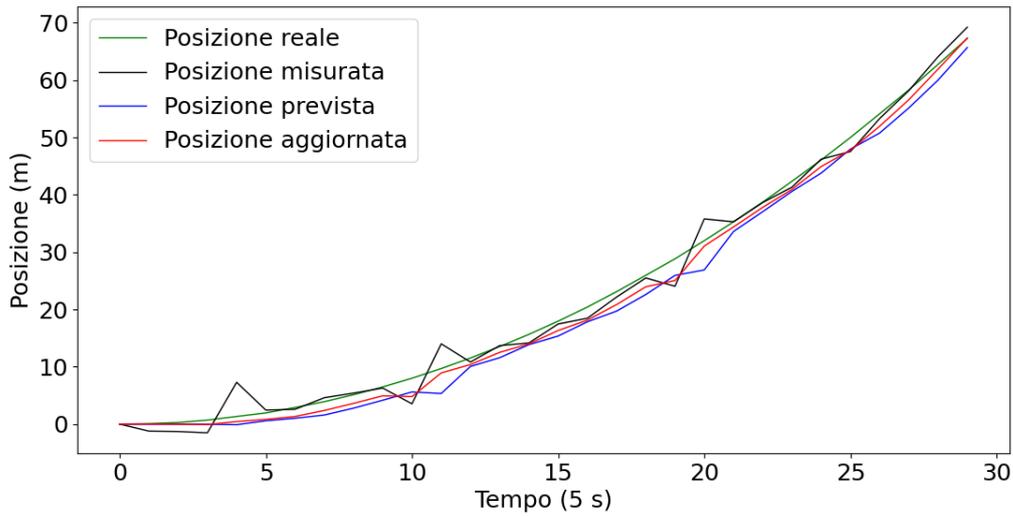


Figura (2.8): L'immagine rappresenta gli andamenti della posizione reale, misurata, prevista e aggiornata nel caso di un sistema che evolve secondo un moto parabolico, approssimato tramite un modello evolutivo a velocità costante.

Come anche nel paragrafo precedente, l'effetto del rumore browniano sulla posizione reale non è osservabile dal grafico, perché anche in questo caso $\sigma_q^2 \sim \Delta t^5$. Rispetto al grafico riportato in Figura (2.6), tuttavia, si osserva che la posizione prevista oscilla molto meno, non ha quasi mai derivata prima negativa e segue una traiettoria più fedele alla posizione reale rispetto al caso in cui la matrice \mathbf{F} è costruita secondo lo stesso modello evolutivo del sistema. Ciò accade perché l'algoritmo percepisce come errore sistematico al pari del rumore browniano sul vettore di stato la differenza tra il modello teorico con cui si costruisce la matrice della transizione di stato e il moto reale del punto materiale. Questa spiegazione risulta ancora più chiara analizzando ciò che succede all'innovazione, la cui evoluzione temporale è riportata in Figura (2.9).

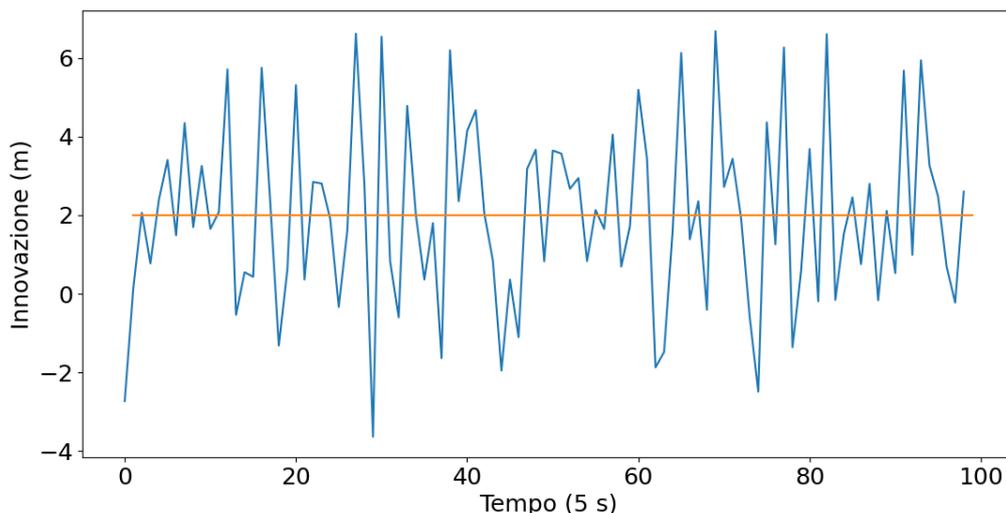


Figura (2.9): Nell'immagine è rappresentato l'andamento temporale dell'innovazione, con la linea blu, comparato con la media effettuata su tutti i valori, rappresentata dalla linea arancione.

Il valor medio dell'innovazione in questo caso è ben diverso dal valore medio atteso. La ragione consiste proprio nel fatto che il filtro percepisce come errore la differenza tra il valore misurato e il valore atteso. Tuttavia questa differenza non è dovuta, come nei casi precedenti, a un rumore browniano distribuito gaussianamente con media nulla, ma all'approssimazione effettuata in fase di costruzione del modello; per tale ragione, dal momento che la traiettoria del moto parabolico cresce più velocemente della traiettoria del moto lineare, sarà una differenza tendenzialmente positiva; non è sempre positiva perché si deve tener conto anche del rumore browniano sul vettore di stato e dell'errore sulla misura, ma la media della differenza, che consiste proprio nell'innovazione, è positiva, dal momento che la misura è mediamente maggiore della previsione. Quantificando il discorso, si ha che

$$E[\nu_{t+1}] = E[\mathbf{H} \cdot \mathbf{x}_{t+1} + z_{t+1} - \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t}] \quad (2.51)$$

Poiché l'errore sulla misura ha media nulla segue che

$$E[\nu_{t+1}] = E[q_{t+1}] - E[q_{t+1|t}] \quad (2.52)$$

dove q_{t+1} indica la posizione reale al tempo $t + 1$ e $q_{t+1|t}$ la posizione prevista. Poiché la posizione reale evolve secondo un moto uniformemente accelerato si ha che

$$q_{t+1} = q_t + v_t \Delta t + \frac{1}{2} a \Delta t^2 + w_{q_{t+1}} \quad (2.53)$$

dove $w_{q_{t+1}}$ indica il rumore browniano sulla posizione al tempo $t + 1$; d'altra parte si ha

$$q_{t+1|t} = q_t + v_{t|t} \Delta t \quad (2.54)$$

dal momento che la posizione prevista risulta dall'applicazione della matrice approssimativa \mathbf{F} alla posizione aggiornata, dove $v_{t|t}$ è la velocità aggiornata, che in generale è

diversa da quella reale ed è sempre crescente, perché, siccome lo spazio effettivamente percorso nell'intervallo Δt è sempre maggiore di quello previsto, il filtro cerca di adeguare le ipotesi sulla velocità alla posizione misurata, supponendo di avere, in ogni singolo tratto, un moto rettilineo con una velocità sempre maggiore iterazione dopo iterazione dell'algoritmo, cercando di compensare lo scarto che si calcola e provando a minimizzarlo. Alla luce di quanto detto si ha che

$$\begin{aligned} E[\nu_{t+1}] &= E[(v_t - v_{t|t})\Delta t + \frac{1}{2}a\Delta t^2 + w_{qt+1}] = \\ &= E[(v_t - v_{t|t})\Delta t + \frac{1}{2}a\Delta t^2] > 0 \end{aligned} \tag{2.55}$$

Quindi, tornando a osservare ciò che succede nel grafico di Figura (2.8), per quanto detto la misura è mediamente maggiore della previsione; ciò comporta che la posizione aggiornata, dovendo essere sempre compresa tra la posizione misurata e la posizione prevista, sarà mediamente maggiore di quest'ultima; dal momento che la posizione prevista prolunga secondo il modello con cui si è costruita la matrice della transizione di stato il trend seguito dalla posizione aggiornata nell'intervallo di tempo precedente, la traiettoria sarà sempre crescente. In altre parole, poiché la posizione prevista aggiusta la propria traiettoria in base alle fluttuazioni della posizione aggiornata, siccome quest'ultima è mediamente maggiore, le previsioni tenderanno ad adeguarsi provando ad aumentare sempre di più il coefficiente angolare del tratto lineare descritto in un intervallo di tempo.

Si riporta infine l'andamento temporale della prima componente del Kalman gain in Figura (2.10); rispetto al caso trattato nel precedente paragrafo il valore limite sui cui si stabilizza è molto minore, dal momento che nel computo della matrice di covarianza prevista si trascurano la correlazione tra la posizione e l'accelerazione, la correlazione tra la velocità e l'accelerazione e la varianza dell'accelerazione. D'altra parte, poiché gli intervalli Δt sono stati assunti piccoli, la posizione prevista riesce ad adeguarsi correttamente alla posizione reale; ciò consente di avere delle buone stime tramite la previsione, dal momento che il filtro le considera più affidabili rispetto alle misure; in altre parole, l'errore dovuto all'approssimazione viene ritenuto più accettabile rispetto all'errore sulla misura, per cui, per intervalli piccoli, si può concludere che localmente l'approssimazione lineare del moto parabolico è buona.

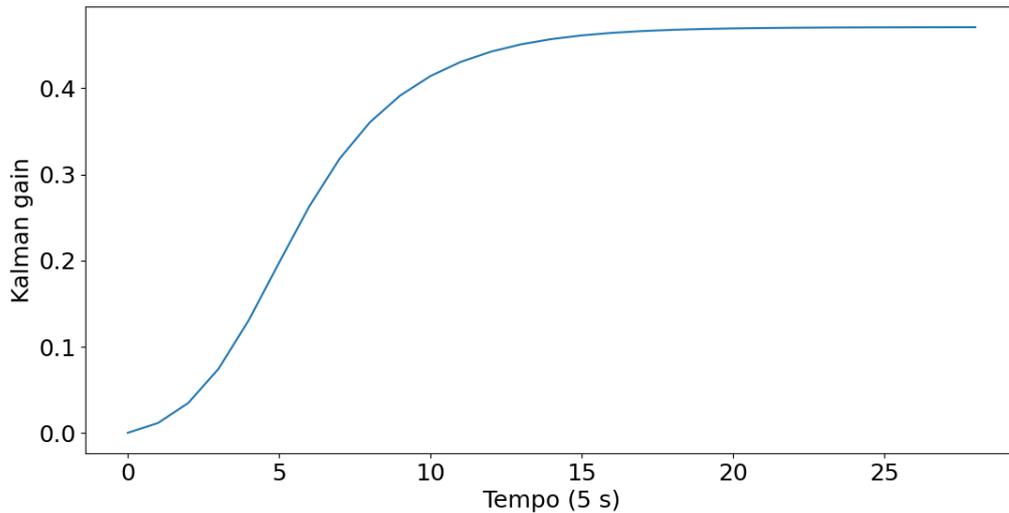


Figura (2.10): L'immagine rappresenta l'evoluzione temporale del Kalman gain sulla posizione.

Qualora si aumenti l'ampiezza degli intervalli, le previsioni diventano meno accurate e il discostamento della posizione prevista dalla posizione reale è ritenuto meno accettabile dell'errore sulla misura: ripetendo la simulazione con $\Delta t = 0.5$ si ottiene

$$\mathbf{K}_t^{(1)} \rightarrow (\approx 0.9)$$

per cui si osserva che l'approssimazione eseguita rende il filtro fortemente sensibile all'ampiezza dei tratti su cui si effettua l'approssimazione stessa dell'evoluzione del sistema. Lo studio di ciò che accade per un modello approssimativo è propedeutico all'analisi, presentata nel prossimo paragrafo, di ciò che accade applicando il filtro di Kalman a un oscillatore armonico la cui evoluzione temporale è approssimata con un modello ad accelerazione costante.

2.4 Oscillatore armonico

In questo paragrafo si mostra come si comporta il filtro di Kalman nel caso di un oscillatore armonico unidimensionale. Anche in questo caso, come nella simulazione trattata nel precedente paragrafo, si opera tramite una approssimazione. Il punto materiale si muove, al netto del rumore, secondo le leggi del moto armonico, come riportato di seguito:

$$\mathbf{x}(t) = \begin{pmatrix} q(t) \\ v(t) \\ a(t) \end{pmatrix} = \begin{pmatrix} A \cdot \text{sen}(\omega t + \varphi_0) \\ A\omega \cdot \text{cos}(\omega t + \varphi_0) \\ -A\omega^2 \cdot \text{sen}(\omega t + \varphi_0) \end{pmatrix} + \mathbf{w}(t) \quad (2.56)$$

dove A indica l'ampiezza delle oscillazioni, ω la velocità angolare, φ_0 la fase iniziale dell'oscillazione e $\mathbf{w}(t)$ denota il rumore sul valore vero del vettore di stato. Come si osserva dalle relazioni giustappunto riportate, sebbene l'accelerazione non sia costante, tuttavia la si assumerà tale nell'elaborazione del modello tramite il quale si costruisce

la matrice di transizione di stato, per cui proprio in questo consisterà l'approssimazione. Dunque come matrice \mathbf{F} si avrà la stessa ottenuta nel caso del moto uniformemente accelerato, ossia

$$\mathbf{F} = \begin{pmatrix} 1 & \Delta t & \frac{1}{2}\Delta t^2 \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \end{pmatrix} \quad (2.57)$$

In questo caso tuttavia, si effettua una ulteriore approssimazione, ossia costruire la matrice \mathbf{Q} , ossia la matrice di covarianza del rumore browniano, cosiccome il vettore $\mathbf{w}(t)$ stesso, assumendo che il moto del punto materiale sia uniformemente accelerato. La ulteriore approssimazione consiste proprio nel fatto che la variazione dell'accelerazione che si ha nel moto del punto materiale viene considerata come un rumore browniano sull'accelerazione stessa, oltre considerare come errore sistematico la discrepanza che si ha nella posizione prevista e reale sulla base dei due diversi modelli di evoluzione temporale. Ciò, come si osserverà nel seguito, porta a delle previsioni meno accurate rispetto al caso trattato nel precedente paragrafo, con un conseguente aumento della prima componente del Kalman gain. Questa approssimazione è tuttavia ragionevole, perché le variazioni dell'accelerazione saranno mediamente nulle qualora si consideri un numero intero di periodi: la variazione dell'accelerazione che si ha con l'inversione della direzione del moto in corrispondenza di un massimo sarà compensata con la variazione in corrispondenza di un minimo nelle oscillazioni.

Occorre inoltre, prima di proseguire e commentare il funzionamento del filtro di Kalman nel caso in esame, effettuare altre due precisazioni aggiuntive. La prima riguarda la matrice di covarianza: a differenza degli altri casi in tale circostanza si è posto

$$\mathbf{P} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \quad (2.58)$$

La diagonale principale è nulla, perché si è assunto che le condizioni iniziali siano conosciute con certezza e che siano cioè distribuite secondo una distribuzione con varianza nulla. In questo caso esistono, tuttavia, delle correlazioni tra di esse, giacché dipendono tutte da un unico parametro, che è φ_0 , ossia la fase iniziale. Ponendo $\varphi_0 = 0$ si ottengono le seguenti condizioni iniziali:

$$\begin{cases} q(0) = A \cdot \text{sen}(0) = 0 \\ v(0) = A\omega \cdot \text{cos}(0) = A\omega a(0) = -A\omega^2 \cdot \text{sen}(0) = 0 \end{cases} \quad (2.59)$$

Poiché l'accelerazione e la posizione iniziali sono in fase, mentre rispetto a esse la velocità iniziale è in opposizione di fase², si ha che

$$\begin{aligned} \sigma_{qa} &= 1 \\ \sigma_{qv} &= -1 \\ \sigma_{va} &= -1 \end{aligned} \quad (2.60)$$

²Ciò implica che, qualora si scegliesse una φ_0 diversa dallo zero, al crescere del valore della posizione iniziale cresce il valore dell'accelerazione iniziale e al decrescere del valore della posizione iniziale decresce quello dell'accelerazione iniziale; per la velocità iniziale vale il discorso opposto, ossia al crescere del valore di posizione e accelerazioni iniziale diminuisce quello della velocità iniziale e viceversa.

dove si è denotato con σ_{qa} la covarianza tra posizione e accelerazione iniziali, con σ_{qv} la covarianza tra posizione e velocità iniziali e con σ_{va} la covarianza tra velocità e accelerazione iniziali. Le relazioni riportate nell'Equazione (2.60) spiegano perché si è costruita la matrice di covarianza iniziale come riportato nell'Equazione (2.58).

La seconda precisazione riguarda l'ampiezza degli intervalli temporali. Nei casi precedentemente esaminati si è posto $\Delta t = 0.2$, anche se nulla avrebbe impedito di porre, per esempio, $\Delta t = 1$; per il moto periodico occorre invece rispettare il Teorema del campionamento di Nyquist. Esso afferma che, dato un segnale sinusoidale, perché si possa ricostruire in maniera accurata il segnale stesso ed evitare fenomeni di aliasing, la frequenza di campionamento deve essere tra le cinque e le dieci volte maggiore della frequenza del segnale medesimo. Nel caso in esame, come si potrà osservare dal grafico dell'evoluzione temporale del sistema successivamente riportato, il periodo del moto armonico è $T_0 \simeq 3 s$; dovendo essere il periodo di campionamento, che corrisponde all'inverso della frequenza di campionamento, tra le cinque e le dieci volte minore del periodo del segnale, si deve avere che $\nu_C < 0.6 s$, dove ν_C è proprio la frequenza di campionamento. Per tale ragione si è posto, come grandezza adimensionale, $\Delta t = 0.1$.

Terminate le premesse, dal momento che l'errore sulle misure non presenta differenze rispetto a quanto esplicitato nell'Equazione (2.1), si può ora discutere dell'applicazione del filtro di Kalman al caso di un oscillatore armonico unidimensionale; in Figura (2.11) è riportata l'evoluzione temporale del sistema. I valori utilizzati per gli errori sono $\lambda = 30 m^2$ e $\sigma_\varepsilon = 4 m$; come valore dell'ampiezza delle oscillazione si è adoperato $A = 50 m$ e la velocità angolare è stata settata al valore $\omega = 2 s^{-1}$.

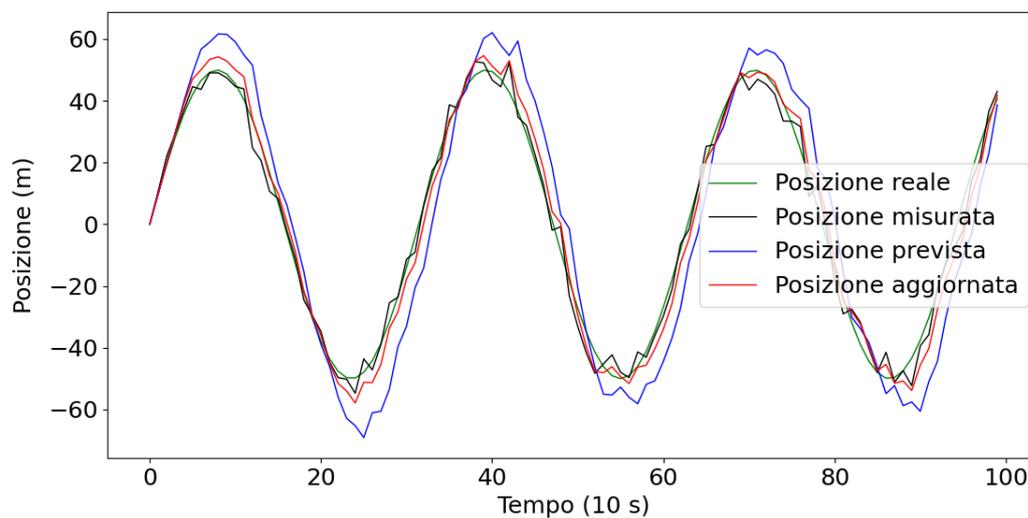


Figura (2.11): L'immagine rappresenta l'evoluzione temporale della posizione reale in verde, misurata in nero, prevista in blu e aggiornata in rosso.

Dal grafico si osserva come la traiettoria prevista si discosti leggermente dalla traiettoria misurata. Ciò avviene perché come modello con cui si è costruita la matrice di transizione di fase si è adoperato il moto uniformemente accelerato e si è considerata, per altro, la variazione dell'accelerazione nel moto reale del punto come rumore distribuito con media nulla sull'accelerazione prevista, che si suppone costante. In particolare il

discostamento maggiore lo si ha in prossimità dei punti di massimo e di minimo, dove la variazione della velocità è massima e proprio in quei tratti l'approssimazione effettuata risulta meno accurata e la previsione fatica maggiormente ad adeguarsi alla posizione reale, essendo costretta a seguire un andamento quadratico nel tempo.

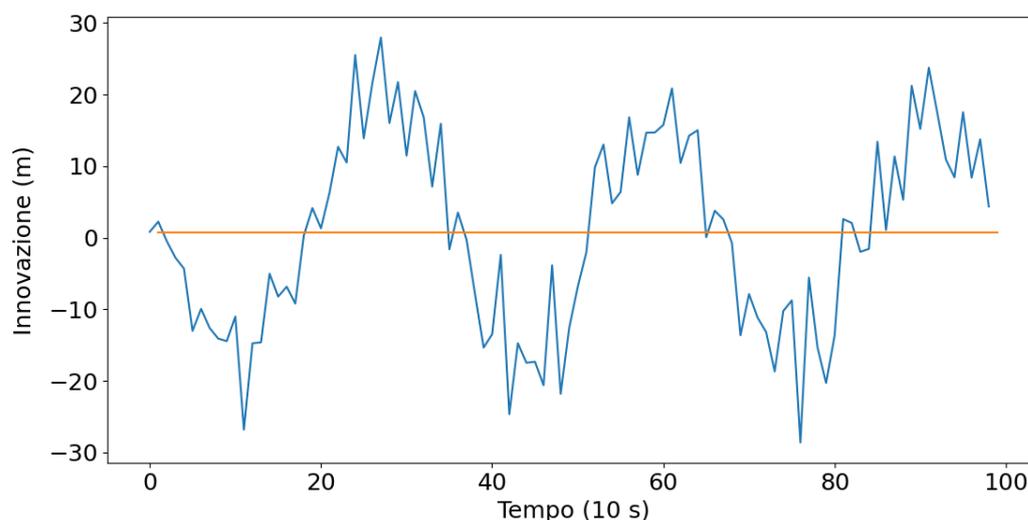


Figura (2.12): In figura è riportato l'andamento temporale dell'innovazione, rappresentato dalla linea blu, confrontato con la media, indicata con la linea arancione.

Per quanto riguarda l'innovazione, in questo caso si osserva che il valore della media è compatibile con il valore atteso, anche se si ha un errore nella costruzione del modello; si osserva anche che l'innovazione ha essa stessa un andamento periodico. Il modo in cui il rumore browniano sulla posizione influisce sull'innovazione, è visibile nelle contenute fluttuazioni che disturbano leggermente il trend periodico dell'innovazione; in altre parole la perturbazione causata dall'errore sistematico è piccola perché il rumore browniano è molto minore dell'errore commesso con l'approssimazione del modello. Infatti, come si può notare dal grafico, al netto di questo rumore sulla posizione, per un primo tratto l'innovazione ha un trend negativo. Infatti si ha che

$$a(t) = -A\omega^2 \cdot \text{sen}(\omega t) \quad (2.61)$$

per cui al crescere del tempo, a causa del segno negativo, l'accelerazione decresce, mentre secondo il modello teorico adoperato essa dovrebbe rimanere costante: questa divergenza motiva il trend decrescente dell'innovazione. Ciò accade finché $\omega t < \pi$. Infatti in π il seno cambia segno, per cui l'accelerazione inizia a crescere e inizia un tratto di trend positivo. In generale l'andamento periodico dell'innovazione rispecchia l'andamento periodico dell'accelerazione. Poiché l'accelerazione è una funzione sinusoidale e l'integrale tra 0 e 2π della funzione seno fornisce risultato nullo, considerando un numero intero di periodi l'innovazione sarà complessivamente nulla: infatti in tal caso i contributi positivi dell'innovazione compensano quelli negativi, mentre il contributo medio dato dall'errore sistematico è nullo, essendo esso distribuito normalmente con media nulla.

Si conclude questa trattazione mostrando, in Figura (2.13), l'andamento temporale Kalman gain.

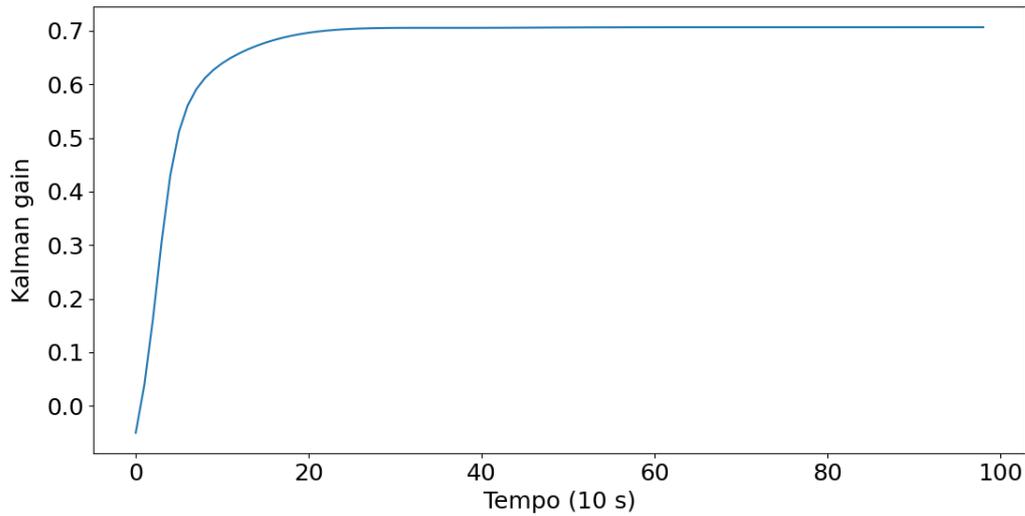


Figura (2.13): L'immagine rappresenta l'andamento temporale del Kalman gain nel caso dell'oscillatore armonico. Non si osservano rilevanti differenze rispetto ai casi trattati in precedenza.

Esso, come negli altri casi, si stabilizza rapidamente sul valore limite, anche se rispetto all'approssimazione mostrata nel precedente paragrafo, per quanto sia ulteriormente ridotta l'ampiezza dell'intervallo Δt , il Kalman gain attribuisce un maggiore peso alle misure: ciò vuol dire che l'approssimazione viene ritenuta peggiore dell'errore sulla misura, dal momento che sono state evidenziate alcune discrepanze tra la posizione reale e quella prevista alquanto consistenti. L'errore che si commette quindi approssimando un moto parabolico con un moto rettilineo è quindi minore rispetto a quello commesso approssimando un moto armonico localmente con un moto uniformemente accelerato; rispetto al caso precedente, per altro, è stato ridotto il rumore browniano ed è stato aumentato l'errore sulla misura. Ciononostante la posizione prevista riesce comunque a seguire la posizione reale e il Kalman gain riesce a stabilizzarsi su un valore limite per il quale si tiene conto in parte anche delle previsioni effettuate, a riprova del fatto che il filtro riesce a funzionare correttamente.

Dopo aver concluso i test del filtro di Kalman su questi quattro casi fisici di crescente difficoltà, nel prossimo capitolo ci si concentrerà su un argomento di principio completamente diverso, ossia sui titoli di borsa, che però, come si vedrà, può essere ricondotto agli esempi già analizzati.

Capitolo 3

Serie storiche di rendimenti

In questa parte si presenterà una applicazione del filtro di Kalman estranea all'ambito per il quale fu inizialmente studiato e concepito; in particolare, se ne testerà l'efficacia in materia finanziaria. Per poter arrivare, tuttavia, a discutere sul funzionamento del filtro di Kalman relativamente ai titoli di borsa, è necessario prima introdurre il modello che verrà utilizzato.

3.1 L'ipotesi di random walk

Nel 1953 Maurice Kendall presenta alla Royal Statistical Society un paper dal titolo *The Analysis of Economic Time Series. Part 1: Prices*, nel quale viene esposta una importante teoria nell'ambito dei mercati finanziari nota come *ipotesi di random walk*. Secondo tale teoria, come suggerisce il nome stesso che ha assunto, i prezzi degli strumenti di borsa evolvono nel tempo secondo un modello di random walk, tanto che nella propria pubblicazione Kendall paragona ogni serie storica dei prezzi a "un essere errante, quasi come se la sorte, una volta alla settimana, scrivesse un numero a caso e lo sommasse al prezzo corrente, determinando così il prezzo della settimana successiva" [Kendall 1953]. Questo vuol dire che le variazioni nei prezzi delle azioni sono casuali e che dunque non sono predicibili. All'origine di questa dottrina economica si ha la *teoria del mercato efficiente* di Eugene Fama, presentata nel 1969 in una pubblicazione dal titolo *Efficient Capital Markets: A Review of Theory and Empirical Work*, in cui si identifica come efficiente un mercato cui partecipano molti investitori, ciascuno dei quali si comporta in maniera razionale e ambisce a massimizzare i propri profitti cercando di prevedere i valori di mercato futuri dei singoli titoli avendo a disposizione le medesime informazioni rispetto ai propri competitors; contestualmente, vengono individuate tre tipologie di mercato efficiente, di seguito illustrate seguendo quanto teorizzato in [Fama 1969].

Nei mercati *efficienti in forma debole* i prezzi dei titoli rispecchiano tutte e sole le informazioni contenute nelle serie storiche dei prezzi passati. In questi mercati realizzare profitti sistematici è teoricamente impossibile, dal momento che, se lo studio dei prezzi passati potesse essere sfruttato per prevedere le variazioni di prezzo future, un qualsiasi investitore non avrebbe alcuna difficoltà a realizzare ingenti guadagni; ciò tuttavia non accade, perché qualora gli investitori tentassero di trarre vantaggio dalle informazioni contenute nei prezzi passati, i prezzi attuali si adeguerebbero immediatamente, impedendo così qualsiasi opportunità di profitto. Quindi, nei mercati efficienti in forma

debole, tutte le informazioni date dai prezzi passati, vengono assorbite solamente dai prezzi attuali senza condizionare quelli futuri; di fatti, non appena viene individuata una sistematicità nell'andamento dei prezzi, tale sistematicità viene meno e di conseguenza le variazioni di prezzo in un periodo non influenzano le variazioni in un periodo successivo e questo implica che i prezzi seguano un cammino casuale.

I mercati *efficienti in forma semiforte*, invece, riflettono non solo le informazioni contenute nelle serie storiche dei prezzi, ma anche quelle disponibili al pubblico, per esempio tramite giornali o, in tempi moderni, siti finanziari. Anche in questo caso realizzare profitti sistematici è teoricamente impossibile, perché le informazioni che possiede un investitore qualunque sono a portata di qualsiasi altra persona; se l'investitore quindi cercasse di sfruttarle per realizzare profitti sistematici, potrebbero effettuare la stessa operazione tutte le altre persone attive nei mercati finanziari. I prezzi si adeguerebbero di conseguenza, annullando ogni possibilità di profitto, e verrebbe meno l'andamento sistematico. Quindi anche nei mercati efficienti in forma semiforte i prezzi futuri seguono teoricamente un andamento casuale.

Nei mercati *efficienti in forma forte*, infine, i prezzi rispecchiano tutte le informazioni che si possono acquisire, anche quelle disponibili solamente agli insider, ossia alle persone più vicine agli amministratori delle aziende quotate in borsa, che sono gli unici a possedere tutti i dati e le informazioni che si possono conoscere. Anche questa forma di efficienza non consente di effettuare guadagni sistematici. Infatti un insider può sfruttare la conoscenza di informazioni riservate per effettuare guadagni ingenti nell'immediato, cosa che è teoricamente impossibile per le altre due tipologie di mercato; tuttavia, tramite il meccanismo di imitazione collettiva¹, presto l'informazione diviene di dominio pubblico, riconducendosi in questo modo al caso dell'efficienza semiforte e per un investitore che non sia un insider non è quindi possibile ottenere profitti superiori a quelli del mercato.

Lo stesso Fama fu un importante sostenitore dell'ipotesi di random walk, asserendo, in un articolo intitolato *Random Walks in Stock Market Prices*, che le serie delle variazioni nei prezzi sono "prive di memoria" [Fama 1965]; in altre parole, tali variazioni sono variabili distribuite secondo una determinata distribuzione di probabilità e sono indipendenti le une dalle altre, per cui la variazione di prezzo al giorno n non influisce minimamente nella variazione al giorno $n + 1$. Per accreditare ulteriormente questa tesi, si propone un'analisi basata sul rendimento di un titolo azionario, una grandezza concettualmente e analiticamente connesso alle variazioni dei prezzi dei titoli medesimi.

Denotando con S_{n-1} il prezzo di chiusura del titolo di chiusura nel giorno $n - 1$ e con S_n il prezzo di chiusura nel giorno n si ha che, seguendo quanto illustrato in [Brealey 2015]

$$r_n = \frac{S_n - S_{n-1}}{S_{n-1}} \quad (3.1)$$

dove r_n è proprio il rendimento, che viene dunque a indicare il guadagno effettuato rispetto al valore di partenza del titolo. Se la variazione del prezzo è casuale, allora ci si aspetta, sulla base di quanto si può osservare dall'espressione dei rendimenti, che anche questi siano casuali, cioè indipendenti l'uno dall'altro. A sostegno di questa tesi, si sono raccolti i prezzi di chiusura giornalieri delle azioni di *Intel Corporation* dal 01 gennaio 2019 al 01 gennaio 2022 e si sono conseguentemente calcolati i rendimenti, graficando

¹Secondo il meccanismo di imitazione collettiva, quando si osserva un insider assumere nel mercato un determinato atteggiamento, si segue tale comportamento.

quindi il rendimento relativo al giorno n in funzione del rendimento relativo al giorno $n - 1$, cioè in funzione del rendimento nel giorno precedente. Il risultato ottenuto è riportato in Figura (3.1).

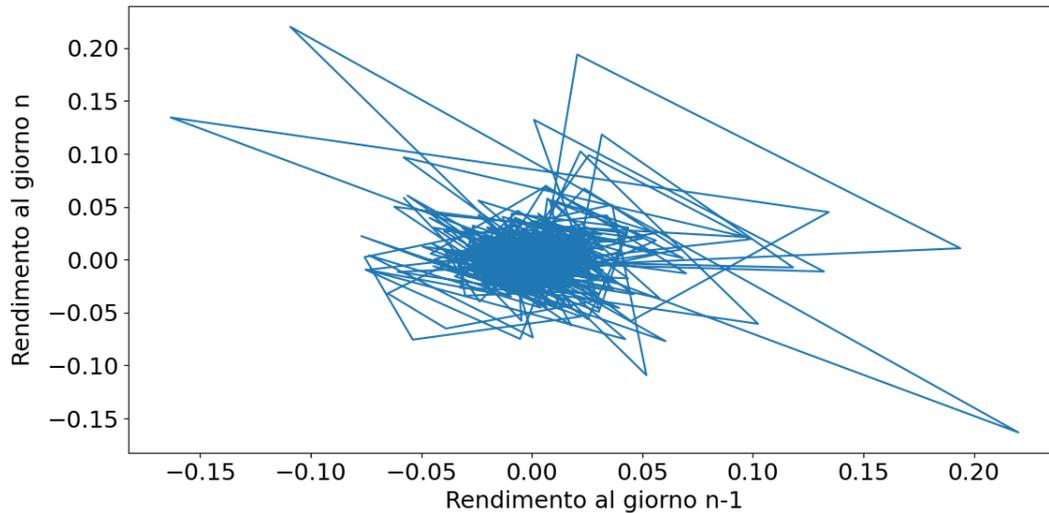


Figura (3.1): L'immagine rappresenta la traiettoria dei rendimenti nello spazio dell'evoluzione dei rendimenti medesimi. In altre parole i punti che hanno come ascissa il rendimento al giorno $n - 1$ e come ordinata il rendimento al giorno n sono stati connessi, disegnando la traiettoria tipica di un random walk bidimensionale.

Fonte dei dati: <https://finance.yahoo.com/>

Nel grafico è rappresentata la traiettoria del rendimento nello spazio dell'evoluzione dei rendimenti stessi; sostanzialmente la traiettoria disegnata rappresenta il percorso effettuato dal rendimento nell'evoluzione da un giorno al suo successivo e quello che si osserva è una traiettoria totalmente casuale che dimostra una scorrelazione tra le grandezze in ascissa e ordinata. Qualora vi fosse stata, infatti, una dipendenza del rendimento nel giorno n da quello nel giorno precedente, nel grafico si sarebbe dovuto osservare un maggiore addensamento di curve in una determinata direzione, cosa che evidentemente non si verifica.

Si è altresì calcolata la correlazione tra tutti i punti rappresentati nel grafico di Figura (3.1), ricordando che il coefficiente di correlazione lineare tra due grandezze omogenee x e y è definito, facendo riferimento a [Fornasini 2008] come

$$\alpha = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (3.2)$$

dove N indica il numero di misure e \bar{x} e \bar{y} indicano i valori medi delle due grandezze in esame. In particolare, α è contenuto nell'intervallo $[-1, 1]$ e per valori prossimi al limite superiore le grandezze crescono una al crescere dell'altra, per valori prossimi al limite inferiore decrescono una al crescere dell'altra, per valori prossimi allo zero sono indipendenti; inoltre il valore numerico del coefficiente di correlazione lineare permette

di determinare la probabilità che, considerando due grandezze che sono indipendenti certamente, si abbia un coefficiente di correlazione lineare in modulo maggiore o uguale a quello effettivamente ottenuto, indicata con $P(|r| > r_0)$. Per il caso in esame si è ottenuto $\alpha \simeq -0.23$ considerando tutto il campione di 755 misure, che coprono un intervallo temporale di tre anni; considerando invece solamente le prime 50 misure, che coprono un periodo di circa due mesi si è ottenuto $\alpha \simeq -0.06$ e per tale valore si ha come valore tabulato $P(|r| > r_0) > 50\%$. Come si nota, i rendimenti sono maggiormente correlati considerando scale temporali maggiori: questo è dovuto al fatto che sul lungo periodo i rendimenti scontano l'influenza delle condizioni generali dell'economia e gli effetti esterni di grande portata, come nel caso in esame possono essere stati la crisi pandemica da COVID-19 e le previsioni sulla spirale inflazionistica del 2022. Dunque è ragionevole supporre che i rendimenti sul breve periodo siano indipendenti l'uno dall'altro e che lo siano quindi anche le variazioni giornaliere dei prezzi dei titoli azionari.

3.2 L'esperimento di Malkiel

In questo capitolo si presenterà un complemento volto ad approfondire e accreditare ulteriormente l'ipotesi di random walk. Uno dei maggiori sostenitori di tale dottrina è Burton Malkiel, professore di economia presso l'Università di Princeton e autore del saggio *A Random Walk Down Wall Street*, nel quale arriva ad affermare che "Una scimmia bendata che lancia freccette contro le pagine di un giornale finanziario potrebbe selezionare un portafoglio che andrebbe bene come quello accuratamente selezionato dagli esperti" [Malkiel 2015], supportando quindi l'ipotesi di random walk fino alle sue estreme conseguenze. Malkiel è anche famoso per un esperimento visivo che porta il suo nome, sempre a sostegno della teoria in esame, illustrato anche in [Malkiel 2015]: l'andamento temporale di una azione le cui variazioni in termini di prezzo siano determinate dal lancio di una moneta presenta la stessa forma dell'andamento temporale di un titolo in borsa; tale esperimento viene quindi riproposto in questa sede in una variante più accurata.

Sia data una azione valutata inizialmente 100\$ e si supponga di lanciare settimanalmente un dado non truccato per determinare la variazione di prezzo dell'azione. Se esce 1, il valore dell'azione aumenterà del 5%; se esce 2 il valore aumenterà del 3%; se esce 3 il valore aumenterà dell'1%; se esce 4 il valore diminuirà dell'1%; se esce 5 il valore diminuirà del 3%; se esce 6 il valore diminuirà del 5%. Il grafico dell'andamento temporale del valore di un'azione così determinato effettuando 60 iterazioni viene paragonato in Figura (3.2) con le misure effettuate con frequenza mensile del prezzo di chiusura dell'indice di borsa *Standard&Poors500* nell'arco temporale dal 01 febbraio 2002 al 01 febbraio 2007; i valori di chiusura dell'indice sono stati normalizzati in maniera che il valore iniziale fosse uguale a quello dell'azione.

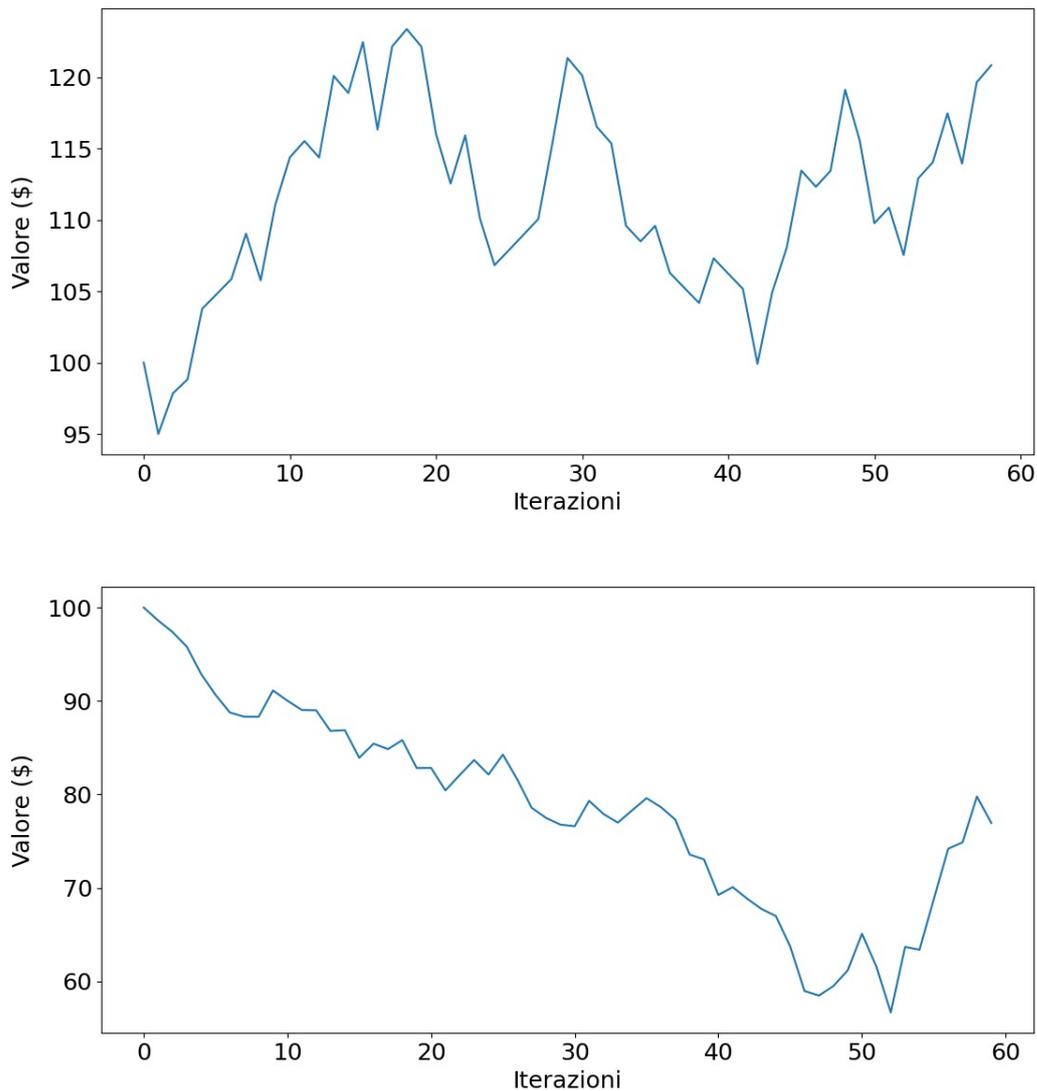


Figura (3.2): Uno dei due grafici rappresenta l'andamento dell'indice Standard & Poors500, l'altro del valore di una azione la cui evoluzione è determinata dal lancio di un dado. Si provi a indovinare quale dei due grafici rappresenta l'indice di borsa e quale l'evoluzione determinata dal lancio del dado; la soluzione è fornita nell'Appendice. Fonte dei dati: <https://finance.yahoo.com/>

Pur non essendo due grafici perfettamente sovrapponibili, la forma che assumono le traiettorie è simile e, senza avere informazioni, non sarebbe possibile individuare, semplicemente guardando i grafici, quale dei due riguarda l'indice Standard&Poors500 e quale l'azione. Questo dimostra che è corretto supporre che le variazioni nei prezzi delle azioni o nei punti degli indici siano, sul breve periodo, del tutto casuali. Dopo aver portato sufficienti prove a supporto di questa teoria finanziaria, si vuole ora procedere individuando un modello che possa descrivere l'evoluzione temporale dei prezzi.

3.3 Random Walk Asset Pricing

Per ottenere una formula con la quale prezzare i titoli, o equivalentemente, per ottenere un modello di evoluzione temporale per i rendimenti dei medesimi, si segue quanto illustrato in [Requião da Cunha 2022]. Come punto di partenza, il migliore strumento matematico di cui si dispone è costituito dai *processi di Lévy*, ovvero delle particolari successioni di variabili aleatorie indicizzate dal tempo, siano esse W_t . Prima di procedere occorre definire uno spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ su cui si lavora, sia questo costituito rispettivamente dallo spazio campionario, dalla σ -algebra degli eventi e dalla misura di probabilità. I processi in esame sono caratterizzati dalle seguenti proprietà:

- il valore iniziale del processo è nullo, ossia $W_0 = 0$;
- gli incrementi sono indipendenti, ossia scelti due intervalli temporali disgiunti, cioè a intersezione nulla, siano essi $[t_0, t_1]$ e $[t_2, t_3]$, si ha che, definendo $X = W_{t_1} - W_{t_0}$ e $Y = W_{t_3} - W_{t_2}$, allora X e Y sono variabili aleatorie indipendenti;
- gli incrementi sono stazionari, ossia si ha che $W_{t_1} - W_{t_0} = W_{t_1-t_0}$;
- la misura di probabilità \mathbb{P} dello spazio di probabilità su cui si definisce il problema è continua, ossia

$$\lim_{\Delta t \rightarrow 0} \mathbb{P}(|W_{t+\Delta t} - W_t| > \varepsilon) = 0 \quad \forall \varepsilon > 0$$

Una classe particolare dei processi di Lévy sono i *processi di Wiener*. Essi hanno la particolarità di avere gli incrementi distribuiti con media nulla e con varianza data da $\sigma^2 = t - s$.

Per fissare le idee su alcuni concetti utili per meglio comprendere la formula che si adotterà per studiare come il filtro di Kalman si comporta con i titoli di borsa, si consideri per il momento un modello di mercato molto semplice, in cui il valore di un titolo rischioso indicizzato dal tempo e indicato con S_t , al tempo successivo può avere un incremento, ossia un ritorno maggiore di 1, indicato con u , oppure un decremento, ossia un ritorno compreso tra 0 e 1, indicato con d ; in generale, quindi, indicando con r_t il ritorno al tempo t sul titolo rischioso, si ha che

$$r_t \in \{u, d\} \tag{3.3}$$

e si possono verificare le due seguenti situazioni:

$$S_{t+1} = \begin{cases} uS_t & \text{se } r_t = u \\ dS_t & \text{se } r_t = d \end{cases} \tag{3.4}$$

In questo modo si può costruire il cosiddetto **albero binomiale**: indicando con p la probabilità che r_t sia pari a u si può costruire il grafico in Figura (3.3).

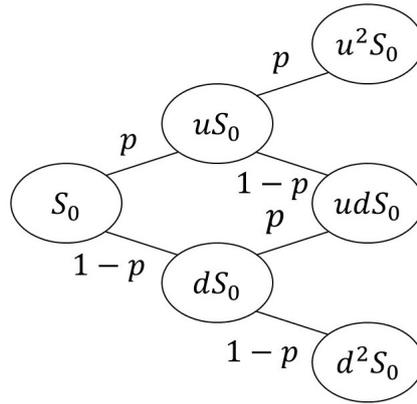


Figura (3.3): L'immagine rappresenta un albero binomiale, dove i nodi costituiscono i possibili valori ai tempi t_1, t_2, \dots del titolo rischioso, mentre gli archi raffigurano le possibili traiettorie che il titolo rischioso può seguire.

In riferimento alla Figura (3.3), per trovare la probabilità che S_t assuma un determinato valore, occorre moltiplicare, poiché si considerano i ritorni come variabili aleatorie indipendenti, le probabilità sugli archi della traiettoria seguita dal titolo rischioso. In generale si può scrivere

$$S_t = u^k d^{t-k} S_0 \quad (3.5)$$

per un determinato $k \in \mathbb{N}$. Quello che anche si evince dalla Figura (3.3) è che esiste più di una traiettoria tramite la quale si perviene al medesimo valore di S_t ; in altre parole, la funzione che va dallo spazio delle traiettorie allo spazio dei valori del titolo rischioso è una funzione non iniettiva. Per poter calcolare la molteplicità di tale funzione, ossia il numero di traiettorie che restituiscono il medesimo valore di S_t , occorre calcolare in quanti diversi modi si possono disporre k termini u e $t - k$ termini d nella sequenza di t termini che denotano la serie storica dei rendimenti: questo viene computato tramite il coefficiente binomiale. In altre parole, tale serie equivale a un campione di misure distribuite secondo una distribuzione binomiale.

$$\mathbb{P}(S_t = u^k d^{t-k} S_0) = \binom{t}{k} p^k (1-p)^{t-k} \quad (3.6)$$

dove la misura di probabilità \mathbb{P} viene quantificata dalla costante parametrica p , che può essere stimata a livello statistico eseguendo uno studio su un campione molto grande di dati e distribuiti in un intervallo di tempo sufficientemente ampio di tempo, in modo che la stima non sia inquinata dalla dipendenza dalle condizioni di mercato in un determinato periodo storico. Comunque già da questo fatto si evince quanto sia difficile trattare il pricing dei titoli, che rispecchia una realtà estremamente variegata e complessa, con un modello banale come quello binomiale.

Quanto visto è, a ogni modo, quello che accade in un modello di mercato molto semplice, dove gli incrementi e i decrementi sono tutti uguali. Risulta evidente, tuttavia, dall'esperienza, che gli incrementi possono variare tra loro e di quantità non trascurabili. Si rende quindi necessario implementare un modello matematico più realistico.

Prima però risulta necessario spendere qualche parola riguardo la capitalizzazione degli interessi. Nel modello che si andrà a costruire si adotta una capitalizzazione composta, ossia si considera che gli interessi maturino sugli interessi. Si può dimostrare² quindi che un titolo sicuro, come un titolo di stato o un mutuo bancario, evolve secondo la seguente legge:

$$B_t = B_0 e^{rt} \quad (3.7)$$

dove r è il tasso di interesse e B_0 è il valore dell'investimento. Tale processo stocastico viene utilizzato per attualizzare il valore di un ammontare di denaro nel futuro. In altre parole, il valore di un euro al tempo t futuro equivale a e^{-rt} euro nel presente: se si pone $B_0 = e^{-rt}$, infatti, al tempo t si ha che $B_t = 1$. La ragione per cui il valore del denaro cambia nel tempo è da ricercarsi nel fatto che al tempo presente è sempre possibile investire in titoli sicuri che restituiscono un ritorno piccolo ma probabilisticamente certo, in quanto l'evento F sul quale lo Stato dichiara default determinando il crollo del valore dei titoli del tesoro lo si considera, almeno sul breve periodo, di misura nulla.

Nell'ipotesi di mercati efficienti, come è stato ampiamente argomentato, il valore che assumono i rendimenti è casuale, ossia la serie storica dei rendimenti è un random walk. Si consideri un limite continuo di tempo e si supponga di attualizzare il valore del titolo rischioso tramite un titolo sicuro. Si ottiene quindi:

$$\lim_{\Delta t \rightarrow 0} S_{t+\Delta t} = \lim_{\Delta t \rightarrow 0} S_t e^{r\Delta t} \quad (3.8)$$

e applicando all'esponenziale lo sviluppo al primo ordine in serie di Taylor, dove l'espansione è limitata al primo ordine in quanto, supponendo che il titolo rischioso evolva secondo un random walk, da un istante di tempo al successivo si può considerare S_t come una funzione lineare del tempo, riscrivendolo quindi come $S(t)$, si ottiene

$$\lim_{\Delta t \rightarrow 0} S(t + \Delta t) = \lim_{\Delta t \rightarrow 0} S(t)(1 + r\Delta t) \quad (3.9)$$

Si perviene quindi alla seguente uguaglianza:

$$\lim_{\Delta t \rightarrow 0} (S(t + \Delta t) - S(t)) = \lim_{\Delta t \rightarrow 0} rS(t)\Delta t \quad (3.10)$$

ossia

$$dS(t) = rS(t)dt \quad (3.11)$$

Quindi ci si attende che il titolo rischioso abbia una componente di drift la cui ragione è da ricercarsi nella possibilità di un reinvestimento, teoricamente supposto continuo, del capitale in titoli sicuri. Inoltre, come già detto, si suppone che il titolo S abbia una componente di random walk, ossia

$$dS(t) \propto S(t)dW(t) \quad (3.12)$$

dove $W(t)$ è un moto browniano; in particolare si tratta di un moto browniano standard, ovvero tale per cui la distanza tra due punti consecutivi è, a meno di una costante, distribuita secondo una distribuzione normale standard, cioè con media nulla e varianza unitaria. Il titolo rischioso di borsa sarà quindi matematicamente rappresentato da un

²Per la dimostrazione di questo fatto si rimanda a [Pascucci 2009].

moto browniano geometrico, che equivale a un moto rettilineo uniforme in cui la traiettoria viene disturbata da un rumore browniano; in altre parole, esso dovrà soddisfare la seguente equazione differenziale:

$$dS(t) = \mu S(t)dt + \sigma S(t)dY(t) \quad (3.13)$$

dove si considerano il drift μ e la varianza σ come costanti, mentre dY denota il differenziale di una distribuzione normale standard. Questa rappresentazione del titolo è un processo di Wiener e costituisce il modello di asset pricing noto come **modello di Black-Scholes**. Il passaggio da fare ora, per ottenere una equazione utilizzabile a tempo discreto, di cui si necessita per testare il filtro di Kalman, è separare le variabili e integrare, ossia

$$\int_{S_t}^{S_{t+\Delta t}} \frac{dS(t')}{S(t')} = \mu \int_t^{t+\Delta t} dt' + \sigma \int_{Y_t}^{Y_{t+\Delta t}} dY(t') \quad (3.14)$$

dove si ha $\Delta t \rightarrow 0$; si ottiene quindi

$$\log\left(\frac{S_{t+\Delta t}}{S_t}\right) = \mu\Delta t + \sigma N_t \quad (3.15)$$

con $N_t \sim N(0; 1)$, che è giustificato ricordando che $Y(t')$ è un processo di variabili aleatorie tali che l'incremento tra due elementi consecutivi della serie storica sia una variabile aleatoria distribuita come una normale con media nulla e varianza unitaria; ovviamente è necessario indicizzare la variabile N_t perché essa varia a ogni istante temporale e costituisce proprio il rumore del moto browniano. Osservando l'espansione di $S(t + \Delta t)$ data dall'Equazione (3.10) e sostituendola nella precedente uguaglianza si ottiene

$$\log\left(\frac{S_t + r\Delta t}{S_t}\right) = \mu\Delta t + \sigma N \quad (3.16)$$

Considerando che Δt è molto piccolo, che r , denotando un tasso di interesse, è un parametro compreso tra 0 e 1 e sapendo che $\log(1 + x) \approx x$ per $x \approx 0$, si ha che

$$r\Delta t = \mu S_t \Delta t + \sigma N \quad (3.17)$$

che equivale, sempre per l'Equazione (3.10), alla seguente uguaglianza:

$$S_{t+1} = S_t + \mu S_t \Delta t + \sigma S_t N_t \quad (3.18)$$

Da questa equazione si può ricavare la relazione di riferimento per i ritorni sui titoli azionari, ricordando l'Equazione (3.1):

$$r_{t+1} = \mu\Delta t + \sigma N_t \quad (3.19)$$

Si può quindi facilmente osservare come ci si sia ricondotti a una situazione analoga al moto rettilineo uniforme dove la traiettoria rettilinea dei valori veri viene disturbata da un elemento di rumore esterno.

3.4 Il filtro di Kalman e i titoli di borsa

In riferimento dell'Equazione (3.19) si vede che la traiettoria che ci si attende sia seguita dai ritorni sui titoli azionari è un moto rettilineo uniforme in cui la traiettoria rettilinea viene disturbata da un rumore costituito da un moto browniano. Per poter effettuare le simulazioni e testare dunque il filtro di Kalman occorre stimare i due parametri μ e σ e in particolare, perché sia rispettata l'ipotesi di mercati efficienti, ci si attende che la costante di drift sia compatibile con lo zero. Il primo parametro viene ricavato effettuando una regressione lineare sulla serie storica dei rendimenti, il secondo per definizione è la relativa deviazione standard, dunque facilmente calcolabile. Si evidenzia, per mantenere un buon livello di correttezza formale, che sia la deviazione standard sia la costante di drift sono grandezze adimensionali, in quanto sono calcolate rispetto ai rendimenti, definiti come rapporto tra due termini aventi stessa unità di misura e quindi adimensionali. In questo caso particolare si avrà che la misura e il valore vero, a differenza delle simulazioni effettuate in precedenza, sono uguali, cioè non si ha errore sulla misura.

3.4.1 Metodologia

Il filtro di Kalman è stato applicato a due serie storiche di rendimenti, una relativa a un solo titolo azionario, in particolare si è scelto quello emesso da *Walmart Inc.*, che sarà nel seguito indicata con S^1 e l'altra relativa a una serie storica risultante da una media dei rendimenti normalizzati, perché siano tra loro confrontabili, di dieci titoli azionari, che sarà indicata nel seguito con S^2 ; in questo modo si è ipotizzato di poter osservare quali sono gli effetti dovuti a un bacino maggiore di titoli e di poter constatare se tale diversificazione fosse in qualche modo legata all'efficienza del mercato. I rendimenti sono stati calcolati su serie di valori risultanti dalla media del prezzo delle azioni di apertura e del prezzo di chiusura. Dunque, sono stati dapprima raccolti i dati relativi al periodo compreso tra il 25 settembre 2019 e il 25 settembre 2022 con cadenza settimanale dei seguenti emittenti scelti tra quelli quotati nella Borsa di Wall Street: *Walmart Inc.*, *The Boeing Company*, *Visa Inc.*, *Berkshire Hathaway Inc.*, *The Walt Disney Company*, *Amazon.com Inc.*, *ExxonMobil Corporation*, *The Coca-Cola Company*, *Adobe Inc.* e *JPMorgan Chase & Co.*; di ciascuno di questi titoli sono state calcolate le serie temporali dei ritorni, rappresentate in Figura (3.4).

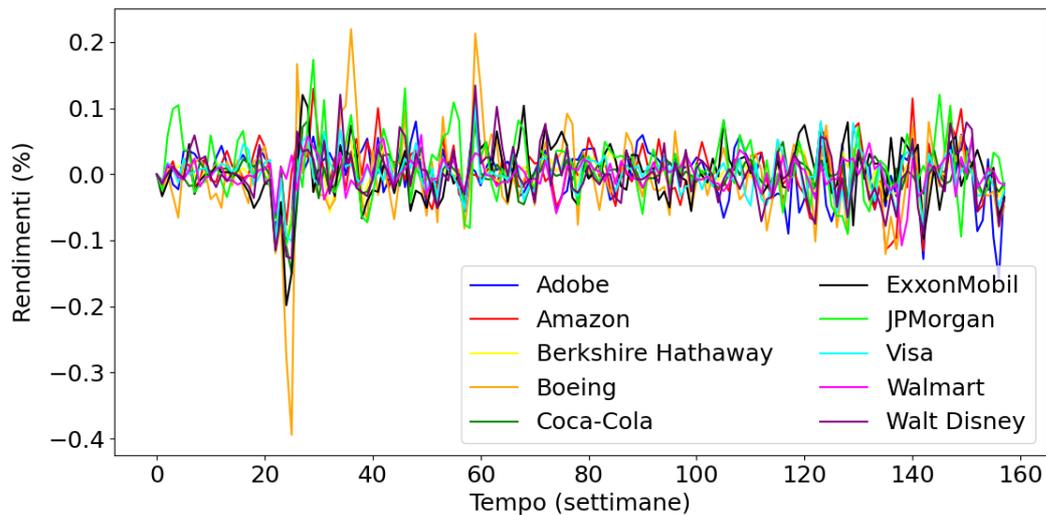


Figura (3.4): L'immagine rappresenta, per avere una visione di insieme e per dimostrare quindi, almeno visivamente, che i titoli azionari scelti sono sufficientemente non correlati tra loro, i rendimenti percentuali dei medesimi in funzione del tempo, espresso in settimane.

Fonte dei dati: <https://finance.yahoo.com/>

Come si osserva dall'immagine appena riportata, tuttavia, i rendimenti di alcuni titoli, come quelli emessi da *Boeing*, subiscono nel corso del tempo delle variazioni nettamente maggiori rispetto ad altri, come *Walmart*. Finché si lavora con la serie di un solo titolo non si pongono problemi, ma i valori ottenuti non vanno bene per la costruzione di S^2 , in quanto i titoli scelti non sono di principio confrontabili e si rende dunque necessario normalizzarli: le serie temporali dei rendimenti sono state dunque divise per la loro deviazione standard, in modo da ottenere delle serie temporali aventi tutte deviazione standard unitaria e quindi confrontabili. In questo modo, per di più, si pone, per la serie storica su cui verrà applicato il filtro di Kalman, $\sigma = 1$, in accordo con la proprietà dei processi di Wiener di avere deviazione standard pari all'intervallo temporale su cui si effettua il campionamento.

Il grafico risultante è mostrato in Figura (3.5). Nel caso della serie S^2 è importante sottolineare che, mediando i ritorni dei dieci titoli scelti, si ottiene una deviazione standard sui ritorni diversa dal valore unitario: in generale, infatti, nulla impone che la media di serie temporali aventi deviazione standard pari a un determinato valore, abbia il medesimo valore. Dunque, per ottenere una maggiore uniformità, i ritorni della serie S^2 sono stati rinormalizzati, dividendo ciascuno di essi per la deviazione standard, in modo da ottenere una serie temporale avente $\sigma = 1$.

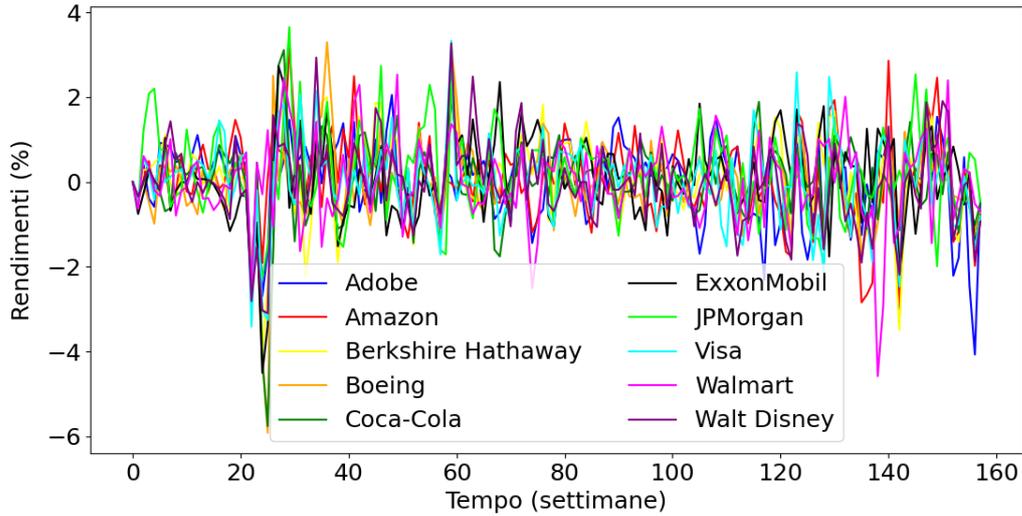


Figura (3.5): L'immagine rappresenta le serie storiche dei rendimenti normalizzati in modo che abbiano tutti varianza unitaria e siano dunque confrontabili tra loro.

Per poter stimare il parametro μ , come anticipato, si effettua una regressione lineare, applicando le seguenti formule tratte da [Fornasini 2008], dove si pone $y = A + Bx$ e dove dunque B indica il coefficiente angolare e A l'intercetta all'origine, con N che indica il numero delle misure effettuate:

$$A = \frac{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} \quad (3.20)$$

$$B = \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} \quad (3.21)$$

Gli errori dati dalle seguenti formule:

$$\sigma_A = \sqrt{\frac{1}{N-2} \sum_{i=1}^N (y_i - A - Bx_i)^2} \sqrt{\frac{\sum_{i=1}^N x_i^2}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}} \quad (3.22)$$

$$\sigma_B = \sqrt{\frac{1}{N-2} \sum_{i=1}^N (y_i - A - Bx_i)^2} \sqrt{\frac{N}{N \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2}} \quad (3.23)$$

Nel caso in esame le x sono le settimane, cioè gli istanti di tempo. Chiaramente ci si attende che sia $A = 0$ per entrambe le serie storiche, dal momento che al tempo iniziale non si ha rendimento, ovvero lo si pone uguale a zero; i valori ottenuti dall'applicazione delle formule sopra riportate alle serie costruite sono:

$$A(S^1) = 0.15 \pm 0.16$$

dunque compatibile con il valore atteso;

$$A(S^2) = 0.24 \pm 0.16$$

che non risulta, per una ragione spiegata nel seguito, compatibile con il valore nullo atteso. Venendo poi alla costante di drift, si ottiene

$$\mu(S^1) = -0.0014 \pm 0.0017$$

$$\mu(S^2) = -0.0021 \pm 0.0037$$

entrambi compatibili, come ci si attende, con il valore nullo atteso; infatti nell'ipotesi di mercati efficienti i rendimenti si muovono secondo un moto browniano semplice, ossia secondo un moto browniano geometrico con costante di drift nulla; d'altro canto il periodo di tempo coperto è sufficientemente lungo perché il valore delle azioni non sia condizionato da circostanze particolari di mercato, che possono determinare il comportamento temporaneamente non efficiente di un mercato, come l'annuncio riguardo gli utili dell'azienda, ma al contempo sufficientemente breve per non risentire della crescita che sul lungo periodo ci si attende dai mercati, dovuta a fenomeni come l'inflazione e al fatto che, banalmente, le aziende non virtuose escono dai mercati in seguito a fallimenti o acquisizioni e i titoli disponibili sono dunque quelli emessi da enti finanziari sul lungo periodo in crescita.

Come si osserva, la costante di drift ottenuta per S^2 è, in valore assoluto, maggiore di quella ottenuta per S^1 ; graficamente si ha una linea di regressione lineare maggiormente inclinata verso il basso per S^2 rispetto a S^1 e ciò è dovuto al fatto che nella media effettuata per ottenere la serie S^2 sono stati considerati anche i rendimenti di titoli azionari, come quelli di *Boeing* che hanno avuto una maggiore decrescita nel periodo considerato. Per questa ragione, se si traccia la linea di regressione a partire dal valore finale della serie, si ottiene un valore al tempo iniziale che è compreso nell'intervallo $[0, 1]$, con un margine di errore statistico che non comprende il valore nullo: ciò spiega, a livello puramente intuitivo, perché il valore $A(S^2)$ non è compatibile con il valore atteso. Tuttavia, nelle applicazioni di interesse per la trattazione in esame, ciò non ha un impatto rilevante, per cui si considera tale problema trascurabile.

3.4.2 Risultati

Prima di procedere alle applicazioni del filtro di Kalman, sono stati inseriti i rendimenti che costituiscono le due serie storiche in esame in istogrammi per verificare anche graficamente, innanzitutto, che i valori fossero distribuiti secondo una distribuzione normale. Il risultato grafico è mostrato in Figura (3.6).

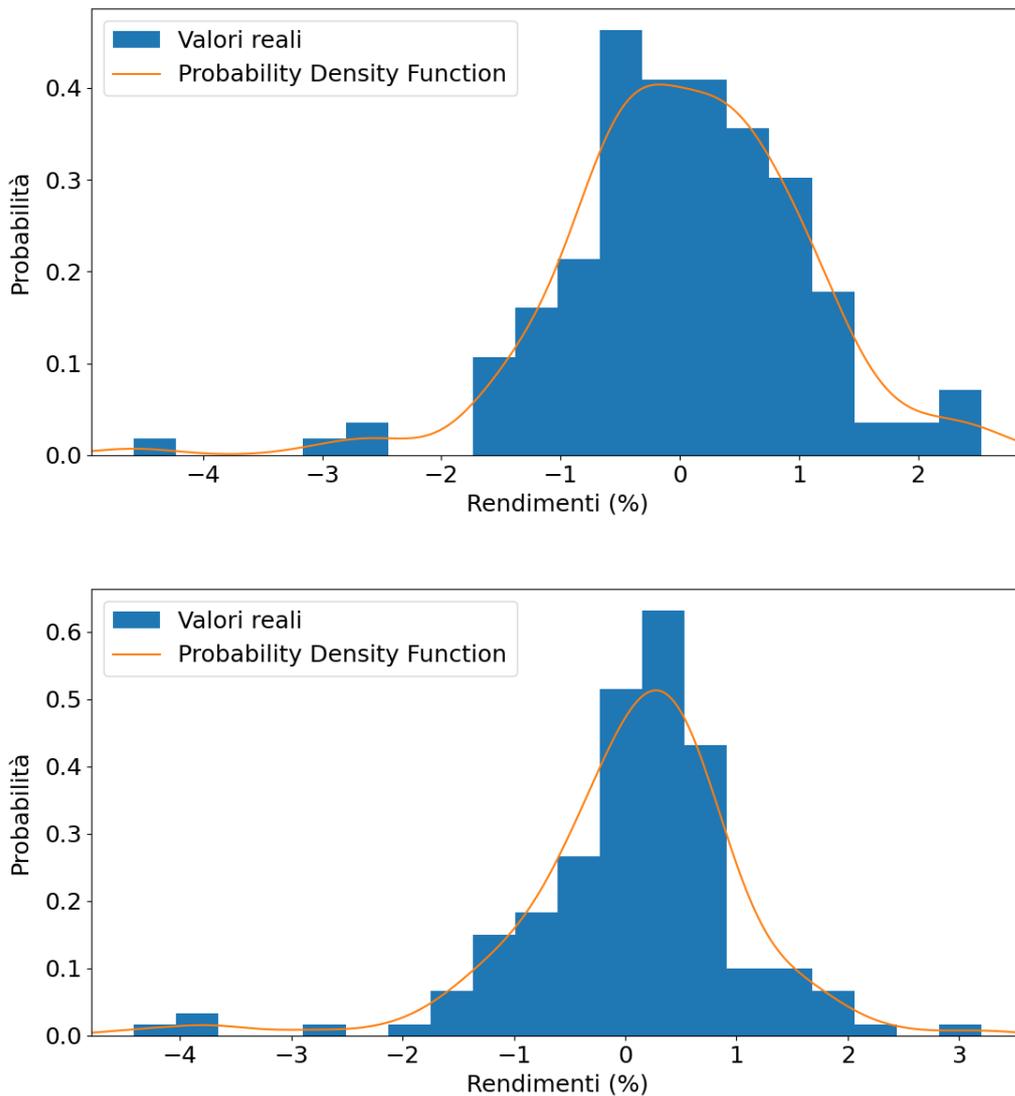


Figura (3.6): L'immagine mostra due diversi istogrammi riempiti rispettivamente quello in alto con i dati della serie storica S^1 e quello in basso con i dati della serie storica S^2 . Gli istogrammi sono stati costruiti suddividendo i valori sull'asse delle ascisse in 20 bin, un numero non troppo piccolo perché non emerga la forma caratteristica dell'istogramma stesso, ma nemmeno troppo elevato e che permette quindi di accorpate le piccole differenze tra valori molto simili di rendimento. Gli istogrammi sono sovrapposti alla curva della Probability Density Function generata a partire dai dati in esame.
 Fonte dei dati: <https://finance.yahoo.com/>

Come si può osservare dai grafici appena mostrati, lavorando con una media dei rendimenti di dieci titoli azionari anziché con uno solo, la forma gaussiana dell'istogramma è più evidente, giacché mediando su più valori si livellano le anomalie che caratterizzano i singoli titoli e che sono dovute a episodi particolari che riguardano le rispettive aziende emittenti; d'altro canto, guardando alle Probability Density Function sovrapposte in entrambi i casi agli istogrammi, si constata che quella generata dai dati della serie S^2 è

più simile a una gaussiana, rispetto a quella generata dai dati della serie S^1 : ciò dimostra, come già anticipato da una semplice ipotesi di lavoro, che il mercato si comporta in maniera più efficiente nei confronti di una media di rendimenti rispetto al caso del rendimento di un unico titolo.

Per poter quantificare ulteriormente il comportamento efficiente del mercato, si applica quindi, a entrambe le serie storiche, il filtro di Kalman. Poiché, come già spiegato, il modello di riferimento è sostanzialmente analogo a un moto rettilineo uniforme, si deve fare riferimento al framework teorico precedentemente mostrato a riguardo, adeguatamente sostituendo la posizione con il rendimento e la velocità con le costanti di drift calcolate. In particolare, sia per la serie S^1 sia per la S^2 , la deviazione standard dei rendimenti è $\sigma_r = 1$, indicata così per non confonderla con l'errore statistico sulla costante di drift; di conseguenza si deve avere

$$\lambda \frac{\Delta t^3}{3} = 1 \quad (3.24)$$

ed essendo $\Delta t = 1$ in quanto si considera la variazione unitaria di settimane, si ha

$$\lambda = 3 \quad (3.25)$$

e questo implica che

$$\sigma_\mu = \lambda \Delta t = 3 \quad (3.26)$$

Ciò dimostra come il modello assuma che ci possa essere una variazione statistica sulla costante di drift molto maggiore di quella che si ha nella realtà e ciò rende il fatto che si siano ottenute costanti di drift non nulle ancora più accettabile. Inoltre, poiché, come già detto, non esiste errore sulla misura, ovvero la misura coincide con il valore vero, osservando l'Equazione (1.58) e l'Equazione (1.64), ci si aspetta che il Kalman gain sui rendimenti sia costantemente unitario e di conseguenza che il valore aggiornato equivalga sempre al valore vero. Infatti si ha $\sigma_\varepsilon = 0$ costantemente, in quanto l'errore sulla misura è nullo. Di conseguenza

$$\mathbf{K}_{t+1}^{(1)} = \mathbf{H} \cdot \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T (\mathbf{H} \cdot \mathbf{P}_{t+1|t} \cdot \mathbf{H}^T)^{-1} = 1 \quad (3.27)$$

e quindi, ricordando che la misura equivale al valore vero,

$$\boldsymbol{\mu}_{t+1|t+1} = \boldsymbol{\mu}_{t+1|t} + (1, 0) \cdot (\mathbf{x}_{t+1} - \mathbf{H} \cdot \boldsymbol{\mu}_{t+1|t+1}) = \mathbf{x}_{t+1} \quad (3.28)$$

Per questa stessa ragione la varianza aggiornata sui rendimenti è sempre nulla, dal momento che, nel caso in esame,

$$\mathbf{I} - \mathbf{K}_{t+1} \cdot \mathbf{H} = \begin{pmatrix} 0 & 0 \\ -\mathbf{K}_{t+1}^{(2)} & 1 \end{pmatrix} \quad (3.29)$$

per cui la prima componente della matrice risultante dal prodotto dell'operatore sopra mostrato con la matrice di covarianza prevista è sempre nulla.

Il grafico risultante dall'applicazione del filtro di Kalman alle serie è riportato in Figura (3.7).

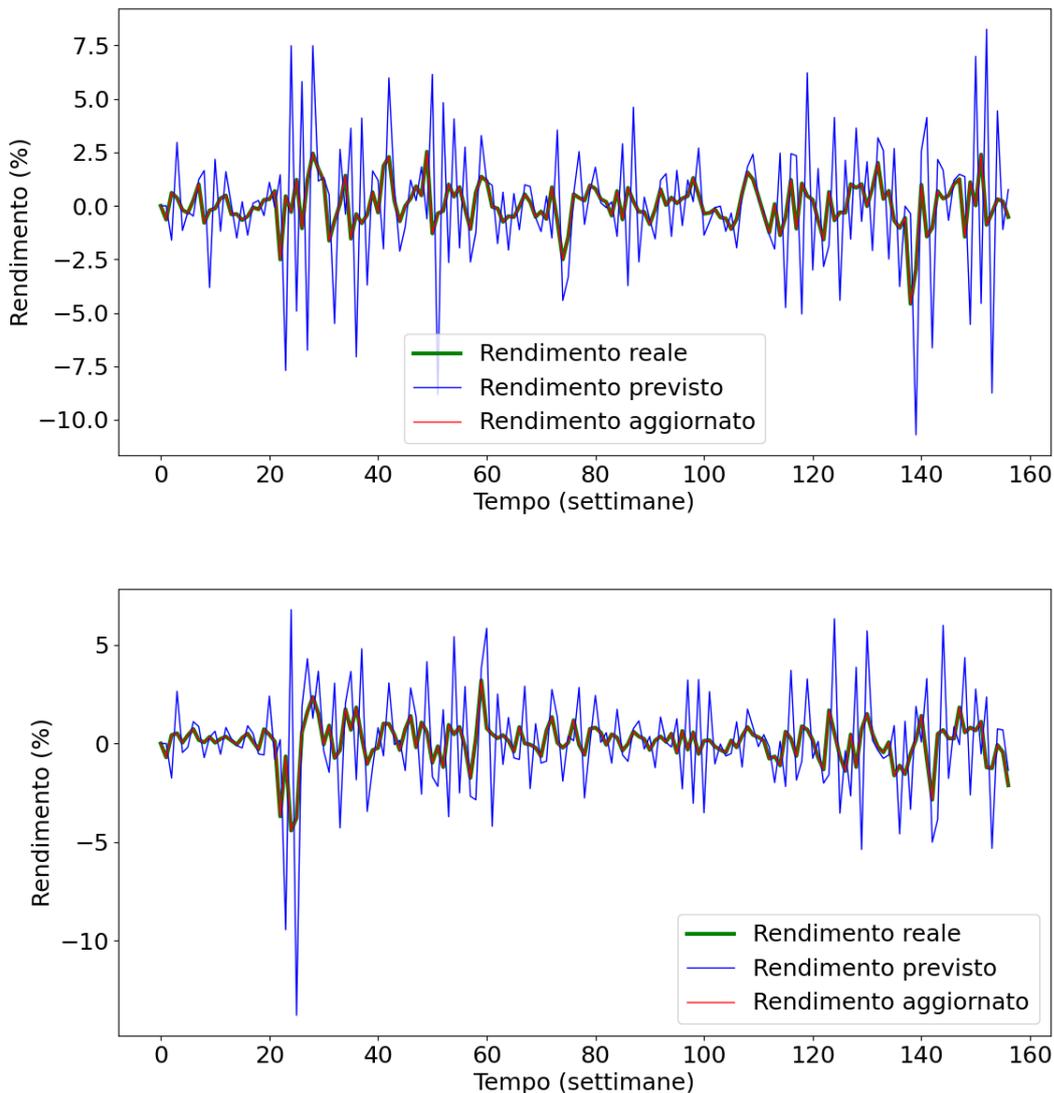


Figura (3.7): L'immagine in alto rappresenta la serie temporale S^1 , mentre quella in basso la serie S^2 , entrambe sovrapposte con le previsioni effettuate sui relativi dati tramite il filtro di Kalman e con gli aggiornamenti.

Nei grafici sopra mostrati si può osservare il peculiare comportamento del rendimento previsto: poiché la costante di drift è molto ridotta, quando si hanno delle oscillazioni è come se l'algoritmo, istante per istante, ipotizzasse l'esistenza di un trend che esula dal modello lineare e tentasse di seguirlo effettuando una previsione al tempo successivo. In parole semplici, quando si ha, per esempio, un decremento del rendimento, l'algoritmo ipotizza che vi sia un decremento anche all'istante successivo, come se i rendimenti potessero muoversi secondo una traiettoria diretta verso il basso, seguendo quindi un trend decrescente fino al tempo in cui si verifica invece un aumento, a partire dal quale si correggono le previsioni, ipotizzando che la direzione seguita dalla traiettoria prosegua verso l'alto e via così per tutte le fluttuazioni attorno alla traiettoria media dei rendimenti, che è lineare con una pendenza quasi nulla.

Per mostrare, in conclusione, una applicazione rudimentale di questo studio sui dati, si supponga di adottare una banale strategia di investimento, in cui si acquista al tempo $t = 0$ un titolo rischioso, vendendolo al tempo finale $t = T$. L'idea è quella di ottenere, a partire dalla serie dei rendimenti previsti, la serie dei valori previsti dei titoli. Si consideri per il momento la serie dei titoli di *Walmart*. Nella simulazione precedentemente effettuata, i rendimenti previsti sono stati ottenuti a partire dai rendimenti normalizzati; per un confronto tra il valore reale del titolo e quello previsto sarebbe opportuno applicare il filtro di Kalman alla serie dei rendimenti non normalizzati, sia essa indicata con S'^1 ; si ottiene quindi una nuova serie dei rendimenti previsti, grazie alla quale, applicando la formula ricorsiva

$$S_{t+1} = (1 + r_t)S_t \quad (3.30)$$

a partire dal valore iniziale S_0 , si ricava la serie dei valori previsti dei titoli. Chiaramente il valore iniziale sarà il medesimo per entrambe le serie, cioè sia per quella dei valori reali sia per quella dei valori previsti, in quanto al tempo $t = 0$ non è possibile effettuare una previsione, ovvero essa coincide con il valore reale. Per i titoli di *Walmart* si ottengono, indicando con T il tempo finale, i dati riportati in Tabella (3.1). Nella tabella si è riservata una colonna in cui si sono calcolati i valori reale e previsto al tempo finale nel caso in cui il valore iniziale venga normalizzato a 100\$, per un migliore confronto visivo con quanto verrà fatto per la serie S^2 .

	Valori non normalizzati	Valori normalizzati
$S_0(\text{\$})$	118.81	100.00
$S_{T, \text{reale}}(\text{\$})$	130.97	110.24
$S_{T, \text{previsto}}(\text{\$})$	106.27	89.45

Tabella (3.1): La tabella mostra i valori reale e previsto al tempo finale e il valore al tempo iniziale sia normalizzati sia non normalizzati ottenuti a partire dalla serie S'^1 .

Si sarebbero potuti ottenere comunque i medesimi valori anche a partire dalla serie S^1 , ossia la serie dei rendimenti normalizzati, moltiplicando la serie dei rendimenti previsti quindi ottenuta per la deviazione standard dei dati della serie temporale dei rendimenti normalizzati. Per quanto riguarda, invece, la serie dei rendimenti S^2 , per come è stata costruita, non è possibile effettuare una operazione analoga, in quanto non è possibile riottenere i rendimenti previsti marginali dei dieci titoli a partire dal rendimento congiunto. Si può comunque procedere in una maniera formalmente meno corretta, ma che permette di visualizzare una applicazione anche nel caso di una serie costruita a partire da più titoli. Quello che si può fare è normalizzare i valori dei titoli a un comune valore iniziale, sia esso $S_0 = 100.00\text{\$}$, mediare quindi i valori dei titoli normalizzati e ricavare la serie dei rendimenti a partire dalla media dei valori dei titoli normalizzati, sia questa serie S'^2 . In altre parole, rispetto al processo di costruzione di S^2 le procedure di ricavo dei rendimenti e normalizzazione sono scambiate. In Figura (3.8) è riportato per completezza il grafico che mostra l'andamento temporale della serie S'^2 e delle serie dei rendimenti previsti e aggiornati risultanti dall'applicazione del filtro di Kalman alla serie S'^2 .

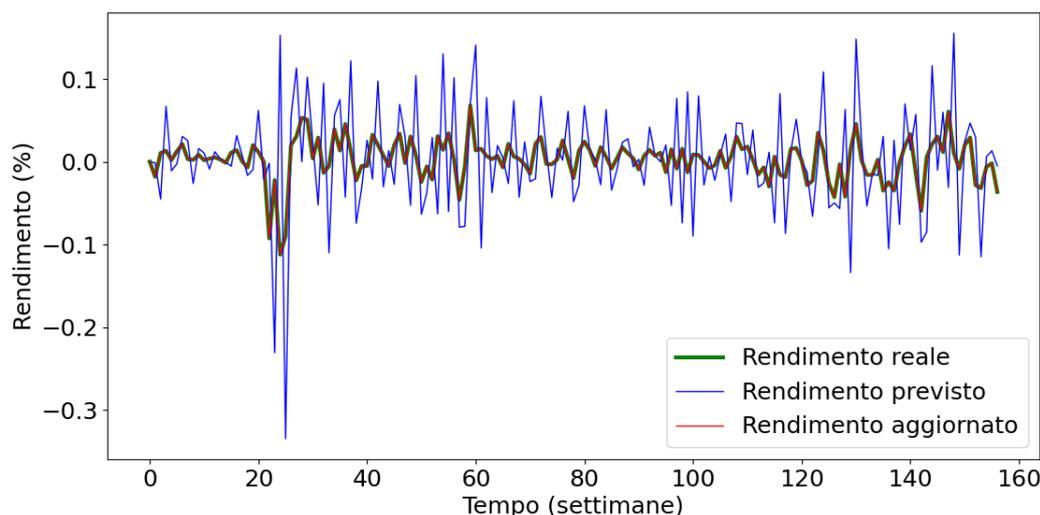


Figura (3.8): L'immagine rappresenta l'andamento temporale della serie dei rendimenti S^{r2} e, in seguito all'applicazione a questa del filtro di Kalman, delle serie dei rendimenti previsti e aggiornati quindi ottenute.

Fonte dei dati: <https://finance.yahoo.com/>

Valgono per questo grafico le considerazioni riportate anche per la Figura (3.7). Tramite la simulazione effettuata per ottenere quest'ultima, tuttavia, si perviene ai risultati riportati in Tabella (3.2), dove chiaramente non è possibile operare una distinzione tra valori normalizzati e non normalizzati.

	Valori normalizzati
$S_0(\text{\$})$	100.00
$S_{T, \text{reale}}(\text{\$})$	136.08
$S_{T, \text{previsto}}(\text{\$})$	116.83

Tabella (3.1): La tabella mostra i valori reale e previsto al tempo finale e il valore al tempo iniziale ottenuti a partire dalla serie S^{r2} .

In entrambi i casi, sia per il titolo singolo sia per la media di più titoli, tramite il filtro di Kalman si ottiene una sottostima del valore effettivo al tempo finale del titolo, con un esito tuttavia profondamente diverso nei due casi: infatti, mentre per la serie S^2 si prevede comunque una crescita del valore del titolo, anche se minore di quella che si ha nella realtà, per la serie S^1 il filtro di Kalman prevede una decrescita del valore del titolo. Questo risultato supporta un'ipotesi introdotta nel corso del presente capitolo, ossia che un titolo singolo si comporta in maniera meno efficiente rispetto a un titolo costruito come media di più titoli. Comunque, per quanto effettivamente ci sia un malfunzionamento del filtro di Kalman per i rendimenti dei titoli emessi da *Walmart*, si tratta di un errore non eccessivamente grave, dal momento che in finanza è sempre meglio non sfruttare una opportunità di guadagno, senza tuttavia incorrere in perdite, piuttosto che sovrastimare una opportunità di guadagno, con il rischio di incorrere in perdite. I casi presentati in questo sottoparagrafo riguardano tuttavia due serie di rendimenti di

titoli che hanno valore reale complessivamente crescente nel periodo considerato, per cui potrebbe sorgere spontanea la seguente domanda: come si comporta il filtro di Kalman nel caso di titoli che abbiano valore complessivamente decrescente?

3.4.3 Approfondimento

Per completezza, come approfondimento a quanto mostrato nel precedente sottoparagrafo, si costruisce una serie di rendimenti non normalizzati S^3 , seguendo lo stesso procedimento adottato per la costruzione della serie S^1 , in accordo con la teoria precedentemente introdotta, a partire dai titoli emessi da *Boeing*, i quali assumono valori complessivamente decrescenti. Per la serie trattata in questo sottoparagrafo si è ottenuto come valore della costante di drift

$$\mu(S^3) = 0.0002 \pm 0.0017$$

che è molto prossimo allo zero, per cui ci si aspetta da questa serie di rendimenti un comportamento estremamente efficiente.

In Figura (3.9) è riportato l'istogramma riempito con gli elementi della serie S^3 e il grafico che mostra l'andamento temporale dei valori veri della medesima e delle serie dei rendimenti previsti e aggiornati ottenute tramite l'applicazione del filtro di Kalman alla serie S^3 .

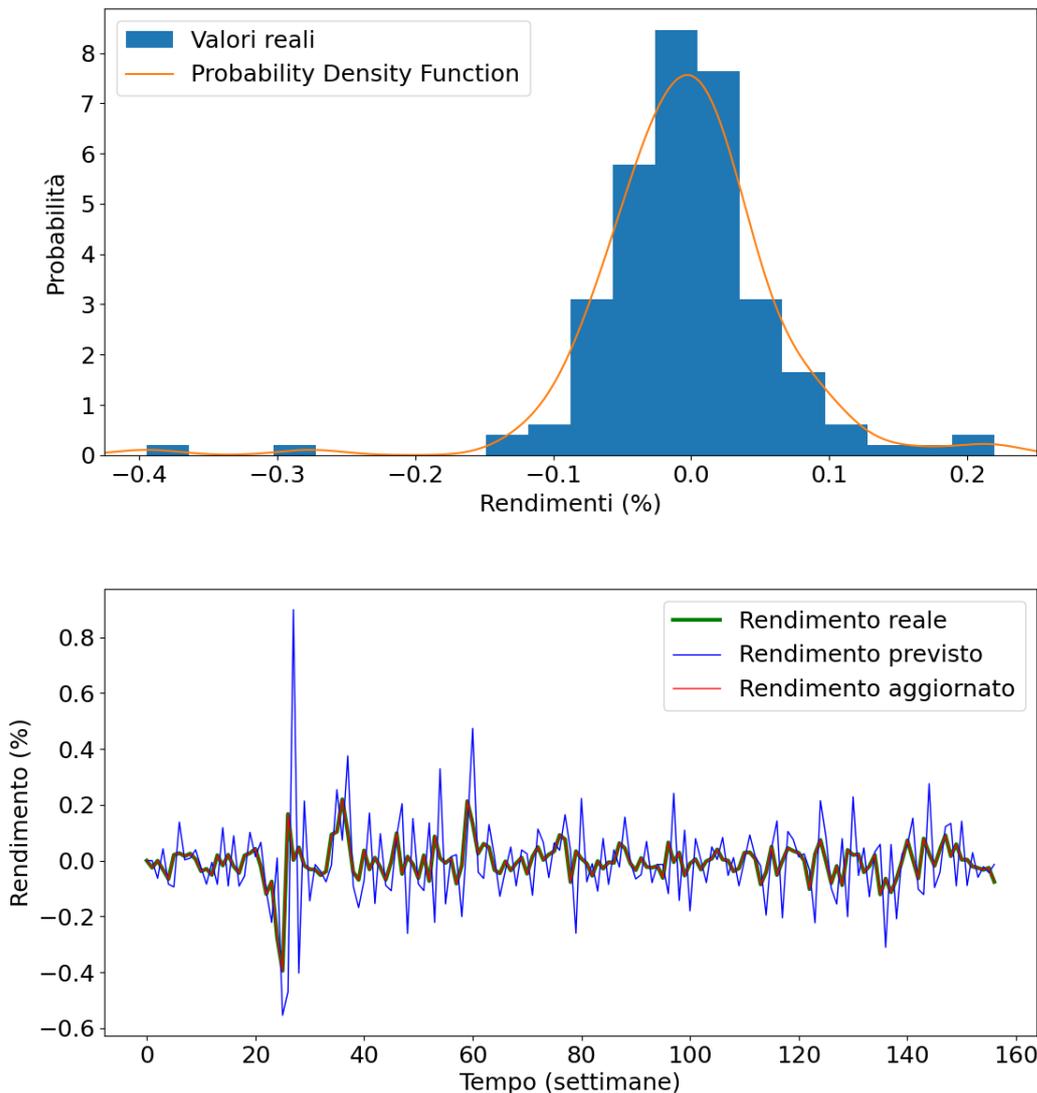


Figura (3.9): In alto è riportato l'istogramma che mostra la distribuzione dei rendimenti della serie S^3 , in basso è riportato l'andamento temporale dei valori veri dei rendimenti della stessa serie, sovrapposto all'andamento delle serie dei rendimenti previsti e aggiornati ottenuti con il filtro di Kalman.

Dall'istogramma si osserva che la distribuzione dei rendimenti ottenuti dai valori dei titoli emessi da *Boeing* è molto simile a una gaussiana centrata in zero, a riprova del comportamento molto efficiente della serie in esame. Dal grafico si osserva che il rendimento previsto, a parte in alcuni tratti in cui si hanno delle forti fluttuazioni del rendimento reale, si adegua, come anche nei casi precedenti, abbastanza bene all'andamento di quest'ultimo, con delle fluttuazioni maggiori che sono date proprio dal fatto che l'errore assunto dal filtro sulla costante di drift è molto maggiore di quello che si ha statisticamente, per cui, siccome i valori reali oscillano molto, le previsioni si adeguano assumendo delle grandi variazioni della costante di drift.

Costruendo infine, secondo la stessa procedura adottata in precedenza, la serie dei

valori previsti a partire da quella dei valori previsti, si sono ottenuti i valori riportati in Tabella (3.3).

	Valori non normalizzati	Valori normalizzati
$S_0(\text{\$})$	383.15	100.00
$S_{T, \text{reale}}(\text{\$})$	132.84	34.67
$S_{T, \text{previsto}}(\text{\$})$	29.19	7.62

Tabella (3.1): La tabella mostra i valori reale e previsto al tempo finale e il valore al tempo iniziale sia normalizzati sia non normalizzati ottenuti a partire dalla serie S^1 .

Si osserva che anche in questo caso il valore finale previsto tramite il filtro di Kalman è molto inferiore al valore reale, anche se si tratta di un risultato tutto sommato accettabile, in quanto l'importante è che, per la serie in esame, il valore finale non venisse sovrastimato. Una ipotesi che possa spiegare perché il filtro di Kalman tenda a fornire stime inferiori ai valori reali è che, avendo comunque coperto con il campionamento un periodo non brevissimo, i rendimenti si comportino di fatto in maniera meno efficiente dell'andamento che potrebbero assumere sul breve periodo. Tuttavia, per avere maggiore consapevolezza e completezza di informazioni riguardo il funzionamento e l'applicabilità del filtro di Kalman al settore finanziario si dovrebbero adottare strategie di investimento ben più complesse, per esempio inserendo delle transazioni tra il tempo iniziale e finale, in modo da poter valutare l'efficacia delle previsioni in maniera più corretta, ma ciò esula dall'ambito di studio di questa tesi.

Conclusione

In questo lavoro come argomento principale è stato presentato il filtro di Kalman, un algoritmo che permette, conoscendo la misura di una grandezza in esame al tempo t , di effettuare una previsione del valore vero al tempo $t + 1$ e di aggiornare tale previsione una volta effettuata la $(t + 1)$ -esima misura.

Si è testato il funzionamento di tale strumento prima applicandolo a punti materiali soggetti a moti fisici, osservando sempre come la misura aggiornata venga stimata tramite una media pesata della posizione misurata e della posizione prevista e come quest'ultima tenda a seguire una traiettoria costituita da prolungamenti, secondo il modello adottato, del percorso seguito nell'intervallo temporale precedente dalla posizione aggiornata; si è contestualmente osservato il comportamento del Kalman gain, che si stabilizza già dopo poche iterazioni dell'algoritmo su un valore limite, una volta che anche le matrici di covarianza prevista e aggiornata si sono stabilizzate su valori asintotici.

Infine si è introdotta l'ipotesi dei mercati efficienti e si è mostrato il modello di Black-Scholes, con il quale si è ricondotta l'evoluzione temporale dei titoli di borsa a un moto browniano geometrico, ossia un moto rettilineo in cui la traiettoria è disturbata da un rumore browniano. Assumendo questo modello, infine, a partire dal valore dei titoli emessi da diverse aziende quotate in borsa, tramite il filtro di Kalman si è previsto il valore al tempo finale dei titoli ottenendo un risultato tutto sommato buono nel caso di una serie costruita come media di più titoli, mentre nel caso di un titolo singolo si è stimato un valore in contrasto con il prezzo reale, dimostrando sia che il rendimento dei valori mediati di più titoli si comporta in maniera più efficiente rispetto a un singolo titolo, sia come, per strategie di investimento eccessivamente semplici, l'algoritmo in esame funziona in maniera solo parzialmente corretta anche per l'evoluzione temporale di titoli di borsa.

Appendice

Prima di mostrare i codici delle simulazioni effettuate, si fornisce la soluzione al quesito posto contestualmente all'illustrazione dell'esperimento di Malkiel in Figura (3.2): l'andamento dell'indice *Standard&Poors500* è rappresentato nel grafico in basso, mentre l'andamento dell'azione la cui evoluzione è determinata dal lancio del dado è rappresentato nel grafico in alto.

Moto rettilineo uniforme

```
import numpy as n
import matplotlib.pyplot as plt

# COSTANTI E SUCCESSIONI
# successione dei vettori di stato
x_ = []
# successione delle misure
y_ = []
# successione delle previsioni dei vettori di stato
x_pred_ = []
# successione delle previsioni delle matrici di covarianza
P_pred_ = []
# successione degli aggiornamenti dei vettori di stato
x_up_ = []
# successione degli aggiornamenti delle matrici di covarianza
P_up_ = []
# successione delle innovazioni
N_ = []
# successione del Kalman gain
K_ = []
# intervallo temporale
deltaT = 0.2
# costante del rumore browniano
c = 30
# varianza della distribuzione dell'errore sulla posizione
sigma_wq = c * (deltaT ** 3) / 3
# varianza della distribuzione dell'errore sulla velocità
sigma_wv = c * deltaT
```

```

# matrice della transizione di stato
F = a.array([[1, deltaT], [0, 1]])
# matrice di errore sulla covarianza
Q = c * a.array([[ (deltaT ** 3) / 3, (deltaT ** 2) / 2 ],
                 [ (deltaT ** 2) / 2, deltaT]])
# deviazione standard dell'errore sulla misura
sigma_eta = 2
# matrice dell'errore sulla misura
R = sigma_eta
# vettore di transizione in uscita
H = a.array([1, 0])
# velocità
v = 5

# STATI INIZIALI
# si pone nulla la matrice di covarianza iniziale
x_pred = a.array([0, v]).T
x_.append(x_pred)
y_.append(x_pred[0])
x_pred_.append(x_pred)
P_pred = a.array([[0, 0], [0, 0]])
P_pred_.append(P_pred)
x_up = a.array([0, v])
x_up_.append(x_up)
P_up = a.array([[0, 0], [0, 0]])
P_up_.append(P_up)
P = P_up

# ITERAZIONI E APPLICAZIONE DEL FILTRO
for i in range(1, 100):
    x = a.array([i * deltaT * v, v]).T
    w_q = a.random.normal(0, a.sqrt(sigma_wq))
    w_v = a.random.normal(0, a.sqrt(sigma_wv))
    w = a.array([w_q, w_v]).T
    x_.append(x + w)
    # previsione
    x_pred = a.dot(F, x_up)
    x_pred_.append(x_pred)
    P_pred = a.dot(F, a.dot(P, F.T)) + Q
    P_pred_.append(P_pred)
    z = a.random.normal(0, sigma_eta)
    y = a.dot(H, x + w) + z
    y_.append(y)
    # innovazione
    n = y - a.dot(H, x_pred)
    N_.append(n)

```

```

# varianza dell'innovazione
S = sigma_eta + a.dot(H, a.dot(P_pred, H.T))
# Kalman gain
K = a.dot(P_pred, H.T) / S
K_.append(K)
# aggiornamento
x_up = x_pred + K * n
x_up_.append(x_up)
M = a.array([[1, 0], [0, 1]]) - a.outer(K, H)
P_up = a.dot(M, a.dot(P_pred, M.T)) + \
        a.dot(K, a.dot(R, K.T))
P_up_.append(P_up)
P = P_up
print(P_up)

# GRAFICO POSIZIONI
x_print, = plt.plot(a.transpose(x_) [0], label="Posizione reale",
                    color='green', linewidth=1)
y_print, = plt.plot(y_, label="Posizione misurata",
                    color='black', linewidth=1)
x_pred_print, = plt.plot(a.transpose(x_pred_) [0],
                         label="Posizione prevista",
                         color='blue', linewidth=1)
x_up_print, = plt.plot(a.transpose(x_up_) [0],
                       label="Posizione aggiornata",
                       color='red', linewidth=1)
plt.xlabel('Tempo (5 s)', fontsize=18, labelpad=5)
plt.ylabel('Posizione (m)', fontsize=18, labelpad=10)
plt.legend(handles=[x_print, y_print, x_pred_print,
                    x_up_print], fontsize=18)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()
# GRAFICO INNOVAZIONE
plt.plot(N_)
plt.xlabel('Tempo (5 s)', fontsize=18, labelpad=5)
plt.ylabel('Innovazione (m)', fontsize=18, labelpad=10)
plt.plot(a.linspace(1, len(N_), len(N_)), len(N_) * [a.mean(N_)])
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()
# GRAFICO KALMAN GAIN
plt.plot(a.transpose(K_) [0])
plt.xlabel('Tempo (5 s)', fontsize=18, labelpad=5)
plt.ylabel('Kalman gain', fontsize=18, labelpad=10)
plt.xticks(fontsize=18)

```

```
plt.yticks(fontsize=18)
plt.show()
```

Moto uniformemente accelerato

```
import numpy as n
import matplotlib.pyplot as plt

# COSTANTI E SUCCESSIONI
# successione dei vettori di stato
x_ = []
# successione delle misure
y_ = []
# successione delle previsioni dei vettori di stato
x_pred_ = []
# successione delle previsioni delle matrici di covarianza
P_pred_ = []
# successione degli aggiornamenti dei vettori di stato
x_up_ = []
# successione degli aggiornamenti delle matrici di covarianza
P_up_ = []
# successione delle innovazioni
N_ = []
# successione del Kalman gain
K_ = []
# intervallo temporale
deltaT = 0.2
# deviazione standard della distribuzione dell'errore sulla misura
sigma_eta = 2
# costante dell'errore sistematico
c = 50
# varianza della distribuzione dell'errore sulla posizione
sigma_wq = c * (deltaT ** 5) / 20
# varianza della distribuzione dell'errore sulla velocità
sigma_wv = c * (deltaT ** 3) / 3
# varianza della distribuzione dell'errore sull'accelerazione
sigma_wa = c * deltaT
# matrice della transizione di stato
F = n.array([[1, deltaT, (deltaT ** 2) / 2],
             [0, 1, deltaT], [0, 0, 1]])
# matrice di errore sulla covarianza
Q = c * n.array([[ (deltaT ** 5) / 20, (deltaT ** 4) / 8,
                   (deltaT ** 3) / 6], [(deltaT ** 4) / 8,
                   (deltaT ** 3) / 3, (deltaT ** 2) / 2],
                   [(deltaT ** 3) / 3, (deltaT ** 2) / 2, deltaT]])
# matrice dell'errore sulla misura
```

```

R = sigma_eta
# vettore di transizione in uscita
H = a.array([1, 0, 0])
# velocità iniziale
v = 3
# accelerazione
l = 2

# STATI INIZIALI
# si pone nulla la matrice di covarianza iniziale
x_pred = a.array([0, v, l]).T
x_.append(x_pred)
y_.append(x_pred[0])
x_pred_.append(x_pred)
P_pred = a.array([[0, 0, 0], [0, 0, 0], [0, 0, 0]])
P_pred_.append(P_pred)
x_up = a.array([0, v, l]).T
x_up_.append(x_up)
P_up = a.array([[0, 0, 0], [0, 0, 0], [0, 0, 0]])
P_up_.append(P_up)
P = P_up

# ITERAZIONI E APPLICAZIONE DEL FILTRO
for i in range(1, 100):
    x = a.array([l * i * i * deltaT * deltaT / 2,
                l * i * deltaT, l]).T
    w_q = a.random.normal(0, a.sqrt(sigma_wq))
    w_v = a.random.normal(0, a.sqrt(sigma_wv))
    w_a = a.random.normal(0, a.sqrt(sigma_wa))
    w = a.array([w_q, w_v, w_a]).T
    x_.append(x + w)
    # previsione
    x_pred = a.dot(F, x_up)
    x_pred_.append(x_pred)
    P_pred = a.dot(F, a.dot(P, F.T)) + Q
    P_pred_.append(P_pred)
    z = a.random.normal(0, sigma_eta)
    y = a.dot(H, x) + z
    y_.append(y)
    # innovazione
    n = y - a.dot(H, x_pred)
    N_.append(n)
    # varianza dell'innovazione
    S = sigma_eta + a.dot(H, a.dot(P_pred, H.T))
    # Kalman gain
    K = a.dot(P_pred, H.T) / S

```

```

K_.append(K)
# aggiornamento
x_up = x_pred + K * n
x_up_.append(x_up)
M = a.array([[1, 0, 0], [0, 1, 0], [0, 0, 1]]) - \
      a.outer(K, H)
P_up = a.dot(M, a.dot(P_pred, M.T)) + \
      a.dot(K, a.dot(R, K.T))
P_up_.append(P_up)
P = P_up
print(P_up)

# GRAFICO POSIZIONI
x_print, = plt.plot(a.transpose(x_)[0], label="Posizione reale",
                  color='green', linewidth=1)
y_print, = plt.plot(y_, label="Posizione misurata",
                  color='black', linewidth=1)
x_pred_print, = plt.plot(a.transpose(x_pred_)[0],
                        label="Posizione prevista",
                        color='blue', linewidth=1)
x_up_print, = plt.plot(a.transpose(x_up_)[0],
                      label="Posizione aggiornata",
                      color='red', linewidth=1)
plt.xlabel('Tempo (5 s)', fontsize=18, labelpad=5)
plt.ylabel('Posizione (m)', fontsize=18, labelpad=10)
plt.legend(handles=[x_print, y_print,
                  x_pred_print, x_up_print], fontsize=18)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()
# GRAFICO INNOVAZIONE
plt.plot(N_)
plt.xlabel('Tempo (5 s)', fontsize=18, labelpad=5)
plt.ylabel('Innovazione (m)', fontsize=18, labelpad=10)
plt.plot(a.linspace(0, len(N_), len(N_)), len(N_) * [a.mean(N_)])
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()
# GRAFICO KALMAN GAIN
plt.plot(a.transpose(K_)[0])
plt.xlabel('Tempo (5 s)', fontsize=18, labelpad=5)
plt.ylabel('Kalman gain', fontsize=18, labelpad=10)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()

```

Modello approssimativo

```
import numpy as a
import matplotlib.pyplot as plt

# COSTANTI E SUCCESSIONI
# successione dei vettori di stato
x_ = []
# successione delle misure
y_ = []
# successione delle previsioni dei vettori di stato
x_pred_ = []
# successione delle previsioni delle matrici di covarianza
P_pred_ = []
# successione degli aggiornamenti dei vettori di stato
x_up_ = []
# successione degli aggiornamenti delle matrici di covarianza
P_up_ = []
# successione delle innovazioni
N_ = []
# successione del Kalman gain
K_ = []
# intervallo temporale
deltaT = 0.2
# deviazione standard della distribuzione dell'errore sulla misura
sigma_eta = 2
# costante dell'errore sistematico
c = 50
# varianza della distribuzione dell'errore sulla posizione
sigma_wq = c * (deltaT ** 5) / 20
# varianza della distribuzione dell'errore sulla velocità
sigma_wv = c * (deltaT ** 3) / 3
# varianza della distribuzione dell'errore sull'accelerazione
sigma_wa = c * deltaT
# matrice della transizione di stato
F = a.array([[1, deltaT, 0], [0, 1, 0], [0, 0, 0]])
# matrice di errore sulla covarianza
Q = c * a.array([[ (deltaT ** 5) / 20, (deltaT ** 4) / 8,
                  (deltaT ** 3) / 6], [(deltaT ** 4) / 8,
                  (deltaT ** 3) / 3, (deltaT ** 2) / 2],
                  [(deltaT ** 3) / 3, (deltaT ** 2) / 2, deltaT]])
# matrice dell'errore sulla misura
R = sigma_eta
# vettore di transizione in uscita
H = a.array([1, 0, 0])
# velocità iniziale
```

```

v = 0
# accelerazione
l = 4

# STATI INIZIALI
# si pone nulla la matrice di covarianza iniziale
x_pred = a.array([0, v, l]).T
x_.append(x_pred)
y_.append(x_pred[0])
x_pred_.append(x_pred)
P_pred = a.array([[0, 0, 0], [0, 0, 0], [0, 0, 0]])
P_pred_.append(P_pred)
x_up = a.array([0, v, l]).T
x_up_.append(x_up)
P_up = a.array([[0, 0, 0], [0, 0, 0], [0, 0, 0]])
P_up_.append(P_up)
P = P_up

# ITERAZIONI E APPLICAZIONE DEL FILTRO
for i in range(1, 100):
    x = a.array([l * i * i * deltaT * deltaT / 2, l * i * deltaT, l])
    w_q = a.random.normal(0, a.sqrt(sigma_wq))
    w_v = a.random.normal(0, a.sqrt(sigma_wv))
    w_a = a.random.normal(0, a.sqrt(sigma_wa))
    w = a.array([w_q, w_v, w_a]).T
    x_.append(x + w)
    # previsione
    x_pred = a.dot(F, x_up)
    x_pred_.append(x_pred)
    P_pred = a.dot(F, a.dot(P, F.T)) + Q
    P_pred_.append(P_pred)
    z = a.random.normal(0, sigma_eta)
    y = a.dot(H, x + w) + z
    y_.append(y)
    # innovazione
    n = y - a.dot(H, x_pred)
    N_.append(n)
    # varianza dell'innovazione
    S = sigma_eta + a.dot(H, a.dot(P_pred, H.T))
    # Kalman gain
    K = a.dot(P_pred, H.T) / S
    K_.append(K)
    # aggiornamento
    x_up = x_pred + K * n
    x_up_.append(x_up)
    M = a.array([[1, 0, 0], [0, 1, 0], [0, 0, 1]]) - \

```

```

        a.outer(K, H)
    P_up = a.dot(M, a.dot(P_pred, M.T)) + \
        a.dot(K, a.dot(R, K.T))
    P_up_.append(P_up)
    P = P_up

# GRAFICO POSIZIONI
x_print, = plt.plot(a.transpose(x_)[0], label="Posizione reale",
                    color='green', linewidth=1)
y_print, = plt.plot(y_, label="Posizione misurata",
                    color='black', linewidth=1)
x_pred_print, = plt.plot(a.transpose(x_pred_)[0],
                         label="Posizione prevista",
                         color='blue', linewidth=1)
x_up_print, = plt.plot(a.transpose(x_up_)[0],
                       label="Posizione aggiornata",
                       color='red', linewidth=1)

plt.xlabel('Tempo (5 s)', fontsize=18)
plt.ylabel('Posizione (m)', fontsize=18)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.legend(handles=[x_print, y_print, x_pred_print,
                    x_up_print], fontsize=18)

plt.show()
# GRAFICO INNOVAZIONE
plt.plot(N_)
plt.xlabel('Tempo (5 s)', fontsize=18)
plt.ylabel('Innovazione (m)', fontsize=18)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.plot(a.linspace(1, len(N_), len(N_)), len(N_) * [a.mean(N_)])
plt.show()
# GRAFICO KALMAN GAIN
plt.plot(a.transpose(K_)[0])
plt.xlabel('Tempo (5 s)', fontsize=18)
plt.ylabel('Kalman gain', fontsize=18)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()

```

Oscillatore armonico

```

import numpy as a
import math
import matplotlib.pyplot as plt

```

```

# COSTANTI E SUCCESSIONI
# velocità angolare
o = 2
# successione dei vettori di stato
x_ = []
# successione delle misure
y_ = []
# successione delle previsioni dei vettori di stato
x_pred_ = []
# successione delle previsioni delle matrici di covarianza
P_pred_ = []
# successione degli aggiornamenti dei vettori di stato
x_up_ = []
# successione degli aggiornamenti delle matrici di covarianza
P_up_ = []
# successione delle innovazioni
N_ = []
# successione del Kalman gain
K_ = []
# intervallo temporale
deltaT = 0.1
# deviazione standard della distribuzione dell'errore sulla misura
sigma_eta = 4
# cotante dell'errore sistematico
c = 30
# varianza della distribuzione dell'errore sulla posizione
sigma_wq = c * (deltaT ** 3) / 3
# varianza della distribuzione dell'errore sulla velocità
sigma_wv = c * (deltaT ** 2) / 2
# varianza della distribuzione dell'errore sull'accelerazione
sigma_wa = c * deltaT
# matrice della transizione di stato
F = a.array([[1, deltaT, (deltaT ** 2) / 2],
             [0, 1, deltaT], [0, 0, 1]])
# matrice di covarianza
P = a.array([[0, -1, 1], [-1, 0, -1], [1, -1, 0]])
# matrice di errore sulla covarianza
Q = c * a.array([[ (deltaT ** 5) / 20, (deltaT ** 4) / 8,
                  (deltaT ** 3) / 6], [(deltaT ** 4) / 8,
                  (deltaT ** 3) / 3, (deltaT ** 2) / 2],
                  [(deltaT ** 3) / 3, (deltaT ** 2) / 2, deltaT]])
# matrice dell'errore sulla misura
R = sigma_eta
# vettore di transizione in uscita
H = a.array([1, 0, 0])
# ampiezza delle oscillazioni

```

```

A_ = 50

# STATI INIZIALI
x_pred = a.array([0, A_ * o, 0]).T
x_.append(x_pred)
y_.append(x_pred[0])
x_pred_.append(x_pred)
P_pred = a.array([[0, 0, 0], [0, 0, 0], [0, 0, 0]])
P_pred_.append(P_pred)
x_up = a.array([0, A_ * o, 0]).T
x_up_.append(x_up)
P_up = a.array([[0, 0, 0], [0, 0, 0], [0, 0, 0]])
P_up_.append(P_up)

# ITERAZIONI E APPLICAZIONE DEL FILTRO
for i in range(1, 100):
    x = a.array([A_ * math.sin(o * i * deltaT),
                A_ * o * math.cos(o * i * deltaT),
                A_ * (o ** 2) * math.sin(o * i * deltaT)]).T
    w_q = a.random.normal(0, a.sqrt(sigma_wq))
    w_v = a.random.normal(0, a.sqrt(sigma_wv))
    w_a = a.random.normal(0, a.sqrt(sigma_wa))
    w = a.array([w_q, w_v, w_a]).T
    x_.append(x + w)
    # previsione
    x_pred = a.dot(F, x_up)
    x_pred_.append(x_pred)
    P_pred = a.dot(F, a.dot(P, F.T)) + Q
    P_pred_.append(P_pred)
    z = a.random.normal(0, sigma_eta)
    y = a.dot(H, x) + z
    y_.append(y)
    # innovazione
    n = y - a.dot(H, x_pred)
    N_.append(n)
    # varianza dell'innovazione
    S = sigma_eta + a.dot(H, a.dot(P_pred, H.T))
    # Kalman gain
    K = a.dot(P_pred, H.T) / S
    K_.append(K)
    # aggiornamento
    x_up = x_pred + K * n
    x_up_.append(x_up)
    M = a.array([[1, 0, 0], [0, 1, 0], [0, 0, 1]]) - \
        a.outer(K, H)
    P_up = a.dot(M, a.dot(P_pred, M.T)) + \

```

```

        a.dot(K, a.dot(R, K.T))
    P_up_.append(P_up)
    P = P_up

# GRAFICO POSIZIONI
x_print, = plt.plot(a.transpose(x_)[0], label="Posizione reale",
                    color='green', linewidth=1)
y_print, = plt.plot(y_, label="Posizione misurata",
                    color='black', linewidth=1)
x_pred_print, = plt.plot(a.transpose(x_pred_)[0],
                         label="Posizione prevista",
                         color='blue', linewidth=1)
x_up_print, = plt.plot(a.transpose(x_up_)[0],
                       label="Posizione aggiornata",
                       color='red', linewidth=1)
plt.xlabel('Tempo (10 s)', fontsize=18, labelpad=5)
plt.ylabel('Posizione (m)', fontsize=18, labelpad=10)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.legend(handles=[x_print, y_print, x_pred_print,
                    x_up_print], fontsize=18)

plt.show()
# GRAFICO INNOVAZIONE
plt.plot(N_)
plt.xlabel('Tempo (10 s)', fontsize=18, labelpad=5)
plt.ylabel('Innovazione (m)', fontsize=18, labelpad=10)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.plot(a.linspace(1, len(N_), len(N_)), len(N_) * [a.mean(N_)])
plt.show()
# GRAFICO KALMAN GAIN
plt.plot(a.transpose(K_)[0])
plt.xlabel('Tempo (10 s)', fontsize=18, labelpad=5)
plt.ylabel('Kalman gain', fontsize=18, labelpad=10)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()

```

Serie storiche di titoli di borsa

Il codice utilizzato per le simulazioni effettuate sulle tre diverse serie S^1 , S^2 e S^3 è assolutamente lo stesso, quello che cambia nei due casi sono semplicemente i dati letti da file in input e il valore della costante di drift.

```
import numpy as a
```

```

import matplotlib.pyplot as plt
import scipy.stats as st

# COSTANTI E SUCCESSIONI
# successione dei vettori di stato
x_ = []
# successione delle previsioni dei vettori di stato
x_pred_ = []
# successione delle previsioni delle matrici di covarianza
P_pred_ = []
# successione degli aggiornamenti dei vettori di stato
x_up_ = []
# successione degli aggiornamenti delle matrici di covarianza
P_up_ = []
# successione delle innovazioni
N_ = []
# successione del Kalman gain
K_ = []
# intervallo temporale
deltaT = 1
# deviazione standard del rumore browniano sui ritorni
sigma_wq = 1.0
# matrice della transizione di stato
F = a.array([[1.0, deltaT], [0.0, 1.0]])
# vettore di transizione in uscita
H = a.array([1.0, 0.0])
# costante di drift
m = -0.0021
# matrice di errore sulla covarianza
Q = 3 * a.array([(deltaT ** 3) / 3,
                 (deltaT ** 2) / 2],
                [(deltaT ** 2) / 2, deltaT])

# STATI INIZIALI
# si pone nulla la matrice di covarianza iniziale
x_pred = a.array([0.0, m]).T
x_.append(x_pred)
x_pred_.append(x_pred)
P_pred = a.array([[0.0, 0.0], [0.0, 0.0]])
P_pred_.append(P_pred)
x_up = a.array([0.0, m]).T
x_up_.append(x_up)
P_up = a.array([[0.0, 0.0], [0.0, 0.0]])
P_up_.append(P_up)
P = P_up

```

```

# COSTRUZIONE DELL'ISTOGRAMMA
s_ = []
s_ = a.genfromtxt("stockdata.dat", float)
plt.hist(s_, density=True, bins=20, label="Valori reali")
mn, mx = plt.xlim()
plt.xlim(mn, mx)
kde_xs = a.linspace(mn, mx, 300)
kde = st.gaussian_kde(s_)
plt.plot(kde_xs, kde.pdf(kde_xs), label="Probability Density Function")
plt.legend(loc="upper left", fontsize=18)
plt.ylabel("Probabilità", labelpad=10, fontsize=18)
plt.xlabel("Rendimenti (%)", labelpad=5, fontsize=18)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()

# ITERAZIONI E APPLICAZIONE DEL FILTRO
for i in range(1, 157):
    x = s_[i]
    y = a.array([x, m]).T
    x_.append(y)
    # previsione
    x_pred = a.dot(F, x_up)
    x_pred_.append(x_pred)
    P_pred = a.dot(F, a.dot(P, F.T)) + Q
    P_pred_.append(P_pred)
    # innovazione
    n = a.dot(H, x_[i] - x_pred)
    N_.append(n)
    # covarianza sull'innovazione
    S = a.inner(H, a.dot(P_pred, H.T))
    # Kalman gain
    K = a.dot(P_pred, H.T) / S
    K_.append(K)
    # previsione
    x_up = x_pred + K * n
    x_up_.append(x_up)
    M = a.array([[1, 0], [0, 1]]) - a.outer(K, H)
    P_up = a.dot(M, a.dot(P_pred, M.T))
    P_up_.append(P_up)
    P = P_up
for i in range(0, 157):
    print((x_pred_[i])[0])

# GRAFICO POSIZIONI
x_print, = plt.plot(a.transpose(x_)[0],

```

```

        label="Rendimento reale",
        color='green', linewidth=3)
x_pred_print, = plt.plot(a.transpose(x_pred_)[0],
        label="Rendimento previsto",
        color='blue', linewidth=1)
x_up_print, = plt.plot(a.transpose(x_up_)[0],
        label="Rendimento aggiornato",
        color='red', linewidth=1)
plt.xlabel('Tempo (settimane)', fontsize=18)
plt.ylabel('Rendimento (%)', fontsize=18)
plt.legend(handles=[x_print, x_pred_print, x_up_print], fontsize=18)
plt.xticks(fontsize=18)
plt.yticks(fontsize=18)
plt.show()
# GRAFICO KALMAN GAIN
plt.plot(a.transpose(K_)[0])
plt.show()

```

Bibliografia

- Brealey R. A.; Myers, S. C. *Principi di Finanza Aziendale*. McGraw-Hill Education, 2015.
- Durbin, J.; Koopman S. J. *Time Series Analysis by Space State Methods*. Oxford University Press, 2012.
- Fama, E. F. «Efficient Capital Markets: A Review of Theory and Empirical Work». In: *The Journal of Finance* 25 (1969), pp. 383–417.
- «Random Walks in Stock Market Prices». In: *Financial Analysts Journal* 5 (1965), pp. 55–59.
- Fornasini, P. *The Uncertainty in Physical Measurement*. Springer, Science & Business Media, 2008.
- Kanieski, J. M. «Kalman Filter-Based Control System for Power Quality Conditioning Devices». In: *IEEE Transactions on Industrial Electronics* 60 (2013), pp. 5214–5227.
- Kendall M. G.; Bradford Hill, A. «The Analysis of Economic Time Series, Part I: Prices». In: *Journal of the Royal Statistical Society. Series A* 116 (1953), pp. 11–34.
- Malkiel, B. G. *A Ranom Walk Down Wall Street*. W. W. Norton & Company, 2015.
- Manoliu M; Tompaidis, S. «Energy futures prices: term structure models with Kalman filter estimation». In: *Applied Mathematical Finance* 9 (2010), pp. 21–43.
- Mohamed, A. H.; Schwarz K. P. «Adaptive Kalman Filtering for INS/GPS». In: *Journal of Geodesy* 73 (1999), pp. 193–203.
- Pascucci A.; Wolfgang, J. R. *Finanza matematica: teoria e problemi per modelli multi-periodali*. Springer, Science & Business Media, 2009.
- Reid I.; Term, H. *Estimation and Kalman Filtering*. Oxford University, Department of Engineering Science. 2001. URL: <https://www.robots.ox.ac.uk/~ian/Teaching/Estimation/LectureNotes2.pdf>.
- Requião da Cunha, C. *Introduction to Econophysics*. CRC Press, Taylor & Francis Group, 2022.

Ringraziamenti

Desidero ringraziare innanzitutto la mia famiglia, che mi è stata vicina in tutti questi anni, nei momenti migliori come in quelli peggiori; rivolgo quindi un abbraccio a mio padre, mia madre, i miei fratelli Diego, Dafne, Alessandro ed Edoardo e tutti i miei nonni. Sapere di avervi sempre accanto in una realtà sempre incerta e piena di sfide è quanto di più importante una persona possa desiderare.

Ringrazio quindi la Prof.ssa Ercolessi e il Dott. Sanavio per aver curato questo lavoro con attenzione e diligenza, svolgendo un ruolo di fondamentale sostegno nella stesura dello stesso. Desidero quindi ringraziare il Prof. Boscherini per aver egregiamente svolto nei miei riguardi la funzione di Tutor del Collegio Superiore in questi tre anni di impegno e studio, ma anche di grandi e impagabili soddisfazioni.

Vorrei ringraziare poi tutti i collegiali, per aver costruito una rete e un ambiente accogliente, familiare, motivante e stimolante, senza il quale non sarei sicuramente riuscito a terminare questo percorso. Un ringraziamento andrebbe poi rivolto a tutti gli amici che mi hanno regalato momenti fantastici. Per menzionarvi tutti occorrerebbe un'altra tesi, ma sappiate che non c'è nessuno di voi a cui non ho pensato nel momento di scrivere questi ringraziamenti. Vista la circostanza, tuttavia, una menzione particolare va a Lorenzo Massarenti, che, da quando l'ho conosciuto una volta entrato al Collegio Superiore, mi ha aiutato a conoscere me stesso, a mettermi alla prova, a dare del mio meglio nei periodi trascorsi a Bologna più di chiunque altro.

Vorrei infine ringraziare la A.S. Roma. Ho iniziato la tesi scrivendo: "La realtà è costituita nel complesso da moltissime grandezze che evolvono nel tempo". Ma ci sono anche alcune cose che nel tempo non cambiano mai e una di queste è l'amore per qualcosa che sin da piccoli ci ha permesso di accenderci gli occhi sempre con lo stesso entusiasmo. Il calcio è una magia, capace, nella sua semplicità, di incantare, di emozionare, di unire le persone, di far piangere, come quella sera del 25 maggio in cui a Tirana abbiamo alzato al cielo la coppa tutti insieme. Voglio ringraziare la Roma perché nel bene e nel male riesce sempre ad accendere in me una scintilla di amore, quel senso di bellezza di un quadro o di una poesia, quel senso di immensità. Spero un giorno di poter venire a lavorare da voi come Data Analyst Manager o qualcosa di simile, chissà.

