

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

SCUOLA DI SCIENZE
Corso di Laurea Magistrale in Matematica

Applicazione delle reti neurali residuali
ai mercati energetici
e loro interpretazione come
problema di controllo ottimo stocastico

Tesi di Laurea in Analisi Stocastica

Relatore:
Prof.
ANDREA COSSO
Correlatore:
Dott.ssa
ELISA RASPANTI

Presentata da:
TERESA ANGELI

II Sessione
Anno Accademico 2020-21

Indice

Introduzione	III
1 Deep Learning: reti neurali residuali	1
1.1 Reti neurali	1
1.1.1 Discesa del gradiente e algoritmo di <i>backpropagation</i>	4
1.2 Reti neurali residuali	7
1.2.1 Formulazione matematica	8
2 Applicazione delle reti neurali residuali ai mercati energetici	10
2.1 Presentazione del problema	10
2.1.1 Raccolta dei dati	11
2.1.2 Allenamento della rete e scelta dei parametri	12
2.1.3 Simulazione di volumi	15
2.2 Punti di forza e limiti	17
2.2.1 Confronto con altri strumenti predittivi	17
3 Problema di controllo	20
3.1 Interpretazione di una rete residuale come problema di controllo	20
3.2 Notazioni ed ipotesi sui coefficienti	21
3.2.1 Notazioni	21
3.2.2 Ipotesi sui coefficienti	23
3.3 Equazione di stato	24
3.4 Funzionale costo e funzione valore	28
3.5 Principio della programmazione dinamica	30
3.5.1 Continuità della funzione valore	32
4 Equazione di Hamilton-Jacobi-Bellman e soluzioni di viscosità	34
4.1 Derivata rispetto alla misura di probabilità	34
4.2 Formula di Itô	37
4.3 Equazione di Hamilton-Jacobi-Bellman	37

4.4 Soluzioni di viscosità	38
4.4.1 Esistenza e unicità	39
Bibliografia	42

Introduzione

Il Deep Learning è uno strumento sempre più rilevante in diversi campi di applicazione del Machine Learning. Per tale ragione, c'è una crescente richiesta di strumenti matematici in grado di analizzare, in modo rigoroso, vari aspetti delle reti neurali profonde.

In quest'ottica, il presente lavoro di tesi ha l'obiettivo di stabilire un legame tra reti neurali residuali, descritte nel Capitolo 1, e teoria del controllo ottimo stocastico, a partire dalla formulazione proposta da E, Han e Li in [4]. Tale legame è formulato e indagato dal punto di vista teorico, attraverso lo sviluppo e lo studio del problema di controllo ottimo stocastico coinvolto. L'idea è quella di interpretare una rete neurale residuale come un problema di controllo ottimo a tempo continuo, supponendo di far tendere all'infinito il numero di strati. In questo modo, l'equazione che regola il passaggio da uno strato all'altro della rete diventa un'equazione differenziale controllata con condizione iniziale aleatoria. Si ha quindi un'identificazione tra i pesi della rete, ovvero i parametri che devono essere calibrati per garantire una buona capacità predittiva, e il processo di controllo soluzione del problema di controllo ottimo.

Dal punto di vista applicativo, si è scelto di utilizzare una rete neurale residuale per risolvere un problema reale legato all'ambito dei mercati energetici, illustrato nel dettaglio nel Capitolo 2. Lo scopo è quello di indagare la relazione tra prezzi e volumi dei sottostanti di interesse e per farlo è stata costruita e allenata una rete residuale attraverso il pacchetto Keras di Python. Lo studio di questo problema è consistito in un'analisi approfondita del contesto e dei dati storici a disposizione e in varie fasi di training e di test della rete. Sono poi stati commentati e interpretati i risultati ottenuti e valutati i punti di forza e i limiti dell'approccio scelto, confrontando lo strumento costruito con altri strumenti predittivi.

Nei Capitoli 3 e 4, la tesi generalizza i risultati riportati in [4] imponendo condizioni più deboli sui coefficienti, che permettono di trattare casi maggiormente significativi per le applicazioni. Si propone poi di studiare il corrispondente problema di controllo ottimo, mostrando che la rete residuale

è associata ad un'opportuna equazione alle derivate parziali di Hamilton-Jacobi-Bellman sullo spazio di Wasserstein delle misure di probabilità con momento secondo finito. Viene inoltre introdotta la nozione di soluzione di viscosità per tale equazione, che coinvolge operatori differenziali recentemente proposti da Pierre-Louis Lions per funzioni definite sullo spazio di Wasserstein.

Il contributo di questo lavoro di tesi è quindi dato da una formulazione rigorosa del legame tra rete residuale e problema di controllo e ad una generalizzazione dei risultati teorici necessari alla caratterizzazione dei parametri ottimali. L'ipotesi che viene infine suggerita è che un ulteriore sviluppo teorico del problema di controllo ottimo possa dare un supporto utile alle applicazioni, attraverso diversi algoritmi di risoluzione o fornendo opportune caratterizzazioni dei parametri ottimali.

Capitolo 1

Deep Learning: reti neurali residuali

Nel seguente capitolo vengono esposti i principali concetti relativi alle reti neurali, in particolare quelle di tipo residuale. I contenuti riportati sono estratti da [5], [6], [7] e [9].

1.1 Reti neurali

Il Deep Learning è un'area del Machine Learning basata sull'utilizzo di reti neurali artificiali, strutture stratificate in grado di apprendere informazioni sulla base di un insieme di dati presi in input. L'aspetto caratteristico di questo tipo di apprendimento è proprio la "profondità", da cui l'aggettivo *deep*, in quanto la presenza dei diversi strati permette l'esecuzione di operazioni sequenziali sui dati in input, che consistono in trasformazioni lineari seguite da funzioni non lineari. Questo meccanismo consente un apprendimento sempre più accurato da parte della rete.

La struttura delle reti neurali artificiali si ispira alle reti neurali biologiche presenti nel nostro cervello. I neuroni, infatti, unità strutturali e funzionali del sistema nervoso, sono connessi tra loro ed è proprio la struttura stratificata che viene a formarsi a consentire un'adeguata risposta agli stimoli elettrici e la completa elaborazione delle informazioni. La connessione tra neuroni è garantita dalle sinapsi, le quali veicolano il segnale da una cellula all'altra. Ogni neurone riceve diversi segnali in ingresso e li elabora facendo una somma pesata. Se il risultato supera una certa soglia, il neurone si attiva; altrimenti rimane in uno stato di riposo.

Le reti neurali artificiali sono costruite per analogia, con lo scopo di imparare ad estrarre informazioni da un insieme di dati. La loro struttura è

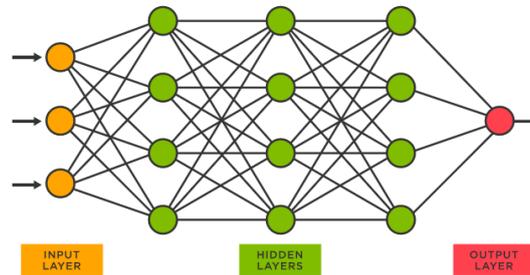


Figura 1.1: Struttura di una rete neurale artificiale *feed-forward*. Gli strati centrali prendono il nome di *strati nascosti* (in inglese *hidden layers*).

costituita da strati formati da singole unità, ciascuna connessa con lo strato precedente e il successivo, come si osserva in Figura 1.1.

Le reti più comuni sono le *feed-forward neural networks*, che presentano connessioni unidirezionali dallo strato di input allo strato di output. Ogni strato prende in input l'output dello strato precedente, pesandolo con opportuni coefficienti associati agli archi della rete. In particolare per andare da uno strato all'altro, le unità coinvolte operano una somma pesata degli input ricevuti dallo strato precedente e applicano al risultato una funzione non lineare f , detta **funzione di attivazione**. Il k -esimo neurone, quindi, riceve gli input x_1, \dots, x_N dai neuroni dello strato precedente, opera una combinazione lineare pesata con pesi θ_{kj} , somma eventualmente un bias b_k e applica la funzione di attivazione f , ottenendo il seguente output

$$f\left(\sum_{j=1}^N \theta_{kj} x_j + b_k\right).$$

Il bias è un parametro che caratterizza il singolo neurone, indipendente dai pesi e dagli input, e va inteso come una soglia che determina se e in quale misura il neurone debba attivarsi. Una funzione di attivazione molto comune, utilizzata anche nel seguito, è la ReLU (*Rectified Linear Unit*), ovvero la parte positiva, data da $f(z) = \max\{f(z), 0\}$, la quale garantisce una buona velocità di apprendimento nei casi con molti strati. Altre funzioni spesso scelte come funzioni di attivazione sono la funzione di Heaviside, la tangente iperbolica e la funzione sigmoidea $f(z) = \frac{1}{1+e^{-z}}$. L'esempio più semplice di rete *feed-forward* è il perceptrone, in Figura 1.2, che presenta un unico neurone in grado di elaborare più input restituendo un solo output.

Lo scopo della rete è quello di modificare i parametri θ e b in modo da restituire un output il più possibile coerente con l'input ricevuto. In partico-

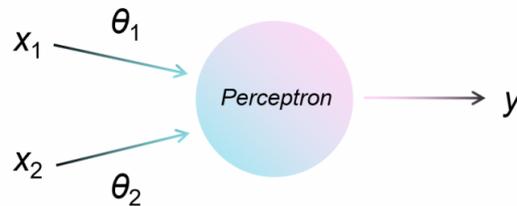


Figura 1.2: Il perceptrone, esempio più semplice di *feed-forward neural network*.

lare, la rete neurale è sottoposta ad una prima fase di allenamento, durante la quale le vengono fornite varie coppie input-output, con l'obiettivo di impararne la relazione. Il set di dati utilizzato, detto *training set*, è costituito da una serie di esempi che consentono alla rete di approssimare la relazione tra input e output e imparare a riproporla nel caso in cui le venga fornito solo un input. Dopo la fase di allenamento, è necessario avere a disposizione un secondo set di dati, detto *validation set*, che permette di confrontare vari set di parametri ottimali ottenuti dall'allenamento, valutando l'errore commesso dalla rete, e scegliere quelli che garantiscono la migliore performance. Viene quindi fornito alla rete solo l'input di questo nuovo insieme di dati ed essa ha il compito di prevedere l'output sulla base di ciò che ha appreso, per poi confrontarlo con l'output effettivo. Infine, una volta fissati i parametri e i pesi, si può testare la rete su un *test set* contenente dati diversi rispetto ai dataset precedenti, con lo scopo di valutare la capacità predittiva della rete. Per approfondire si rimanda a [13]. Questo tipo di approccio prende il nome di *apprendimento supervisionato* e verrà utilizzato nel capitolo successivo per l'addestramento della rete neurale di interesse.

Altri tipi di apprendimento, non approfonditi in questo elaborato, sono l'*apprendimento non supervisionato*, nel quale vengono forniti gli input senza i rispettivi output con lo scopo di far intuire alla rete la struttura dei dati, e l'*apprendimento per rinforzo*, che si basa sull'interazione dinamica con un ambiente esterno, senza l'utilizzo di esempi forniti manualmente.

Gli output prodotti dalla rete possono essere di tipo qualitativo o quantitativo. Si parla rispettivamente di *classificazione* o *regressione*. Nel primo caso la rete neurale ha l'obiettivo di imparare a classificare gli input, ponendo ciascun dato in una opportuna categoria. Un tipico esempio è quello del riconoscimento di immagini, in cui vengono fornite in fase di training coppie formate da un'immagine e la rispettiva etichetta e in un secondo momento viene chiesto alla rete di riconoscere immagini non note appartenenti alle

classi apprese. Nel caso della regressione, invece, la relazione tra input e output è una vera e propria funzione che la rete deve approssimare nel modo più accurato possibile per poter associare a input non noti gli output appropriati. È la stessa idea che viene utilizzata quando si approssima una nube di punti attraverso una regressione lineare.

Come accennato precedentemente, la fase di allenamento serve alla rete per tarare opportunamente i parametri. Questo processo avviene attraverso la minimizzazione della **funzione costo**. Tale funzione prende in considerazione l'output prodotto dalla rete in fase di validazione e l'output effettivo e valuta l'errore. Spesso si utilizza la funzione MSE (*Mean Squared Error*), ovvero la media del quadrato della differenza tra output previsto e output effettivo. L'obiettivo è chiaramente quello di minimizzare l'errore per rendere la previsione della rete più accurata possibile. Questo è possibile modificando i coefficienti θ che, come abbiamo visto, fungono da pesi, e i parametri che rappresentano i bias b (nel seguito supponiamo siano nulli per semplicità). Inizialmente tali parametri vengono settati in modo random o secondo alcune convenzioni e nel corso dell'addestramento vengono ottimizzati.

1.1.1 Discesa del gradiente e algoritmo di *backpropagation*

La modifica dei pesi θ avviene attraverso una tecnica di ottimizzazione, detta discesa del gradiente, combinata con l'algoritmo di *backpropagation*. L'idea è quella di procedere iterativamente nella direzione opposta a quella del gradiente della funzione costo, in modo da minimizzare l'errore. Intuitivamente, l'equazione che regola tale metodo è la seguente:

$$\theta \rightarrow \theta - \eta \nabla \text{Cost}(\theta),$$

dove $\theta \in \mathbb{R}^n$ è il vettore dei parametri da cui parte l'iterazione, $\eta \in \mathbb{R}$ è un parametro detto *learning rate* che regola la velocità di discesa e Cost è la funzione costo, di cui in questo caso si enfatizza la dipendenza dai pesi θ . $\nabla \text{Cost}(\theta)$ è dato dalla media dei gradienti della funzione costo calcolati su tutti i punti del dataset fornito alla rete. Quindi al valore di ogni coefficiente viene iterativamente sottratto il valore della derivata parziale della funzione costo rispetto al coefficiente moltiplicata per il *learning rate*.

Quando la rete presenta molti parametri e la dimensione del training set è elevata, il calcolo del gradiente ad ogni iterazione per ogni punto può risultare molto costoso. Per questo si può utilizzare una variante del metodo di discesa del gradiente, detta *discesa stocastica del gradiente*, in cui invece di

utilizzare la media dei gradienti calcolati su tutti i punti, si sceglie in modo random il gradiente di un singolo punto o la media su un sottoinsieme di punti.

Il calcolo dei gradienti avviene attraverso l'algoritmo di *backpropagation*, ampiamente utilizzato per l'apprendimento supervisionato. Tale algoritmo può essere pensato come un'applicazione pratica della regola della catena per il calcolo della derivata. Esso, infatti, per ricavare il gradiente della funzione costo rispetto agli input, lavora all'indietro a partire dall'ultimo strato della rete in cui viene prodotto l'output e calcolato l'errore facendo uso del valore effettivo di output. Il funzionamento di questo processo è riassunto in Figura 1.3. Dopo il calcolo del gradiente, inizia la fase di aggiornamento dei pesi attraverso il metodo di discesa del gradiente. Vengono quindi modificati i pesi di ciascun neurone, con l'intento di minimizzare il valore della funzione costo. Il procedimento viene ripetuto varie volte finché l'errore si avvicina a zero o comunque fino a raggiungere una soglia accettabile. Tali iterazioni prendono anche il nome di *epoche*. Il numero di epoche è un parametro fissato della rete e può essere pensato come il numero di volte in cui la rete prende in esame lo stesso dataset e modifica iterativamente i pesi.

La tecnica di ottimizzazione appena esposta presenta alcuni limiti che possono compromettere la buona riuscita della fase di training. Può infatti succedere che la ricerca del minimo rimanga bloccata in corrispondenza di un minimo locale e che quindi non si riesca a minimizzare la funzione costo oltre una certa soglia. Un altro problema che si riscontra è quello della scomparsa del gradiente (in inglese *vanishing gradient problem*), causato dalla ripetuta moltiplicazione per valori molto piccoli del gradiente di alcune funzioni di attivazione. Infatti, per la regola della catena, le derivate parziali della funzione costo rispetto ai pesi degli strati in prossimità degli input contengono il prodotto tra le derivate parziali della funzione di attivazione. Se tale prodotto è nullo o molto piccolo, i pesi dei primi strati rimangono costanti o si aggiornano sempre più lentamente, al contrario di quanto accade per gli strati vicini all'output. Questo problema è accentuato dall'aumentare della profondità della rete e può essere affrontato attraverso una scelta opportuna della funzione di attivazione, ad esempio la ReLU, e dell'inizializzazione dei pesi. Tra l'altro, tali scelte contribuiscono ad evitare il problema speculare al precedente di esplosione del gradiente.

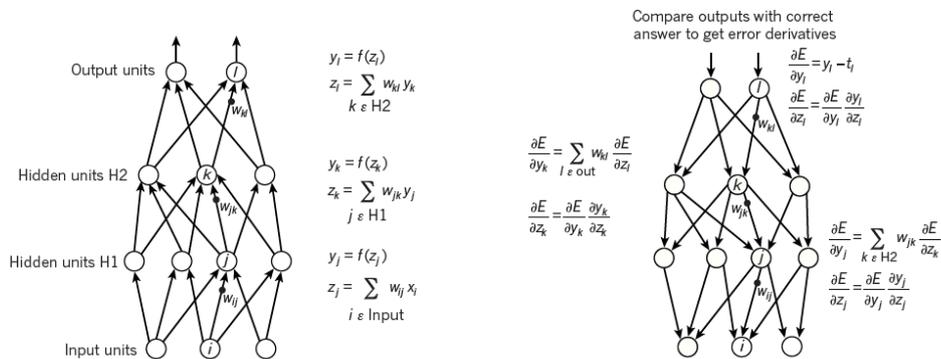


Figura 1.3: A sinistra le equazioni utilizzate per la propagazione "in avanti", ovvero la fase in cui a partire dai valori in input la rete produce gli output, calcolando somme pesate e applicando la funzione di attivazione. A destra le equazioni utilizzate per la propagazione dell'errore "all'indietro". In ogni strato viene calcolata la derivata dell'errore rispetto all'output di ciascun neurone e viene poi convertita in derivata rispetto all'input attraverso la moltiplicazione per il gradiente della funzione di attivazione. L'immagine e la descrizione sono tratte da [9].

1.2 Reti neurali residuali

Le reti neurali residuali o ResNet sono un particolare tipo di CNN (*Convolutional Neural Network*), basato sul concetto del *residual learning*. Una rete neurale convoluzionale (CNN) è una rete neurale artificiale ispirata alla corteccia visiva animale, in cui i neuroni, organizzati in strutture tridimensionali, svolgono specifiche funzioni che combinate insieme consentono l'elaborazione e il riconoscimento di immagini. Tali reti sono costituite da strati di diverso tipo, tra cui quelli convoluzionali, in cui si operano convoluzioni con particolari funzioni, dette filtri, in grado di delineare contorni, profondità, colori. Una struttura comune è data dalla sequenza di blocchi convoluzione-ReLU-convoluzione. Le CNN hanno applicazioni nel riconoscimento di immagini e nell'elaborazione del linguaggio naturale.

Per motivare l'introduzione delle reti neurali residuali, possiamo partire dalla constatazione che aggiungere strati ad una rete neurale non necessariamente migliora la sua performance, come si potrebbe pensare. Consideriamo una rete neurale e una sua versione più profonda, ottenuta aggiungendo alcuni strati. Possiamo costruire una soluzione banale della rete più profonda unendo alla prima rete neurale già allenata alcuni strati che riproducono la funzione identità. L'esistenza di tale soluzione costruttiva suggerisce che una rete più profonda non dovrebbe produrre un errore più ampio di una sua versione meno profonda. Tuttavia gli esperimenti condotti e spiegati in [5] mostrano che non è detto sia possibile trovare una soluzione migliore di quella costruttiva. Tale anomalia deriva dal fatto che l'aumento della profondità della rete enfatizza il problema della scomparsa del gradiente precedentemente descritto.

La struttura di una rete neurale residuale, come abbiamo anticipato, si basa sul concetto del *residual learning* attraverso l'introduzione di un particolare blocco, detto **residual block**, visibile in Figura 1.4. L'idea è quella di far apprendere alla rete, non una trasformazione diretta dell'input in output, ma un certo termine $F(x)$ da sommare all'input per ottenere l'output. Il residual block, infatti, sottopone l'input x alla sequenza convoluzione-ReLU-convoluzione e somma al risultato ottenuto lo stesso x . L'input x viene tenuto in memoria attraverso un meccanismo di *skip connection*, ovvero un collegamento in grado di simulare la funzione identità e di saltare gli strati intermedi trasportando l'input al livello desiderato. In uscita dal blocco si avrà allora $H(x) = F(x) + x$ e ad ogni strato è chiesto di imparare il residuo $F(x)$. L'ipotesi è che sia più immediato imparare il residuo piuttosto che la relazione originale $H(x)$. Per convincerci, possiamo pensare al caso estremo in cui la mappa da apprendere è l'identità: in questo caso sarebbe più sempli-

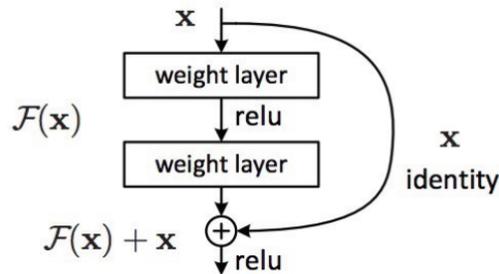


Figura 1.4: Struttura di un residual block. L'addizione avviene attraverso un meccanismo di *skip connection*, rappresentato dalla freccia a destra.

ce annullare il residuo che imparare una funzione identità con una sequenza di strati non lineari. Quindi attraverso la concatenazione di diversi residual blocks, la rete impara il termine residuale $F(x)$ da sommare all'input per avere l'output effettivo. Nel caso in cui input e output abbiano dimensioni differenti, possono essere utilizzate tecniche di zero-padding o di proiezioni, attraverso convoluzioni 1×1 , per creare dimensioni corrispondenti.

L'utilizzo di questo particolare blocco contrasta il problema della scomparsa del gradiente, in quanto, grazie al meccanismo di addizione mediante *skip connection*, durante la *backpropagation* viene creato un percorso alternativo per il gradiente, saltando gli strati intermedi. Questo permette di costruire reti con una profondità più elevata, evitando di comprometterne l'addestramento. Per approfondire questo aspetto si rimanda a [6].

1.2.1 Formulazione matematica

Seguendo l'approccio proposto in [4], possiamo modellizzare una rete neurale residuale come segue. Consideriamo una rete con T strati e sia $x_0 \in \mathbb{R}^d$ il vettore degli input che vengono forniti alla rete. Il passaggio da uno strato al successivo può essere espresso con la seguente equazione alle differenze:

$$x_{t+1} = x_t + f(x_t, \theta_t), \quad t = 0, \dots, T-1, \quad (1.1)$$

dove $f : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^d$ è la funzione di attivazione e θ_t è il vettore dei parametri della rete nello strato t . In questo caso il bias b viene trascurato. L'insieme $\Theta \subseteq \mathbb{R}^m$ contiene i possibili valori del parametro θ e tipicamente nell'ambito del Deep Learning è l'intero spazio \mathbb{R}^m .

L'obiettivo è quello di determinare i parametri $\theta_0, \dots, \theta_{T-1}$ in modo da minimizzare l'errore commesso dalla rete. La funzione costo che valuta tale errore può essere pensata come la somma di due contributi:

- funzione *terminal loss* $\Phi : \mathbb{R}^d \times \mathbb{R}^l \rightarrow \mathbb{R}$ che confronta effettivamente l'output prodotto dalla rete $x_T \in \mathbb{R}^d$ e l'output effettivo $y_0 \in \mathbb{R}^l$, calcolandone la distanza;
- regolarizzatore $L : \Theta \rightarrow \mathbb{R}$, spesso proporzionale alla norma euclidea di θ , che ha lo scopo di evitare situazioni di *overfitting*¹, impedendo al modello di rendere arbitrariamente grandi i suoi parametri.

Dunque il problema può essere esplicitato come segue:

$$\inf_{\theta} \left\{ \Phi(x_T, y_0) + \sum_{t=0}^{T-1} L(\theta_t) \right\} \quad (1.2)$$

con il vincolo (1.1). Si rimanda al Capitolo 3 per maggiori dettagli su questa interpretazione matematica di una rete neurale residuale.

¹L'overfitting si verifica quando la rete, invece di imparare a generalizzare, si limita ad imparare a memoria il training set, diminuendo così il suo livello di accuratezza su dati non noti. Può accadere quando il numero dei parametri è molto maggiore rispetto alle dimensioni del dataset.

Capitolo 2

Applicazione delle reti neurali residuali ai mercati energetici

Nel seguente capitolo viene presentato un problema reale riguardante il mercato dell'energia, che è stato affrontato mediante l'applicazione di una rete neurale residuale. Si ripercorre il lavoro preliminare, si espone il tentativo di risoluzione del problema e si trattano vantaggi e limiti dell'approccio utilizzato.

2.1 Presentazione del problema

Abbiamo studiato un problema reale nell'ambito del mercato dell'energia a cui applicare una rete di tipo residuale. In particolare abbiamo provato ad indagare, attraverso l'utilizzo della rete residuale, la relazione tra prezzi e volumi di sei sottostanti, con l'obiettivo di generare possibili scenari futuri basandosi sui dati storici a disposizione.

Descriviamo il contesto di nostro interesse più nel dettaglio. I prezzi con cui abbiamo lavorato sono prezzi *forward*, ovvero prezzi che vengono stabiliti nel presente e sono riferiti a scadenze future. Dunque conoscendo oggi i prezzi riferiti a vari momenti futuri, si può scegliere quando vendere o acquistare un certo bene che sarà poi ceduto o ottenuto a scadenza. Le scadenze di interesse, nel seguito chiamate anche *delivery*, sono i mesi di un determinato anno (quello in corso o quello successivo) per un insieme di sei sottostanti. I sottostanti considerati nel nostro caso sono:

- PUNBASE: prezzo base dell'elettricità italiana (PUN = Prezzo Unico Nazionale), è dato dalla media dei prezzi di tutte le ore;

- PUNPEAK: prezzo "picco" dell'elettricità italiana, è dato dalla media dei prezzi delle ore di punta (dalle 8 alle 20 dal lunedì al venerdì);
- PSV: prezzo del gas italiano (PSV = Punto di Scambio Virtuale);
- TTF: prezzo del gas olandese, il quale viene scelto come riferimento in quanto più scambiato a livello europeo (TTF = Title Transfer Facility);
- EUA: si riferisce all'autorizzazione necessaria per emettere l'equivalente di una tonnellata di CO_2 durante uno specifico periodo;
- PFOR: indice sulla base del quale sono aggiornati i prezzi del gas, è determinato trimestralmente dall'ARERA (Autorità di Regolazione per Energia Reti e Ambiente) e dipende dalle quotazioni di TTF.

Fissato il mese corrente, quindi, si conoscono i prezzi forward riguardanti i mesi successivi dello stesso anno e tutti i mesi dell'anno successivo, per ogni sottostante. Nei dati a nostra disposizione tali prezzi si aggiornano con granularità giornaliera (nella realtà variano anche all'interno del singolo giorno).

A questi prezzi corrispondono altrettanti volumi, per i sei sottostanti. I volumi hanno segno, cioè sono quantità da vendere o da acquistare, e si aggiornano con granularità mensile. Con lo scorrere dei mesi, per uno stesso mese delivery, il volume da vendere o da acquistare per un certo sottostante potrebbe subire qualche variazione, dovuta alle variazioni di mercato e ad altri fattori quali il periodo dell'anno, la temperatura, le precipitazioni o le necessità dei clienti.

Il problema che abbiamo affrontato è il seguente. Esistono diversi strumenti per simulare l'andamento dei prezzi forward di tutti i sottostanti in maniera correlata e queste simulazioni consentono di generare degli scenari futuri. Se, date le simulazioni di prezzo, si potessero avere anche quelle di volume, si potrebbero generare scenari futuri che tengono conto anche del fatto che i volumi variano se varia il mercato. L'idea è che la rete residuale scelta sia in grado di approssimare la relazione tra prezzi e volumi, così da poter diventare uno strumento di simulazione. Questo approccio parte dal presupposto che ci sia effettivamente una relazione tra queste due grandezze, anche se debole e non facilmente esplicitabile, e che la rete possa impararla se opportunamente allenata.

2.1.1 Raccolta dei dati

In un primo momento abbiamo raccolto i dati storici a nostra disposizione. Gli anni che abbiamo considerato sono 2017-2021. I prezzi sono stati presi

sia relativi all'intero territorio nazionale, sia considerando solo l'area Nord, la quale comprende, in base alla divisione in zone di prezzo stabilita dal GME (Gestore Mercati Energetici), Val D'Aosta, Piemonte, Liguria, Lombardia, Trentino, Veneto, Friuli Venezia Giulia, Emilia Romagna. Oltre a raccogliere questi dati, ci siamo concentrati su alcune variabili di contorno che ci sembrava potessero avere un impatto sul problema. Abbiamo quindi estrapolato anche informazioni riguardanti la temperatura e il periodo dell'anno, per esempio il numero di giorni festivi all'interno del mese. In questo modo abbiamo creato un database più ricco possibile con l'obiettivo di scegliere in fase di test quali fossero i parametri più utili all'allenamento della rete.

Questo compito è stato svolto utilizzando Python (versione 3.8.8). Abbiamo implementato varie funzioni in grado di estrapolare dai file Excel di dati le informazioni di interesse, consentendo diverse scelte parametriche. Abbiamo poi scelto l'architettura della rete residuale da allenare. In particolare abbiamo utilizzato la rete **ResNet50** del pacchetto *Keras*, la cui costruzione si basa principalmente su [5] e [6].

Tale rete fa uso di una funzione di attivazione ReLU (*Rectified Linear Unit*) e nello strato finale abbiamo aggiunto una funzione lineare che consente di trasformare il problema da classificazione a regressione. Per quanto riguarda la discesa stocastica del gradiente viene utilizzato il metodo Adam (*Adaptive Moment Estimation*). Infine per calcolare l'errore si è scelta una funzione *terminal loss* di tipo MSE (*Mean Squared Error*). Nella costruzione della rete neurale sono poi coinvolti alcuni parametri:

- *epochs*: il numero di epoche è il numero di volte in cui la rete "passa" sugli stessi dati. La rete neurale, infatti, viene allenata fornendo un determinato dataset, interamente preso in esame più volte, per ottimizzare iterativamente i pesi;
- *learning rate*: parametro coinvolto nella discesa del gradiente che controlla quanto velocemente vengono modificati i pesi della rete neurale;
- *batch size*: indica la cardinalità dei sottoinsiemi di dati che vengono usati per rendere più efficiente la discesa stocastica del gradiente.

2.1.2 Allenamento della rete e scelta dei parametri

A questo punto ci siamo concentrati sull'allenamento della rete, per poter scegliere quali informazioni fornire e come fissare i parametri. Le nostre valutazioni si sono basate soprattutto sullo studio dell'errore commesso dalla rete, sia in fase di training, quindi durante l'allenamento vero e proprio, che in fase di test. Con fase di test intendiamo la parte finale dell'allenamento,

ovvero il momento in cui la rete prende in esame il *validation set* e fornisce una stima dell'errore (nel seguito chiamato errore di previsione) su un insieme di dati non acquisiti durante il training. Abbiamo usato i dati relativi agli anni 2017-2018 come *training set* e quelli relativi al 2019 come *validation set*.

Da subito abbiamo escluso i dati riguardanti le temperature, che erano stati acquisiti in un primo momento in quanto sicuramente influenti, ma di cui difficilmente si hanno previsioni future affidabili. Questi dati quindi avrebbero potuto supportare l'allenamento della rete, ma non si sarebbero potuti usare nel momento della previsione di scenari futuri.

Dunque la rete è stata allenata fornendo, per ogni coppia di mesi, una coppia di dati (x, y) , dove:

- x è un array di dimensioni $(75,32,1)$ ¹ e rappresenta l'input contenente: la differenza relativa di prezzo per ogni sottostante e per ogni delivery dell'anno di interesse tra i due mesi (quindi $72=12 \times 6$ componenti); i due mesi coinvolti (2 componenti); la distanza tra i mesi (espressa in mesi);
- y è un array di dimensioni $(72,1)$ e rappresenta l'output contenente la differenza relativa di volume per ogni sottostante e per ogni delivery dell'anno di interesse tra i due mesi.

Valutando gli errori commessi dalla rete abbiamo fatto alcune scelte riguardanti il set di informazioni ottimali su cui basare l'allenamento. In particolare abbiamo scelto di considerare solo le coppie di mesi consecutivi, in quanto l'errore di previsione commesso dalla rete risultava decisamente maggiore per coppie di mesi a distanza qualsiasi, come dimostra il grafico in Figura 2.1. Abbiamo anche escluso le informazioni relative alle festività, perché la presenza di questo dato non migliorava la prestazione della rete in modo significativo. Inoltre si è evidenziata una preferenza per i dati di prezzo riguardanti l'intero territorio nazionale, invece che solo il Nord. Infine, abbiamo svolto alcuni test per scegliere se lavorare con la media mensile dei prezzi per ogni sottostante o con i prezzi giornalieri, che ci avrebbero consentito di ampliare le dimensioni del database, con l'auspicio di migliorare l'apprendimento della rete. In Figura 2.2 si può notare che fornendo alla rete i dati relativi ad un numero di giorni sempre maggiore, non si ottiene purtroppo una conseguente diminuzione dell'errore in fase di test, il quale raggiunge in alcuni casi picchi piuttosto alti. In più, aumentando il numero di giorni aumenta in modo consistente anche il tempo necessario per le simulazioni.

¹La presenza del numero 32 è giustificata dal fatto che la rete residuale scelta era predisposta per il riconoscimento di immagini e presentava quindi alcuni vincoli tecnici sulle dimensioni dell'input.

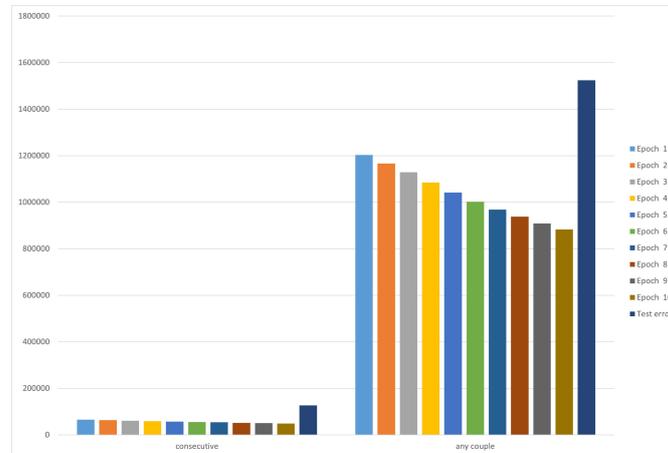


Figura 2.1: Errore in fase di training (per ogni epoca) e in fase di test: a sinistra per coppie di mesi consecutivi, a destra per coppie di mesi a distanza qualsiasi. Si nota che per mesi a distanza qualsiasi l'errore commesso dalla rete ha un ordine di grandezza molto maggiore rispetto al caso con mesi consecutivi, in particolare per quanto riguarda l'errore di previsione.

Abbiamo optato quindi per l'utilizzo della media mensile dei prezzi per ogni sottostante, in quanto l'utilizzo dei prezzi giornalieri non apporta vantaggi significativi.

In questa fase del lavoro abbiamo scelto anche i valori più opportuni per i parametri della rete precedentemente descritti. In particolare abbiamo notato che il valore della *batch size* non era così influente, a patto di prenderlo abbastanza grande rispetto alle dimensioni del dataset. In questo caso il dataset ha meno di un centinaio di elementi e il valore che è risultato più appropriato per la *batch size* è 100, quindi di fatto tutto il dataset viene preso in esame in una sola volta. Per quanto riguarda il numero di epoche abbiamo eletto 15, ossia un valore intermedio tra quelli testati, in quanto aumentare molto il numero di epoche aumentava i tempi di training e non garantiva una diminuzione significativa dell'errore di previsione. Infine il *learning rate* è risultato essere il parametro decisivo per l'errore in fase di test: un valore troppo alto (per esempio 0.1) portava con maggiore probabilità a un errore di previsione alto, dunque abbiamo scelto un valore piccolo, corrispondente a $1e-5$.

Infine abbiamo svolto un'ulteriore indagine riguardante il fattore di normalizzazione opportuno per le variazioni di volume. Abbiamo testato come possibili fattori: il valore del volume noto per ogni sottostante, per ogni delivery; il massimo in valore assoluto al variare delle scadenze sui volumi



Figura 2.2: Errore in fase di training (per ogni epoca) e in fase di test per la media mensile dei prezzi e al variare del numero di giorni considerati. Aumentare il numero di giorni comporta una buona diminuzione dell'errore in fase di training, ma non necessariamente in fase di test, dove si riscontrano picchi consistenti. L'errore in fase di training è riferito all'asse y a sinistra, mentre quello di previsione all'asse secondario.

noti per ogni sottostante; la deviazione standard al variare delle scadenze sui volumi noti per ogni sottostante; tre volte la deviazione standard al variare delle scadenze sui volumi noti per ogni sottostante. Non avendo riscontrato differenze particolarmente significative, abbiamo preferito mantenere la differenza relativa per la sua interpretazione più intuitiva.

2.1.3 Simulazione di volumi

Nell'ultima fase del lavoro abbiamo testato la capacità predittiva della rete neurale costruita. Abbiamo quindi fornito un *test set* contenente i dati relativi agli anni 2020-2021, che non erano stati usati per il training, e abbiamo confrontato l'output prodotto dalla rete con i volumi effettivi a nostra disposizione. Ricordiamo che la rete non simula il valore assoluto del volume, ma, data una coppia di mesi, la variazione relativa del volume. Conoscendo il fattore di normalizzazione, l'output prodotto dalla rete è stato quindi denormalizzato per consentire un confronto con i dati storici.

Nei grafici riportati in Figura 2.3 è rappresentato l'errore percentuale medio, massimo, minimo e in variazione standard commesso dalla rete nella previsione dei volumi per gli anni 2020-2021.

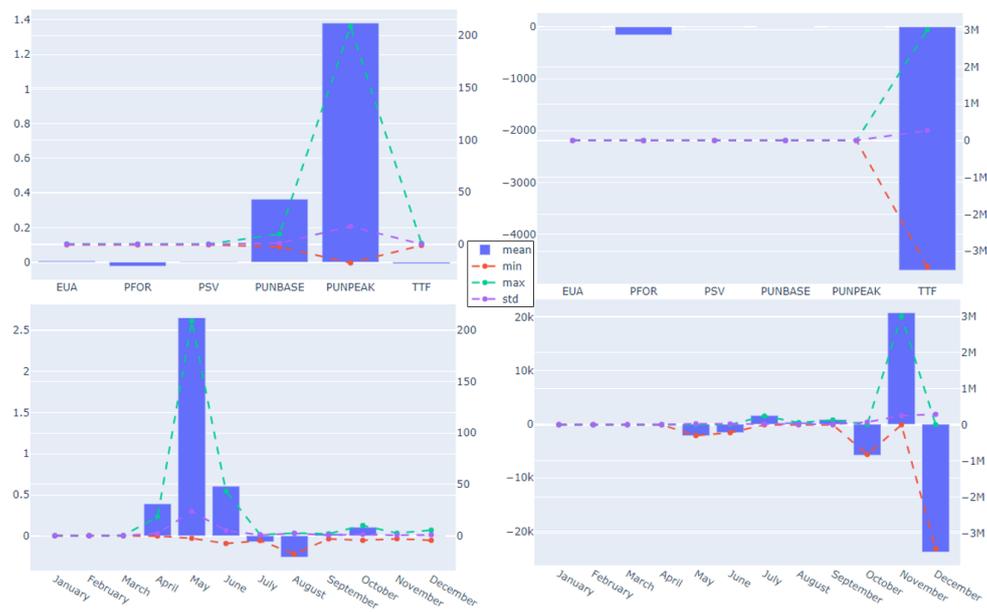


Figura 2.3: Errore percentuale medio, massimo, minimo e in variazione standard per gli anni 2020-2021. In alto aggregato per sottostante, in basso per delivery; a sinistra per l'anno in corso e a destra per l'anno successivo. Si notano errori consistenti nel passaggio da stagione fredda a stagione calda per quanto riguarda l'elettricità e nell'ultimo trimestre dell'anno per quanto riguarda il gas.

Da una prima analisi qualitativa sono emerse le seguenti osservazioni: per quanto riguarda l'elettricità, la parte più difficile da prevedere è il passaggio da stagione fredda a stagione calda, un momento effettivamente delicato in cui possono cambiare i fattori in gioco e quindi i prezzi possono non essere un'indicazione sufficiente per prevedere la variazione di volume. Intervengono ad esempio i dati sulle riserve idriche (quanto ha piovuto, quanta neve si è sciolta) che possono incidere sulla programmazione degli impianti idroelettrici. Per quanto riguarda il gas, invece, la parte più difficile da prevedere è l'ultimo trimestre dell'anno. In effetti tale trimestre è particolarmente rilevante perché coinvolge volumi consistenti e rappresenta l'inizio della stagione fredda. Inoltre è l'inizio dell'anno gas, che va da ottobre a settembre dell'anno successivo, ed è quindi il momento in cui vengono stipulati contratti con nuovi clienti o interrotti contratti già esistenti. Questo dato chiaramente impatta sulla variazione dei volumi e rende la relazione prezzi-volumi più difficilmente approssimabile sulla base della storia passata.

2.2 Punti di forza e limiti

In generale i test svolti hanno mostrato che non è possibile annullare l'errore in fase di test, perché in tutti i casi c'è una soglia che viene rapidamente raggiunta e sotto la quale non si scende. Le cause di questo comportamento, a nostro parere, sono principalmente due: molto probabilmente la relazione input-output che ci siamo proposti di indagare non è così forte da essere ben approssimata dalla rete e la dimensione del dataset utilizzato è ridotta rispetto al numero di parametri della rete, cioè la storia a nostra disposizione non è abbastanza ricca. La relazione prezzi-volumi non è facilmente esplicitabile a causa del fatto che, come è già stato accennato, la variazione di volumi è legata non solo alla variazione dei prezzi, ma anche ad altri fattori riguardanti le temperature, gli impianti, le necessità dei clienti. In particolare, per quanto riguarda la produzione degli impianti, i volumi sono solitamente stabiliti da un ufficio che si occupa di considerare tutti i fattori di cui sopra e di fare previsioni, pertanto la relazione che ne risulta è tutt'altro che deterministica.

2.2.1 Confronto con altri strumenti predittivi

Per testare la capacità predittiva della rete neurale implementata, abbiamo fatto alcuni confronti con altri strumenti. In primo luogo abbiamo cercato di approssimare la relazione prezzi-volumi con una regressione lineare, pur consapevoli della non linearità di tale relazione. Lo scopo era quello di mostrare, almeno a livello qualitativo, che la rete da noi costruita fosse più efficace di

una semplice regressione lineare. Il risultato dei test svolti ha confermato questa ipotesi, in quanto l'errore di previsione dei volumi commesso dalla regressione lineare è risultato molto maggiore di quello commesso dalla rete.

Un altro strumento che abbiamo costruito è quello che abbiamo chiamato "giocatore random". In questo caso non abbiamo utilizzato i dati storici a disposizione, ma semplicemente abbiamo generato delle variazioni di volume random con due diverse distribuzioni. Abbiamo scelto la distribuzione uniforme sull'intervallo $[-1, 1]$ e la distribuzione normale standard, nell'ipotesi che le variazioni di volume da simulare fossero comunque vicine a 0. L'idea era quella di provare a capire se la rete costruita fosse in grado di trarre un reale vantaggio dall'allenamento e quindi dalla storia passata oppure se anche variazioni di volume random in qualche modo fornissero un'adeguata previsione di futuri scenari.

Ciò che è emerso da questo test è che solo in alcuni casi la rete risulta più robusta del giocatore random. In generale però non si è evidenziato un particolare vantaggio nell'utilizzo della rete, come emerge dal confronto in Figura 2.4. Questo conferma l'ipotesi, esposta precedentemente, che i volumi siano influenzati da parecchi fattori e che quindi fornire alla rete solo la variazione di prezzi non consenta una buona capacità predittiva. Inoltre per entrambi gli strumenti ci sono mesi in cui l'errore è particolarmente ampio e questi corrispondono ai casi in cui nella realtà si sono verificati dei picchi nella variazione di volume, visibili in Figura 2.5. Chiaramente fenomeni di questo tipo non sono riproducibili dal giocatore random, ma purtroppo neanche dalla rete, la quale ha la possibilità di imparare solo dai tre anni di storia precedenti.

In conclusione possiamo affermare che tutto il lavoro svolto ci ha portato alla consapevolezza che esiste certamente una relazione tra prezzi e volumi, ma ci sono altri fattori in gioco di cui la rete neurale non sa tenere conto. Lo strumento che abbiamo costruito, quindi, fornisce un supporto leggermente migliore di quello dato da un simulatore random, ma deve essere ancora migliorato per poter garantire una previsione affidabile. Sicuramente la poca storia a disposizione limita la possibilità della rete di imparare e tarare i propri parametri nel migliore dei modi.

L'ipotesi che viene formalizzata nel prossimo capitolo è che il controllo ottimo stocastico possa dare un contributo al nostro problema. L'idea è quella di interpretare la rete neurale attraverso un problema di controllo, studiarne il fondamento teorico e suggerire possibili sviluppi che possano supportare l'interpretazione della nostra applicazione.

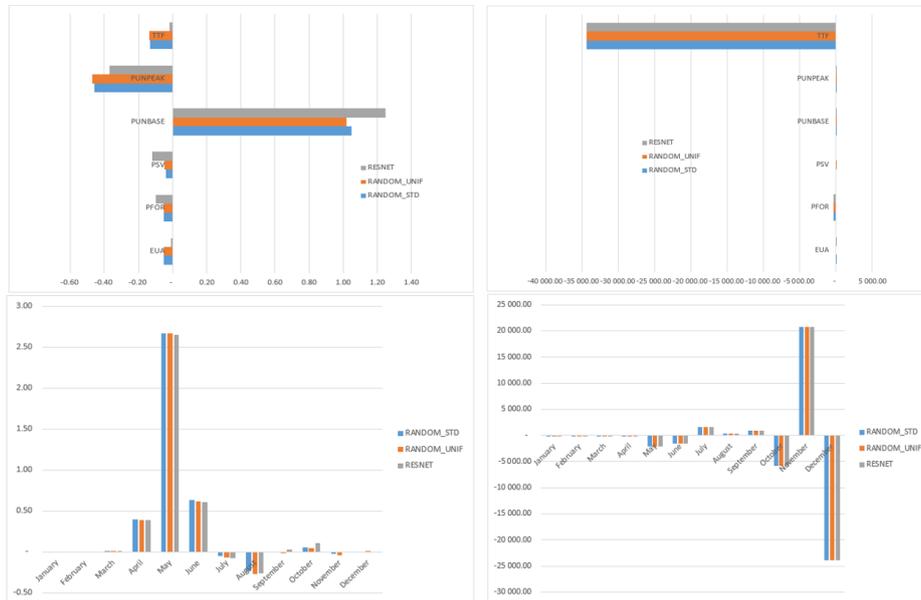


Figura 2.4: Confronto tra rete neurale e giocatore random per gli anni 2020-2021. Viene mostrata la differenza media tra i volumi effettivi e quelli previsti. In alto aggregata per sottostante, in basso per delivery; a sinistra per l'anno in corso e a destra per l'anno successivo.

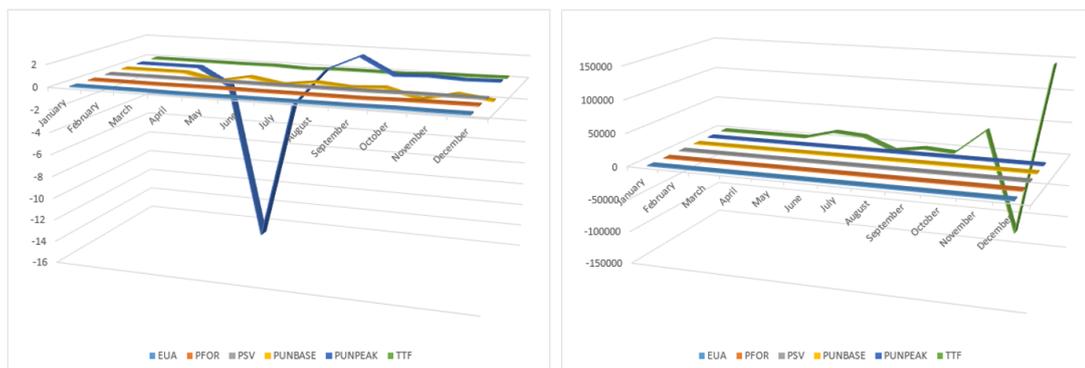


Figura 2.5: Variazioni di volume effettive in media. A sinistra per l'anno in corso, a destra per l'anno successivo. Si notano picchi in particolare per PUNPEAK e TTF, difficilmente riproducibili dagli strumenti predittivi a disposizione.

Capitolo 3

Problema di controllo

Nel seguente capitolo viene presentato il problema di controllo legato alla rete neurale residuale precedentemente introdotta. L'obiettivo è studiare dal punto di vista teorico tale problema e suggerire possibili sviluppi che possano supportare l'applicazione di reti neurali residuali. Il percorso sviluppato si basa prevalentemente su [4].

3.1 Interpretazione di una rete residuale come problema di controllo

Come illustrato nel Capitolo 1, una rete neurale residuale con T strati può essere rappresentata dalla seguente equazione alle differenze:

$$x_{t+1} = x_t + f(x_t, \theta_t), \quad t = 0, \dots, T-1, \quad (3.1)$$

dove x_0 è l'input, x_T è l'output, f è la funzione di attivazione e θ_t è il vettore dei parametri della rete nello strato t . L'obiettivo è scegliere i parametri $\theta_0, \dots, \theta_{T-1}$ in modo tale da minimizzare la distanza tra il valore finale x_T e un determinato valore obiettivo y_0 .

Per formulare il problema di controllo introduciamo un opportuno contesto probabilistico. Sia $(\Omega, \mathcal{F}, \mathbb{P})$ uno spazio di probabilità "sufficientemente ricco" (vedi Lemma 3.2.1). Supponiamo che $x_0 : \Omega \rightarrow \mathbb{R}^d$ e $y_0 : \Omega \rightarrow \mathbb{R}^l$ siano variabili aleatorie con legge congiunta $\mu_0 = \mathbb{P}_{(x_0, y_0)}$. Consideriamo un insieme di controlli ammissibili $\Theta \subseteq \mathbb{R}^m$ con Θ boreliano. Sia $T > 0$ la profondità della rete e siano f, Φ, L rispettivamente la funzione di attivazione, la funzione *terminal loss* e il regolarizzatore introdotti nell'ultima sezione del Capitolo 1. In particolare:

$$f : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^d, \quad \Phi : \mathbb{R}^d \times \mathbb{R}^l \rightarrow \mathbb{R}, \quad L : \Theta \rightarrow \mathbb{R}.$$

Nella descrizione della rete residuale come problema di controllo ottimo a tempo continuo, si suppone che ci sia un'infinità di strati, quindi si sostituisce l'equazione alle differenze (3.1) con la seguente equazione differenziale controllata da $\theta: [0, T] \rightarrow \Theta$:

$$\dot{x}_t = f(x_t, \theta_t), \quad (3.2)$$

con condizione iniziale data dalla variabile aleatoria x_0 . Si ha quindi un'equazione differenziale stocastica, la cui stocasticità deriva solo dalla condizione iniziale.

Considerando come insieme dei controlli $L^\infty([0, T], \Theta)$, il problema di controllo può essere espresso nel modo seguente:

$$\inf_{\theta \in L^\infty([0, T], \Theta)} J(\theta) := \mathbb{E} \left[\Phi(x_T, y_0) + \int_0^T L(\theta_t) dt \right] \quad (3.3)$$

con equazione di stato (3.2). In questo modo la struttura stratificata tipica della rete neurale, che consente di operare successive composizioni, viene rappresentata attraverso un sistema dinamico.

3.2 Notazioni ed ipotesi sui coefficienti

3.2.1 Notazioni

Introduciamo alcune notazioni che ci permetteranno di esplicitare l'equazione di stato del problema di controllo. Consideriamo la coppia $z = (x, y) \in \mathbb{R}^{d+l}$ data dalla concatenazione dell'input $x \in \mathbb{R}^d$ e dell'output $y \in \mathbb{R}^l$. Estendiamo le funzioni precedentemente definite nel modo seguente: $\bar{f}(z, \theta) := (f(x, \theta), 0)$ è la funzione di attivazione $(d+l)$ -dimensionale e $\bar{\Phi}(z) := \Phi(x, y)$ è la funzione *terminal loss*.

Utilizziamo la notazione $L^2(\Omega, \mathbb{R}^{d+l})$ per indicare l'insieme delle variabili aleatorie $(d+l)$ -dimensionali di quadrato sommabili definite su $(\Omega, \mathcal{F}, \mathbb{P})$. Si noti che $L^2(\Omega, \mathbb{R}^{d+l})$ è uno spazio di Hilbert con la norma $\|X\|_{L^2} := \mathbb{E}[|X|^2]^{\frac{1}{2}}$, $X \in L^2(\Omega, \mathbb{R}^{d+l})$, dove $|\cdot|$ indica la norma euclidea su \mathbb{R}^{d+l} . Denotiamo poi con $\mathcal{P}_2(\mathbb{R}^{d+l})$ l'insieme delle misure di probabilità di quadrato sommabili sullo spazio $(\mathbb{R}^{d+l}, \mathcal{B}(\mathbb{R}^{d+l}))$, ovvero

$$\mathcal{P}_2(\mathbb{R}^{d+l}) := \left\{ \mu \in \mathcal{P}(\mathbb{R}^{d+l}) : \int_{\mathbb{R}^{d+l}} (|x|^2 + |y|^2) \mu(dx, dy) < +\infty \right\},$$

dove $\mathcal{P}(\mathbb{R}^{d+l})$ è l'insieme di tutte le misure di probabilità su $(\mathbb{R}^{d+l}, \mathcal{B}(\mathbb{R}^{d+l}))$. Si noti che $X \in L^2(\Omega, \mathbb{R}^{d+l})$ se e solo se $\mathbb{P}_X \in \mathcal{P}_2(\mathbb{R}^{d+l})$, dove \mathbb{P}_X denota la

legge della variabile aleatoria X . Lo spazio $\mathcal{P}_2(\mathbb{R}^{d+l})$ è uno spazio metrico dotato della seguente metrica, detta *distanza di Wasserstein*:

$$W_2(\mu, \nu) := \inf \left\{ \left(\int_{\mathbb{R}^{d+l} \times \mathbb{R}^{d+l}} \|z - w\|^2 \pi(dz, dw) \right)^{\frac{1}{2}} \mid \pi \in \mathcal{P}_2(\mathbb{R}^{d+l} \times \mathbb{R}^{d+l}) : \pi(\cdot \times \mathbb{R}^{d+l}) = \mu, \pi(\mathbb{R}^{d+l} \times \cdot) = \nu \right\}.$$

Si chiede cioè che μ e ν siano le marginali di $\pi \in \mathcal{P}_2(\mathbb{R}^{d+l} \times \mathbb{R}^{d+l})$.

Lo spazio di probabilità $(\Omega, \mathcal{F}, \mathbb{P})$ deve essere "sufficientemente ricco", ovvero l'insieme delle leggi di tutte le variabili aleatorie in $L^2(\Omega, \mathbb{R}^{d+l})$ deve coincidere con $\mathcal{P}_2(\mathbb{R}^{d+l})$. Una condizione necessaria e sufficiente affinché valga questa proprietà è data dal seguente risultato. Si rimanda a [3] per ulteriori dettagli.

Lemma 3.2.1. $(\Omega, \mathcal{F}, \mathbb{P})$ è "sufficientemente ricco" nel senso che

$$\mathcal{P}_2(\mathbb{R}^{d+l}) = \{\mathbb{P}_\xi : \xi \in L^2(\Omega, \mathbb{R}^{d+l})\}$$

se e solo se esiste $U : \Omega \rightarrow \mathbb{R}$ variabile aleatoria \mathcal{F} -misurabile tale che $\mathbb{P}_U = \text{Unif}_{[0,1]}$.

Dimostrazione. (\Rightarrow) Consideriamo lo spazio di probabilità $([0, 1], \mathcal{B}([0, 1]), \text{Unif}_{[0,1]})$. Sia $\{e_n\}_{n \in \mathbb{N}}$ la base canonica di \mathbb{R}^{d+l} e sia $\mathcal{I} : [0, 1] \rightarrow \mathbb{R}^{d+l}$ la mappa definita da

$$\mathcal{I}(r) := r e_1 \quad \forall r \in [0, 1].$$

\mathcal{I} è continua e quindi misurabile. Sia μ la legge della variabile aleatoria \mathcal{I} . Si noti che $\mu \in \mathcal{P}_2(\mathbb{R}^{d+l})$. Dall'ipotesi segue che esiste una variabile aleatoria \mathcal{F} -misurabile $\xi : \Omega \rightarrow \mathbb{R}^{d+l}$ con legge μ . Definiamo la mappa $\Pi : \mathbb{R}^{d+l} \rightarrow \mathbb{R}$ come segue

$$\Pi(x) := \langle x, e_1 \rangle \quad \forall x \in \mathbb{R}^{d+l}.$$

Consideriamo allora la variabile aleatoria $U : \Omega \rightarrow \mathbb{R}$ data da $U := \Pi(\xi)$. Si noti che U è \mathcal{F} -misurabile. Inoltre U ha distribuzione uniforme su $[0, 1]$, infatti vale:

$$\begin{aligned} \mathbb{P}(U \in B) &= \mathbb{P}(\Pi(\xi) \in B) = \mathbb{P}(\langle \xi, e_1 \rangle \in B) = \\ &= \mu(\{x \in \mathbb{R}^{d+l} : \langle x, e_1 \rangle \in B\}) = \\ &= \text{Unif}_{[0,1]}(\{r \in [0, 1] : \langle r e_1, e_1 \rangle \in B\}) = \\ &= \text{Unif}_{[0,1]}(B \cap [0, 1]) = \text{Unif}_{[0,1]}(B) \end{aligned}$$

per ogni sottoinsieme boreliano B di \mathbb{R} .

(\Leftarrow) Fissiamo $\mu \in \mathcal{P}_2(\mathbb{R}^{d+l})$ e consideriamo lo spazio di probabilità $([0, 1], \mathcal{B}([0, 1]), \text{Unif}_{[0,1]})$. Data μ , per il Teorema 3.19 in [8] esiste una variabile aleatoria $\Xi : [0, 1] \rightarrow \mathbb{R}^{d+l}$ con legge uguale a μ . Definiamo allora

$$\xi(\omega) := \Xi(U(\omega)) \quad \forall \omega \in \Omega.$$

ξ è una variabile aleatoria \mathcal{F} -misurabile con legge μ . □

3.2.2 Ipotesi sui coefficienti

Chiediamo che i coefficienti introdotti verifichino le seguenti ipotesi **(H)**, espresse in termini delle loro estensioni $(d+l)$ -dimensionali $\bar{f}(z, \theta) := (f(x, \theta), 0)$ e $\bar{\Phi}(z) := \Phi(x, y)$:

- Funzione di attivazione $f : \mathbb{R}^d \times \Theta \rightarrow \mathbb{R}^d$: boreliana, lipschitziana in x uniformemente rispetto a θ e limitata:

$$|\bar{f}(z, \theta) - \bar{f}(z', \theta)| \leq K|z - z'|,$$

$$|\bar{f}(z, \theta)| \leq M.$$

- Funzione *terminal loss* $\Phi : \mathbb{R}^d \times \mathbb{R}^l \rightarrow \mathbb{R}$: non negativa, boreliana e vale la seguente condizione che implica la crescita quadratica di $\bar{\Phi}$ e la locale lipschitzianità:

$$|\bar{\Phi}(z) - \bar{\Phi}(z')| \leq C(1 + |z| + |z'|)|z - z'|.$$

- Regolarizzatore $L : \Theta \rightarrow \mathbb{R}$: non negativa, boreliana e a crescita quadratica in θ :

$$|L(\theta)| \leq C(1 + |\theta|^2).$$

Osservazione 3.2.1. L'ipotesi di limitatezza della funzione di attivazione risulta necessaria per i risultati proposti in seguito. Nelle applicazioni, come abbiamo visto, si utilizza spesso la funzione ReLU, ovvero la parte positiva, superiormente illimitata. Stiamo quindi operando una sorta di troncamento che risulta lecito tenendo conto del fatto che la costante M può essere comunque arbitrariamente grande e a livello computazionale non sarà possibile andare oltre un certo valore soglia, che dipende dal tipo di computer utilizzato. Un'altra ipotesi che si potrebbe adottare per risolvere questa incoerenza è la limitatezza dell'insieme $\Theta \subseteq \mathbb{R}^m$ in cui prende valore il controllo.

3.3 Equazione di stato

Introduciamo ora l'equazione di stato del nostro problema di controllo. Dati $t \in [0, T]$, $\xi \in L^2(\Omega, \mathbb{R}^{d+l})$ condizione iniziale ed un processo di controllo $\theta \in L^\infty([0, T], \Theta)$, consideriamo per $t \leq s \leq T$:

$$Z_s^{t,\xi,\theta} = \xi + \int_t^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr. \quad (3.4)$$

Diamo la definizione di soluzione di tale equazione.

Definizione 3.3.1. Per ogni $t \in [0, T]$, $\xi \in L^2(\Omega, \mathbb{R}^{d+l})$, $\theta \in L^\infty([0, T], \Theta)$, una *soluzione dell'equazione differenziale stocastica* (3.4) è un processo $Z^{t,\xi,\theta} : [0, T] \times \Omega \rightarrow \mathbb{R}^{d+l}$ continuo e $\mathcal{B}([0, T]) \otimes \mathcal{F}$ -misurabile che verifica:

$$\begin{aligned} \bar{f}(Z^{t,\xi,\theta}, \theta) &\in L^1([t, T]), \\ Z_s^{t,\xi,\theta} &= \xi + \int_t^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr \quad q.c. \quad s \geq t. \end{aligned}$$

Inoltre $Z_s^{t,\xi,\theta} = \xi \quad q.c. \quad s < t$.

Proposizione 3.3.1. Sotto le ipotesi **(H)**, esiste una soluzione di (3.4). Tale soluzione è unica a meno di indistinguibilità¹.

Inoltre vale la *proprietà di flusso* per $0 \leq t \leq \hat{t} \leq s \leq T$:

$$Z_s^{t,\xi,\theta} = Z_s^{\hat{t}, Z_{\hat{t}}^{t,\xi,\theta}, \theta} \quad q.c. \quad s \geq \hat{t}, \quad (3.5)$$

ovvero la restrizione di $Z^{t,\xi,\theta}$ a $[\hat{t}, T]$ e $Z^{\hat{t}, Z_{\hat{t}}^{t,\xi,\theta}, \theta}$ sono processi indistinguibili.

Dimostrazione. Per la dimostrazione di esistenza e unicità della soluzione si rimanda a [10].

Dimostriamo la proprietà di flusso. Per farlo è sufficiente mostrare che i due processi $(Z_s^{t,\xi,\theta})_{s \in [t, T]}$ e $(Z_s^{\hat{t}, Z_{\hat{t}}^{t,\xi,\theta}, \theta})_{s \in [\hat{t}, T]}$ verificano la stessa equazione differenziale stocastica:

$$\begin{aligned} Z_s^{t,\xi,\theta} &= \xi + \int_t^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr = \\ &= \xi + \int_t^{\hat{t}} \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr + \int_{\hat{t}}^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr = \\ &= Z_{\hat{t}}^{t,\xi,\theta} + \int_{\hat{t}}^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr = Z_s^{\hat{t}, Z_{\hat{t}}^{t,\xi,\theta}, \theta}. \end{aligned}$$

¹Due processi $X = (X_t)_{t \in I}$ e $Y = (Y_t)_{t \in I}$ si dicono *indistinguibili* se l'insieme

$$(X = Y) := \{\omega \in \Omega \mid X_t(\omega) = Y_t(\omega) \quad \forall t \in I\}$$

è quasi certo.

Dunque per unicit  della soluzione, i due processi sono indistinguibili. \square

Nei seguenti risultati stabiliamo alcune stime che sono verificate dalla soluzione di (3.4).

Lemma 3.3.1. Sotto le ipotesi **(H)**, valgono le seguenti stime di dipendenza dalle condizioni iniziali per la soluzione $Z^{t,\xi,\theta}$ di (3.4):

$$\mathbb{E} \left[\sup_{t \leq s \leq t+h} |Z_s^{t,\xi,\theta} - \xi|^2 \right] \leq Ch \quad (3.6)$$

$$\mathbb{E} \left[\sup_{t \leq s \leq T} |Z_s^{t,\xi,\theta} - Z_s^{t,\eta,\theta}|^2 \right] \leq C \mathbb{E}[|\xi - \eta|^2] \quad (3.7)$$

per ogni $t \in [0, T]$, $\xi, \eta \in L^2(\Omega, \mathbb{R}^{d+l})$, $\theta \in L^\infty([0, T], \Theta)$ e per qualche costante $C \geq 0$ indipendente da t, ξ, η, θ .

Dimostrazione. Dalla (3.4) segue che

$$Z_s^{t,\xi,\theta} - \xi = \int_t^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr$$

da cui, essendo \bar{f} limitata da una costante M per l'ipotesi **(H)**, si ottiene

$$\begin{aligned} |Z_s^{t,\xi,\theta} - \xi|^2 &= \left| \int_t^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr \right|^2 \leq \\ &\leq \left(\int_t^s |\bar{f}(Z_r^{t,\xi,\theta}, \theta_r)| dr \right)^2 \leq \\ &\leq M^2 (s-t)^2 \leq M^2 T (s-t) = C(s-t). \end{aligned}$$

Quindi

$$\sup_{t \leq s \leq t+h} |Z_s^{t,\xi,\theta} - \xi|^2 \leq Ch$$

che dimostra la (3.6).

Per la (3.7), a partire dalla (3.4) si nota che

$$Z_s^{t,\xi,\theta} - Z_s^{t,\eta,\theta} = \xi - \eta + \int_t^s (\bar{f}(Z_r^{t,\xi,\theta}, \theta_r) - \bar{f}(Z_r^{t,\eta,\theta}, \theta_r)) dr,$$

da cui

$$\begin{aligned} |Z_s^{t,\xi,\theta} - Z_s^{t,\eta,\theta}|^2 &\leq 2 \left(|\xi - \eta|^2 + \left(\int_t^s |\bar{f}(Z_r^{t,\xi,\theta}, \theta_r) - \bar{f}(Z_r^{t,\eta,\theta}, \theta_r)| dr \right)^2 \right) \leq \\ &\leq 2 \left(|\xi - \eta|^2 + T \int_t^s |\bar{f}(Z_r^{t,\xi,\theta}, \theta_r) - \bar{f}(Z_r^{t,\eta,\theta}, \theta_r)|^2 dr \right) \leq \\ &\leq 2|\xi - \eta|^2 + 2K^2 T \int_t^s |Z_r^{t,\xi,\theta} - Z_r^{t,\eta,\theta}|^2 dr, \end{aligned}$$

dove nella seconda disuguaglianza si è sfruttata la lipschitzianità di \bar{f} , ovvero

$$|\bar{f}(z, \theta) - \bar{f}(z', \theta)| \leq K|z - z'|.$$

Per concludere si sfrutta il lemma di Gronwall (cfr. [10]) ponendo

$$v(s) = |Z_s^{t,\xi,\theta} - Z_s^{t,\eta,\theta}|^2$$

e ottenendo quindi

$$v(s) \leq 2|\xi - \eta|^2 e^{2K^2 T(s-t)} \leq 2e^{2K^2 T^2} |\xi - \eta|^2 = C|\xi - \eta|^2$$

che dimostra la (3.7). \square

Proposizione 3.3.2. Sotto le ipotesi **(H)**, per ogni $t, s \in [0, T]$, con $t \leq s$, $\xi, \eta \in L^2(\Omega, \mathbb{R}^{d+l})$, $\theta \in L^\infty([0, T], \Theta)$ valgono le seguenti stime:

$$\mathbb{E} \left[\sup_{t \leq s \leq T} |Z_s^{t,\xi,\theta}|^2 \right] \leq C(1 + \mathbb{E}[|\xi|^2]) \quad (3.8)$$

$$\mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{t,\xi,\theta} - Z_r^{s,\eta,\theta}|^2 \right] \leq C \left(1 + |t - s| + \mathbb{E}[|\xi - \eta|^2] \right) \quad (3.9)$$

con C costante indipendente da t, s, ξ, η, θ .

Dimostrazione. Per l'ipotesi **(H)**, \bar{f} è limitata da una costante M , quindi si ha che

$$\begin{aligned} |Z_s^{t,\xi,\theta}|^2 &= \left| \xi + \int_t^s \bar{f}(Z_r^{t,\xi,\theta}, \theta_r) dr \right|^2 \leq \\ &\leq \left(|\xi| + \int_t^s |\bar{f}(Z_r^{t,\xi,\theta}, \theta_r)| dr \right)^2 \leq \\ &\leq 2 \left(|\xi|^2 + \left(\int_t^s |\bar{f}(Z_r^{t,\xi,\theta}, \theta_r)| dr \right)^2 \right) \\ &\leq 2|\xi|^2 + 2M^2(s-t)^2 \leq \\ &\leq 2|\xi|^2 + 2M^2T^2 = C(1 + |\xi|^2). \end{aligned}$$

Dunque vale la (3.8). Si noti che la costante C non dipende da t, ξ, θ .

La (3.9) segue invece dalle stime di dipendenza dalle condizioni iniziali riportate nel Lemma 3.3.1. Infatti si ha che

$$\begin{aligned}
& \mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{t,\xi,\theta} - Z_r^{s,\eta,\theta}|^2 \right] \leq \\
& \leq 2\mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{t,\xi,\theta} - Z_r^{t,\eta,\theta}|^2 \right] + 2\mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{t,\eta,\theta} - Z_r^{s,\eta,\theta}|^2 \right] \leq \\
& \leq 2\mathbb{E} \left[\sup_{t \leq r \leq T} |Z_r^{t,\xi,\theta} - Z_r^{t,\eta,\theta}|^2 \right] + 2\mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{t,\eta,\theta} - Z_r^{s,\eta,\theta}|^2 \right] \leq \\
& \leq 2C\mathbb{E}[|\xi - \eta|^2] + 2\mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{t,\eta,\theta} - Z_r^{s,\eta,\theta}|^2 \right],
\end{aligned}$$

dove l'ultima disuguaglianza è conseguenza della stima (3.7). Per stimare l'ultimo termine usiamo la proprietà di flusso (3.5), da cui segue che $Z_r^{t,\eta,\theta}$ e $Z_s^{s,Z_s^{t,\eta,\theta},\theta}$ sono indistinguibili su $[s, T]$ e quindi

$$\mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{t,\eta,\theta} - Z_r^{s,\eta,\theta}|^2 \right] = \mathbb{E} \left[\sup_{s \leq r \leq T} |Z_r^{s,Z_s^{t,\eta,\theta},\theta} - Z_r^{s,\eta,\theta}|^2 \right] \leq$$

(per la stima (3.7))

$$\leq C\mathbb{E}[|Z_s^{t,\eta,\theta} - \eta|^2].$$

Infine la tesi segue dalla stima (3.6). \square

Osservazione 3.3.1. Ponendo $\mu := \mathbb{P}_\xi \in \mathcal{P}_2(\mathbb{R}^{d+l})$ possiamo denotare la legge di $Z_s^{t,\xi,\theta}$ con

$$\mathbb{P}_s^{t,\mu,\theta} := \mathbb{P}_{Z_s^{t,\xi,\theta}}. \tag{3.10}$$

Tale notazione è lecita perché $Z_s^{t,\xi,\theta}$ dipende solo dalla legge della variabile aleatoria ξ che rappresenta la condizione iniziale. Questo segue dal fatto che la dimostrazione dell'esistenza della soluzione $Z_s^{t,\xi,\theta}$ si basa sulla costruzione iterativa di una successione di processi (detta *successione di Picard*), la cui legge, ad ogni passo, dipende solo dalla legge della condizione iniziale ξ . Tale dipendenza si mantiene al limite.

Con questa notazione e dalla proprietà di flusso (3.5) mostrata precedentemente segue anche che

$$\mathbb{P}_s^{t,\mu,\theta} = \mathbb{P}_s^{\hat{t}, \mathbb{P}_s^{\hat{t}, \mu, \theta}}, \quad s \geq \hat{t}.$$

3.4 Funzionale costo e funzione valore

Dati $t \in [0, T]$, $\xi \in L^2(\Omega, \mathbb{R}^{d+l})$ con legge μ e $\theta \in L^\infty([0, T], \Theta)$, definiamo il funzionale costo:

$$\begin{aligned} J(t, \mu, \theta) &:= \mathbb{E} \left[\bar{\Phi}(Z_T^{t, \xi, \theta}) + \int_t^T L(\theta_s) ds \right] = \\ &= \int_{\mathbb{R}^{d+l}} \bar{\Phi}(z) \mathbb{P}_T^{t, \mu, \theta}(dz) + \int_t^T L(\theta_s) ds \end{aligned} \quad (3.11)$$

con equazione di stato (3.4). Si noti che $J(\theta)$ definito in (3.3) corrisponde a $J(0, \mu_0, \theta)$ con J appena definito.

Definiamo poi la funzione valore $V : [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l}) \longrightarrow \mathbb{R}$:

$$V(t, \mu) := \inf_{\theta \in L^\infty([0, T], \Theta)} J(t, \mu, \theta). \quad (3.12)$$

Per ogni $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$, il problema di controllo ottimo stocastico con stato μ al tempo iniziale t consiste nel trovare, se esiste, un controllo ottimo, ossia un controllo $\hat{\theta} \in L^\infty([0, T], \Theta)$ tale che

$$V(t, \mu) = J(t, \mu, \hat{\theta}).$$

Notiamo infatti che se $\hat{\theta}$ realizza l'estremo inferiore in (3.3), allora per definizione si ha $J(t, \mu, \hat{\theta}) = V(t, \mu)$.

Proposizione 3.4.1. (Crescita quadratica di J e V). Supponiamo che valgano le ipotesi **(H)** e che $\xi \in L^2(\Omega, \mathbb{R}^{d+l})$ sia una variabile aleatoria con legge $\mu \in \mathcal{P}_2(\mathbb{R}^{d+l})$. Allora per ogni $t \in [0, T]$ e $\theta \in L^\infty([0, T], \Theta)$ vale che

$$|J(t, \mu, \theta)| \leq C(1 + \mathbb{E}[|\xi|^2]) \quad (3.13)$$

con C costante indipendente da t, μ, θ . In particolare vale che

$$|V(t, \mu)| \leq C(1 + \mathbb{E}[|\xi|^2]). \quad (3.14)$$

Dimostrazione. Dalla definizione di J segue che

$$|J(t, \mu, \theta)| \leq \mathbb{E}[|\bar{\Phi}(Z_T^{t, \xi, \theta})|] + \int_t^T |L(\theta_s)| ds.$$

Dalla crescita quadratica di $\bar{\Phi}$ e L e dalla stima (3.8) si ottiene (3.13). La stima (3.14) segue direttamente dalla definizione di funzione valore e dal fatto che (3.13) vale uniformemente rispetto a $\theta \in L^\infty([0, T], \Theta)$. \square

Proposizione 3.4.2. (Continuità di J e V rispetto alla misura). Nelle ipotesi **(H)**, vale che

$$|J(t, \mathbb{P}_\xi, \theta) - J(t, \mathbb{P}_\eta, \theta)| \leq C(1 + \|\xi\|_{L^2} + \|\eta\|_{L^2})\|\xi - \eta\|_{L^2} \quad (3.15)$$

con C costante indipendente da t, ξ, η, θ . In particolare vale che

$$|V(t, \mathbb{P}_\xi) - V(t, \mathbb{P}_\eta)| \leq C(1 + \|\xi\|_{L^2} + \|\eta\|_{L^2})\|\xi - \eta\|_{L^2}. \quad (3.16)$$

Dimostrazione.

$$\begin{aligned} |J(t, \mathbb{P}_\xi, \theta) - J(t, \mathbb{P}_\eta, \theta)| &= \left| \mathbb{E} \left[\bar{\Phi}(Z_T^{t, \xi, \theta}) + \int_t^T L(\theta_s) ds - \bar{\Phi}(Z_T^{t, \eta, \theta}) - \int_t^T L(\theta_s) ds \right] \right| \leq \\ &\leq \mathbb{E} [|\bar{\Phi}(Z_T^{t, \xi, \theta}) - \bar{\Phi}(Z_T^{t, \eta, \theta})|] \leq \end{aligned}$$

(per le ipotesi **(H)** su $\bar{\Phi}$)

$$\begin{aligned} &\leq \mathbb{E} [C(1 + |Z_T^{t, \xi, \theta}| + |Z_T^{t, \eta, \theta}|) |Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}|] = \\ &= C \mathbb{E} [|Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}| + |Z_T^{t, \xi, \theta}| |Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}| + |Z_T^{t, \eta, \theta}| |Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}|] \leq \end{aligned}$$

(per linearità del valore atteso e disuguaglianza di Cauchy-Schwarz)

$$\begin{aligned} &\leq C \left(\mathbb{E} [|Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}|^2]^{\frac{1}{2}} + \mathbb{E} [|Z_T^{t, \xi, \theta}|^2]^{\frac{1}{2}} \mathbb{E} [|Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}|^2]^{\frac{1}{2}} + \right. \\ &\quad \left. + \mathbb{E} [|Z_T^{t, \eta, \theta}|^2]^{\frac{1}{2}} \mathbb{E} [|Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}|^2]^{\frac{1}{2}} \right) = \end{aligned}$$

$$= C \|Z_T^{t, \xi, \theta} - Z_T^{t, \eta, \theta}\|_{L^2} (1 + \|Z_T^{t, \xi, \theta}\|_{L^2} + \|Z_T^{t, \eta, \theta}\|_{L^2}) \leq$$

(usando (3.7) e (3.8))

$$\leq C(1 + \|\xi\|_{L^2} + \|\eta\|_{L^2})\|\xi - \eta\|_{L^2}.$$

L'analogia disuguaglianza per la funzione valore segue dal fatto che

$$|V(t, \mathbb{P}_\xi) - V(t, \mathbb{P}_\eta)| \leq \sup_{\theta \in L^\infty([0, T], \Theta)} |J(t, \mathbb{P}_\xi, \theta) - J(t, \mathbb{P}_\eta, \theta)|$$

e la disuguaglianza mostrata per J è uniforme nel controllo. \square

Per ottenere la continuità globale della funzione valore è necessario introdurre il principio della programmazione dinamica, da cui seguirà la continuità nel tempo.

3.5 Principio della programmazione dinamica

Proposizione 3.5.1. (Principio della programmazione dinamica). Nelle ipotesi **(H)**, per ogni $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$ e $s \in [t, T]$ vale

$$V(t, \mu) = \inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}) \right\} \quad (3.17)$$

dove $\mathbb{P}_s^{t, \mu, \theta}$ è definito in (3.10).

Dimostrazione. Fissato $s \in [t, T]$, poniamo

$$\Lambda(t, \mu) := \inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}) \right\}.$$

Dimostriamo le due disuguaglianze $\Lambda(t, \mu) \leq V(t, \mu)$ e $\Lambda(t, \mu) \geq V(t, \mu)$.

1) Per definizione di $\Lambda(t, \mu)$ si ha

$$\Lambda(t, \mu) \leq \int_t^s L(\theta_r) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}). \quad (3.18)$$

Fissato $\theta \in L^\infty([0, T], \Theta)$, per definizione di funzione valore si ha

$$V(s, \mathbb{P}_s^{t, \mu, \theta}) = \inf_{\beta \in L^\infty([0, T], \Theta)} \mathbb{E} \left[\int_s^T L(\beta_r) dr + \bar{\Phi}(Z_T^{s, Z_s^{t, \xi, \theta}, \beta}) \right]$$

con $\mu \in \mathcal{P}_2(\mathbb{R}^{d+l})$ legge della variabile aleatoria ξ . Scegliendo $\beta = \theta$, otteniamo

$$V(s, \mathbb{P}_s^{t, \mu, \theta}) \leq \mathbb{E} \left[\int_s^T L(\theta_r) dr + \bar{\Phi}(Z_T^{s, Z_s^{t, \xi, \theta}, \theta}) \right].$$

Per la proprietà di flusso (3.5) i processi $(Z_r^{t, \xi, \theta})_{s \leq t \leq T}$ e $(Z_r^{s, Z_s^{t, \xi, \theta}, \theta})_{s \leq t \leq T}$ sono indistinguibili, quindi

$$V(s, \mathbb{P}_s^{t, \mu, \theta}) \leq \mathbb{E} \left[\int_s^T L(\theta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \theta}) \right]. \quad (3.19)$$

Da (3.18) e (3.19) segue che

$$\begin{aligned} \Lambda(t, \mu) &\leq \mathbb{E} \left[\int_t^s L(\theta_r) dr + \int_s^T L(\theta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \theta}) \right] = \\ &= \mathbb{E} \left[\int_t^T L(\theta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \theta}) \right]. \end{aligned}$$

Infine, dall'arbitrarietà di $\theta \in L^\infty([0, T], \Theta)$ otteniamo la tesi

$$\Lambda(t, \mu) \leq \inf_{\theta \in L^\infty([0, T], \Theta)} \mathbb{E} \left[\int_t^T L(\theta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \theta}) \right] = V(t, \mu).$$

2) Sia $\varepsilon > 0$ fissato. Per definizione di $\Lambda(t, \mu)$ esiste $\theta^\varepsilon \in L^\infty([0, T], \Theta)$ tale che

$$\Lambda(t, \mu) \geq \int_t^s L(\theta_r^\varepsilon) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}) - \varepsilon. \quad (3.20)$$

Per definizione di funzione valore

$$V(s, \mathbb{P}_s^{t, \mu, \theta}) = \inf_{\beta \in L^\infty([0, T], \Theta)} \mathbb{E} \left[\int_s^T L(\beta_r) dr + \bar{\Phi}(Z_T^{s, Z_s^{t, \xi, \theta^\varepsilon}, \beta}) \right]$$

con $\mu \in \mathcal{P}_2(\mathbb{R}^{d+l})$ legge della variabile aleatoria ξ . Allora esiste $\beta^\varepsilon \in L^\infty([0, T], \Theta)$ tale che

$$V(s, \mathbb{P}_s^{t, \mu, \theta}) \geq \mathbb{E} \left[\int_s^T L(\beta_r^\varepsilon) dr + \bar{\Phi}(Z_T^{s, Z_s^{t, \xi, \theta^\varepsilon}, \beta^\varepsilon}) \right] - \varepsilon. \quad (3.21)$$

Definiamo per ogni $0 \leq r \leq T$ il processo

$$\gamma_r^\varepsilon := \theta_r^\varepsilon \mathbb{1}_{[0, s]}(r) + \beta_r^\varepsilon \mathbb{1}_{(s, T]}(r).$$

Per la proprietà di flusso (3.5) i processi $(Z_r^{s, Z_s^{t, \xi, \theta^\varepsilon}, \beta^\varepsilon})_{s \leq r \leq T}$ e $(Z_r^{t, \xi, \gamma^\varepsilon})_{s \leq r \leq T}$ sono indistinguibili, quindi la (3.21) diventa

$$V(s, \mathbb{P}_s^{t, \mu, \theta}) \geq \mathbb{E} \left[\int_s^T L(\gamma_r^\varepsilon) dr + \bar{\Phi}(Z_T^{t, \xi, \gamma^\varepsilon}) \right] - \varepsilon.$$

Da quest'ultima disuguaglianza e dalla (3.20) segue che

$$\begin{aligned} \Lambda(t, \mu) &\geq \mathbb{E} \left[\int_t^s L(\theta_r^\varepsilon) dr + \int_s^T L(\gamma_r^\varepsilon) dr + \bar{\Phi}(Z_T^{t, \xi, \gamma^\varepsilon}) \right] - 2\varepsilon = \\ &= \mathbb{E} \left[\int_t^T L(\gamma_r^\varepsilon) dr + \bar{\Phi}(Z_T^{t, \xi, \gamma^\varepsilon}) \right] - 2\varepsilon. \end{aligned}$$

Notiamo che l'uguaglianza è giustificata dal fatto che nell'intervallo $[t, s]$ i processi θ^ε e γ^ε coincidono. Quindi

$$\Lambda(t, \mu) \geq \inf_{\theta \in L^\infty([0, T], \Theta)} \mathbb{E} \left[\int_t^T L(\theta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \theta}) \right] - 2\varepsilon = V(t, \mu) - 2\varepsilon.$$

Dall'arbitrarietà di ε segue che $\Lambda(t, \mu) \geq V(t, \mu)$. \square

3.5.1 Continuità della funzione valore

Lemma 3.5.1. $\forall \varepsilon > 0 \forall (t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$ esiste $\theta^\varepsilon \in L^\infty([0, T], \Theta)$ tale che $\forall s \in [t, T]$

$$\inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}) \right\} \geq \int_t^s L(\theta_r^\varepsilon) dr + V(s, \mathbb{P}_s^{t, \mu, \theta^\varepsilon}) - \varepsilon.$$

In particolare, θ^ε può dipendere da ε, t, μ , ma non dipende da s .

Dimostrazione. Siano $\varepsilon > 0$ e $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$ fissati. Per definizione di funzione valore esiste $\theta^\varepsilon \in L^\infty([0, T], \Theta)$ tale che

$$V(t, \mu) \geq J(t, \mu, \theta^\varepsilon) - \varepsilon.$$

Allora

$$\begin{aligned} & \int_t^s L(\theta_r^\varepsilon) dr + V(s, \mathbb{P}_s^{t, \mu, \theta^\varepsilon}) - \varepsilon = \\ & = \int_t^s L(\theta_r^\varepsilon) dr + \inf_{\beta \in L^\infty([0, T], \Theta)} \mathbb{E} \left[\int_s^T L(\beta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \beta}) \right] - \varepsilon = \\ & = \inf_{\beta \in L^\infty([0, T], \Theta)} \left\{ \mathbb{E} \left[\int_t^s L(\theta_r^\varepsilon) dr + \int_s^T L(\beta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \beta}) \right] \right\} - \varepsilon. \end{aligned}$$

Definiamo per ogni $0 \leq r \leq T$ il processo

$$\gamma_r := \theta_r^\varepsilon \mathbb{1}_{[0, s]}(r) + \beta_r \mathbb{1}_{(s, T]}(r).$$

Dunque segue che

$$\begin{aligned} & \inf_{\beta \in L^\infty([0, T], \Theta)} \left\{ \mathbb{E} \left[\int_t^s L(\theta_r^\varepsilon) dr + \int_s^T L(\beta_r) dr + \bar{\Phi}(Z_T^{t, \xi, \beta}) \right] \right\} - \varepsilon = \\ & = \inf_{\beta \in L^\infty([0, T], \Theta)} \left\{ \mathbb{E} \left[\int_t^s L(\gamma_r) dr + \int_s^T L(\gamma_r) dr + \bar{\Phi}(Z_T^{t, \xi, \gamma}) \right] \right\} - \varepsilon = \\ & = \inf_{\beta \in L^\infty([0, T], \Theta)} \mathbb{E} \left[\int_t^T L(\gamma_r) dr + \bar{\Phi}(Z_T^{t, \xi, \gamma}) \right] - \varepsilon = \\ & = \inf_{\beta \in L^\infty([0, T], \Theta)} J(t, \mu, \gamma) - \varepsilon \leq J(t, \mu, \theta^\varepsilon) - \varepsilon \leq V(t, \mu) = \\ & = \inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}) \right\}, \end{aligned}$$

dove l'ultima uguaglianza segue dal principio della programmazione dinamica (3.17). \square

Proposizione 3.5.2. (Continuità della funzione valore). Nelle ipotesi **(H)**, $\forall \varepsilon > 0 \forall (t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$ esiste $\delta = \delta(\varepsilon, t, \mu) > 0$ tale che

$$|t - s| \leq \delta \quad \Rightarrow \quad |V(t, \mu) - V(s, \mu)| \leq \varepsilon.$$

In particolare la funzione valore è globalmente continua.

Dimostrazione. Siano $\varepsilon > 0$ e $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$ fissati. Sia $s \in [t, T]$. Utilizzando il principio della programmazione dinamica (3.17), si ottiene

$$|V(t, \mu) - V(s, \mu)| = \left| \inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}) \right\} - V(s, \mu) \right| \leq$$

(per il Lemma 3.5.1)

$$\begin{aligned} &\leq \left| \int_t^s L(\theta_r^\varepsilon) dr + V(s, \mathbb{P}_s^{t, \mu, \theta^\varepsilon}) - V(s, \mu) \right| + \varepsilon \leq \\ &\leq \int_t^s L(\theta_r^\varepsilon) dr + |V(s, \mathbb{P}_s^{t, \mu, \theta^\varepsilon}) - V(s, \mu)| + \varepsilon \leq \end{aligned}$$

(per la continuità della funzione valore rispetto alla misura mostrata nella Proposizione 3.4.2)

$$\leq \int_t^s L(\theta_r^\varepsilon) dr + C(1 + \|Z_s^{t, \mu, \theta^\varepsilon}\|_{L^2} + \|\xi\|_{L^2}) \|Z_s^{t, \mu, \theta^\varepsilon} - \xi\|_{L^2} + \varepsilon \leq$$

(per la stima (3.6))

$$\leq \int_t^s L(\theta_r^\varepsilon) dr + C\sqrt{s-t}(1 + \|Z_s^{t, \mu, \theta^\varepsilon}\|_{L^2} + \|\xi\|_{L^2}) + \varepsilon \leq$$

(per la stima (3.8))

$$\leq \int_t^s L(\theta_r^\varepsilon) dr + C\sqrt{s-t}(1 + \|\xi\|_{L^2}) + \varepsilon \leq$$

(per l'ipotesi **(H)** su L)

$$\leq C(s-t)(1 + \|\theta^\varepsilon\|_\infty^2) + C\sqrt{s-t}(1 + \|\xi\|_{L^2}) + \varepsilon,$$

da cui con un'opportuna scelta di δ segue la continuità nel tempo della funzione valore. In particolare, poiché vale la Proposizione 3.4.2 che garantisce la continuità rispetto alla misura (uniforme nel tempo), la funzione valore è globalmente continua. \square

Capitolo 4

Equazione di Hamilton-Jacobi-Bellman e soluzioni di viscosità

Nel seguente capitolo viene presentata la nozione di derivata rispetto alla misura di probabilità, fondamentale per giungere alla formulazione dell'equazione di Hamilton-Jacobi-Bellman per il problema di controllo precedentemente illustrato. L'obiettivo è quello di mostrare che la funzione valore è una soluzione di viscosità di tale equazione.

4.1 Derivata rispetto alla misura di probabilità

Varie nozioni di derivata per funzioni definite su spazi di misure di probabilità sono state introdotte e alcune sono particolarmente utili nella teoria del trasporto ottimo. Nella teoria del controllo ottimo stocastico la definizione più utilizzata è quella introdotta da Pierre-Louis Lions in una serie di lezioni tenute a partire dal 2006 al Collège de France. Per una presentazione di questa nozione di derivata si veda ad esempio [1].

Definizione 4.1.1. Data una funzione $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$, si definisce il suo *sollevamento* come la funzione

$$\begin{aligned} U : L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n) &\rightarrow \mathbb{R} \\ \xi &\mapsto u(\mathbb{P}_\xi). \end{aligned}$$

Poiché $L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ è uno spazio di Banach, possiamo usare la seguente nozione di derivazione.

Definizione 4.1.2. Una funzione $U : L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n) \rightarrow \mathbb{R}$ è *differenziabile nel senso di Frèchet* in ξ se esiste $v \in L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ tale che

$$\lim_{\|h\|_{L^2} \rightarrow 0} \frac{|U(\xi + h) - U(\xi) - \mathbb{E}[\langle v, h \rangle]|}{\|h\|_{L^2}} = 0.$$

Se tale v esiste, allora è unica. Quindi possiamo parlare di differenziale di U in ξ e denotare $DU(\xi) = v$.

Definizione 4.1.3. Una funzione $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ è *differenziabile nel senso di Lions* in μ se esiste $\xi \in L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n)$ con $\mu = \mathbb{P}_\xi$ tale che il sollevamento U sia differenziabile nel senso di Frèchet in ξ .

Definizione 4.1.4. Diciamo che $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ è *differenziabile con continuità nel senso di Lions* se il suo sollevamento è differenziabile con continuità nel senso di Frèchet.

Teorema 4.1.1. Sia $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ differenziabile con continuità nel senso di Lions e sia U il suo sollevamento. Allora

$$\forall \mu \in \mathcal{P}(\mathbb{R}^n) \quad \exists g : \mathbb{R}^n \rightarrow \mathbb{R}^n \quad t.c. \quad DU(\xi) = g(\xi) \quad q.c.$$

$$\forall \xi \in L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n) \quad \text{con} \quad \mathbb{P}_\xi = \mu$$

con g misurabile.

Dimostrazione. La tesi segue dalla Proposizione 5.25 in [1]. □

L'uguaglianza vale quasi certamente, quindi va intesa come $[DU(\xi)](\omega) = g(\xi(\omega))$ per quasi ogni $\omega \in \Omega$. Supponendo che μ sia la legge di ξ , si ha

$$\int_{\mathbb{R}^n} |g(x)|^2 \mu(dx) = \int_{\Omega} g(\xi(\omega)) P(d\omega) = \int_{\Omega} [DU(\xi)](\omega) P(d\omega) < \infty.$$

In particolare g è definita μ -quasi ovunque e $g \in L^2(\mathbb{R}^n, \mu; \mathbb{R}^n)$. Possiamo quindi dare la seguente definizione.

Definizione 4.1.5. Sia $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ differenziabile con continuità nel senso di Lions. Diciamo che la *derivata di u in μ* è la (classe di equivalenza in $L^2(\mathbb{R}^n, \mu; \mathbb{R}^n)$ della) funzione misurabile $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita nel Teorema 4.1.1 e la denotiamo con $\partial_\mu u(\mu)$.

Notiamo che si sceglie di usare la notazione di derivata parziale rispetto alla misura in quanto solitamente abbiamo a che fare con funzioni del tipo $u(t, \mu)$.

Specifichiamo dunque la costruzione appena vista nel caso di nostro interesse. La funzione valore $V(t, \mu)$ definita in (3.12) ha infatti come dominio $[0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$. Ci concentriamo quindi su funzioni del tipo $u(t, \mu)$ definite su $[0, T] \times \mathcal{P}_2(\mathbb{R}^n)$ e ci interessa la nozione di derivata rispetto al tempo e rispetto alla misura di probabilità.

La derivata parziale rispetto al tempo è definita come al solito, ovvero diciamo che u è *differenziabile nel tempo* in (t, μ) se esiste finito il seguente limite

$$\partial_t u(t, \mu) := \lim_{h \rightarrow 0} \frac{|u(t+h, \mu) - u(t, \mu)|}{h},$$

con $h \rightarrow 0^+$ (risp. $h \rightarrow 0^-$) se $t = 0$ (risp. $t = T$).

La definizione di derivata rispetto alla misura in questo caso è data a tempo fissato: $u(t, \cdot)$ deve essere differenziabile con continuità nel senso di Lions per ogni $t \in [0, T]$ e la derivata di u rispetto alla misura in (t, μ) è la funzione $\partial_\mu u(\mu)(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definita μ -quasi certamente. Il seguente risultato dà una condizione sufficiente affinché la derivata rispetto alla misura sia continua e univocamente determinata. In tal caso si può parlare di derivata di una specifica funzione, invece che di classe di equivalenza in $L^2(\mathbb{R}^n, \mu; \mathbb{R}^n)$.

Lemma 4.1.1. Sia $u : \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R}$ differenziabile con continuità nel senso di Lions. Per ogni $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ esiste al più una versione, ovvero una ridefinizione su un insieme di misura nulla, di $\partial_\mu u(\mu)(\cdot) \in L^2(\mathbb{R}^n, \mu; \mathbb{R}^n)$ tale che la funzione

$$\begin{aligned} \partial_\mu u : \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (\mu, x) &\longmapsto \partial_\mu u(\mu)(x) \end{aligned}$$

sia congiuntamente continua. In particolare tale versione è continua per ogni $\mu \in \mathcal{P}_2(\mathbb{R}^n)$ fissata.

Dimostrazione. La tesi segue dal Lemma 4.9 in [3]. □

Nel nostro caso, che coinvolge anche il tempo, possiamo affermare che se $U : [0, T] \times L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n) \rightarrow \mathbb{R}$ è differenziabile con continuità nel senso di Fréchet (rispetto al suo secondo argomento) e la mappa $\partial_\mu u(t, \cdot, \cdot) : \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ è continua per ogni $t \in [0, T]$, allora la mappa

$$\begin{aligned} \partial_\mu u : [0, T] \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (t, \mu, x) &\longmapsto \partial_\mu u(t, \mu)(x) \end{aligned}$$

è definita univocamente.

4.2 Formula di Itô

Definizione 4.2.1. Denotiamo con $\mathcal{C}_2^{1,1}([0, T] \times \mathcal{P}_2(\mathbb{R}^n))$ l'insieme delle funzioni $u(t, \mu) : \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n \rightarrow \mathbb{R}$ tali che

- esistono le funzioni

$$\begin{aligned}\partial_t u &: [0, T] \times \mathcal{P}_2(\mathbb{R}^n) \rightarrow \mathbb{R} \\ \partial_\mu u &: [0, T] \times \mathcal{P}_2(\mathbb{R}^n) \times \mathbb{R}^n \rightarrow \mathbb{R}^n;\end{aligned}$$

- $u, \partial_t u, \partial_\mu u$ sono continue;
- $\partial_\mu u$ soddisfa la seguente condizione di crescita:

$$|\partial_\mu u(t, \mu)(x)| \leq C(1 + |x|^2).$$

Osservazione 4.2.1. Per i risultati di unicità nella sezione precedente, $\partial_\mu u(t, \mu)(\cdot)$ è una fissata funzione, ben definita su tutto \mathbb{R}^n per ogni $\mu \in \mathcal{P}_2(\mathbb{R}^n)$.

Teorema 4.2.1. (Formula di Itô). Sia $u \in \mathcal{C}_2^{1,1}([0, T] \times \mathcal{P}_2(\mathbb{R}^n))$, $t \in [0, T]$ e $\xi \in L^2(\Omega, \mathcal{F}, P; \mathbb{R}^n)$. Sia $F : [0, T] \times \Omega \rightarrow \mathbb{R}^n$ un processo limitato $\mathcal{B}([0, T]) \otimes \mathcal{F}$ -misurabile e consideriamo il processo

$$X_s = \xi + \int_t^s F_r dr \quad s \in [t, T].$$

Allora per ogni $s \in [t, T]$ vale che

$$u(s, \mathbb{P}_{X_s}) = u(t, \mathbb{P}_\xi) + \int_t^s \partial_t u(r, \mathbb{P}_{X_r}) dr + \int_t^s \mathbb{E}[\langle F_r, \partial_\mu u(r, \mathbb{P}_{X_r})(X_r) \rangle] dr. \quad (4.1)$$

Dimostrazione. La tesi segue dalla Proposizione 5.102 in [1]. \square

4.3 Equazione di Hamilton-Jacobi-Bellman

Dal principio della programmazione dinamica si ottiene un'equazione alle derivate parziali di cui la funzione valore è una soluzione in senso opportuno. Possiamo ricavare in modo formale tale equazione supponendo che V sia sufficientemente regolare e applicando la formula di Itô (4.1) a $V(s, \mathbb{P}_s^{t, \mu, \theta})$ tra t ed s , ottenendo

$$\begin{aligned}V(s, \mathbb{P}_s^{t, \mu, \theta}) &= V(t, \mu) + \int_t^s \partial_t V(r, \mathbb{P}_r^{t, \mu, \theta}) dr + \\ &+ \int_t^s \mathbb{E}[\langle \bar{f}(Z_r^{t, \xi, \theta}, \theta_r), \partial_\mu V(r, \mathbb{P}_r^{t, \mu, \theta})(Z_r^{t, \xi, \theta}) \rangle] dr,\end{aligned} \quad (4.2)$$

dove $\mu = \mathbb{P}_\xi$. Allora dal principio della programmazione dinamica (3.17) si ha che

$$\begin{aligned} V(t, \mu) &= \inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + V(s, \mathbb{P}_s^{t, \mu, \theta}) \right\} = \\ &= \inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + V(t, \mu) + \int_t^s \partial_t V(r, \mathbb{P}_r^{t, \mu, \theta}) dr + \right. \\ &\quad \left. + \int_t^s \mathbb{E}[\langle \bar{f}(Z_r^{t, \xi, \theta}, \theta_r), \partial_\mu V(r, \mathbb{P}_r^{t, \mu, \theta})(Z_r^{t, \xi, \theta}) \rangle] dr \right\}, \end{aligned}$$

da cui

$$\inf_{\theta \in L^\infty([0, T], \Theta)} \left\{ \int_t^s L(\theta_r) dr + \int_t^s \partial_t V(r, \mathbb{P}_r^{t, \mu, \theta}) dr + \int_t^s \mathbb{E}[\langle \bar{f}(Z_r^{t, \xi, \theta}, \theta_r), \partial_\mu V(r, \mathbb{P}_r^{t, \mu, \theta})(Z_r^{t, \xi, \theta}) \rangle] dr \right\} = 0.$$

Scegliendo $s = t+h$ e dividendo per h , passando formalmente al limite $h \rightarrow 0$, si ottiene la seguente equazione su $[0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$, nota come **equazione di Hamilton-Jacobi-Bellman**:

$$\begin{cases} \partial_t V(t, \mu) + \inf_{\vartheta \in \Theta} \{ L(\vartheta) + \mathbb{E}[\langle \bar{f}(\xi, \vartheta), \partial_\mu V(t, \mu)(\xi) \rangle] \} = 0 \\ V(T, \mu) = \mathbb{E}[\bar{\Phi}(\xi)] \end{cases} \quad (4.3)$$

con $\xi \in L^2(\Omega, \mathbb{R}^{d+l})$ tale che $\mathbb{P}_\xi = \mu$.

4.4 Soluzioni di viscosità

In generale non si hanno soluzioni classiche dell'equazione di Hamilton-Jacobi-Bellman (4.3). Per questo è necessario introdurre una nozione di soluzione debole, che prende il nome di **soluzione di viscosità**, ideata da M. Crandall e P.-L. Lions all'inizio degli anni '80. In questo caso, ci interessa la nozione di soluzione di viscosità per l'equazione di Hamilton-Jacobi-Bellman nello spazio di Wasserstein.

Definizione 4.4.1. Una funzione continua $u : [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l}) \rightarrow \mathbb{R}$ si dice *sotto-soluzione di viscosità* (risp. *sopra-soluzione di viscosità*) dell'equazione di Hamilton-Jacobi-Bellman (4.3) se

- $u(T, \mu) \leq \mathbb{E}[\bar{\Phi}(\xi)]$ (risp. \geq) con $\mu = \mathbb{P}_\xi$;

- per ogni $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$ e per ogni funzione test $\psi \in \mathcal{C}_2^{1,1}([0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l}))$ tale che $u(t, \mu) = \psi(t, \mu)$ e la mappa $u - \psi$ ha un massimo (risp. minimo) globale in (t, μ) vale che

$$\partial_t \psi(t, \mu) + \inf_{\vartheta \in \Theta} \left\{ L(\vartheta) + \mathbb{E}[\langle \bar{f}(\xi, \vartheta), \partial_\mu \psi(t, \mu)(\xi) \rangle] \right\} \geq 0$$

(risp. \leq) con $\mu = \mathbb{P}_\xi$.

Definizione 4.4.2. Una funzione continua $u : [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l}) \rightarrow \mathbb{R}$ si dice *soluzione di viscosità* dell'equazione di Hamilton-Jacobi-Bellman (4.3) se è sotto-soluzione e sopra-soluzione di viscosità di (4.3).

4.4.1 Esistenza e unicità

In questa sezione vogliamo mostrare che, sotto opportune ipotesi, la funzione valore è l'unica soluzione di viscosità dell'equazione di Hamilton-Jacobi-Bellman (4.3). Per ulteriori approfondimenti si vedano [2] e [3].

Proposizione 4.4.1. Nelle ipotesi **(H)**, la funzione valore V definita in (3.12) è soluzione viscosa dell'equazione di Hamilton-Jacobi-Bellman (4.3).

Dimostrazione. La funzione valore V è continua per quanto dimostrato nella Proposizione 3.5.2 e vale $V(T, \mu) = \mathbb{E}[\bar{\Phi}(\xi)]$ per definizione. Mostriamo le proprietà di sotto e sopra-soluzione di viscosità.

Passo 1. Vogliamo mostrare che V è sotto-soluzione di viscosità di (4.3). Siano $(t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l})$ e $\psi \in \mathcal{C}_2^{1,1}([0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l}))$ funzione test tale che $V(t, \mu) = \psi(t, \mu)$ e la mappa $V - \psi$ ha un massimo globale in (t, μ) . Fissiamo un processo di controllo costante identicamente uguale a $\vartheta \in \Theta$. Dal principio della programmazione dinamica (3.17) si ha che

$$V(t, \mu) \leq \int_t^{t+h} L(\vartheta) dr + V(t+h, \mathbb{P}_{t+h}^{t, \mu, \vartheta}).$$

Poiché $V(t, \mu) = \psi(t, \mu)$ e $V - \psi$ ha massimo globale in (t, μ) , si ha che

$$\psi(t, \mu) \leq \int_t^{t+h} L(\vartheta) dr + \psi(t+h, \mathbb{P}_{t+h}^{t, \mu, \vartheta}),$$

da cui, applicando la formula di Itô (4.1) a $\psi(t+h, \mathbb{P}_{t+h}^{t, \mu, \vartheta})$ e dividendo per h , si ottiene

$$0 \leq \frac{1}{h} \left[\int_t^{t+h} \left(L(\vartheta) + \partial_t \psi(r, \mathbb{P}_r^{t, \mu, \vartheta}) + \mathbb{E}[\langle \bar{f}(Z_r^{t, \xi, \vartheta}, \vartheta), \partial_\mu \psi(r, \mathbb{P}_r^{t, \mu, \vartheta})(Z_r^{t, \xi, \vartheta}) \rangle] \right) dr \right],$$

dove $\mu = \mathbb{P}_\xi$. Sfruttando la regolarità della funzione test ψ e le ipotesi **(H)**, possiamo passare al limite per $h \rightarrow 0^+$ ottenendo

$$\partial_t \psi(t, \mu) + L(\vartheta) + \mathbb{E}[\langle \bar{f}(\xi, \vartheta), \partial_\mu \psi(t, \mu)(\xi) \rangle] \geq 0,$$

da cui segue la proprietà di sotto-soluzione di viscosità per arbitrarietà di $\vartheta \in \Theta$.

Passo 2. Vogliamo mostrare che V è sopra-soluzione di viscosità di (4.3). Siano $(t, \mu) \in [0, T) \times \mathcal{P}_2(\mathbb{R}^{d+l})$ e $\psi \in \mathcal{C}_2^{1,1}([0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l}))$ funzione test tale che $V(t, \mu) = \psi(t, \mu)$ e la mappa $V - \psi$ ha un minimo globale in (t, μ) . Per ogni $\varepsilon > 0$, dal principio della programmazione dinamica (3.17) tra t e $t + \varepsilon$, abbiamo che esiste $\theta^\varepsilon \in L^\infty([0, T], \Theta)$ tale che

$$V(t, \mu) \geq \int_t^{t+\varepsilon} L(\theta_r^\varepsilon) dr + V(t + \varepsilon, \mathbb{P}_{t+\varepsilon}^{t, \mu, \theta^\varepsilon}) - \varepsilon^2.$$

Poiché $V(t, \mu) = \psi(t, \mu)$ e $V - \psi$ ha minimo globale in (t, μ) , si ha che

$$\psi(t, \mu) \geq \int_t^{t+\varepsilon} L(\theta_r^\varepsilon) dr + \psi(t + \varepsilon, \mathbb{P}_{t+\varepsilon}^{t, \mu, \theta^\varepsilon}) - \varepsilon^2,$$

da cui, applicando la formula di Itô (4.1) a $\psi(t + \varepsilon, \mathbb{P}_{t+\varepsilon}^{t, \mu, \theta^\varepsilon})$ e dividendo per ε , si ottiene

$$0 \geq \frac{1}{\varepsilon} \left[\int_t^{t+\varepsilon} \left(L(\theta_r^\varepsilon) + \partial_t \psi(r, \mathbb{P}_r^{t, \mu, \theta^\varepsilon}) + \mathbb{E}[\langle \bar{f}(Z_r^{t, \xi, \theta^\varepsilon}, \theta_r^\varepsilon), \partial_\mu \psi(r, \mathbb{P}_r^{t, \mu, \theta^\varepsilon})(Z_r^{t, \xi, \theta^\varepsilon}) \rangle] \right) dr \right] - \varepsilon,$$

dove $\mu = \mathbb{P}_\xi$. Sommando e sottraendo le quantità opportune si ha che

$$0 \geq \frac{1}{\varepsilon} \left[\int_t^{t+\varepsilon} \left(L(\theta_r^\varepsilon) + \partial_t \psi(t, \mu)(\xi) + \mathbb{E}[\langle \bar{f}(\xi, \theta_r^\varepsilon), \partial_\mu \psi(t, \mu)(\xi) \rangle] \right) dr \right] - \varepsilon + \rho(\varepsilon),$$

con

$$\rho(\varepsilon) = \frac{1}{\varepsilon} \left[\int_t^{t+\varepsilon} \mathbb{E}[\langle \bar{f}(Z_r^{t, \xi, \theta^\varepsilon}, \theta_r^\varepsilon), \partial_\mu \psi(r, \mathbb{P}_r^{t, \mu, \theta^\varepsilon})(Z_r^{t, \xi, \theta^\varepsilon}) \rangle - \langle \bar{f}(\xi, \theta_r^\varepsilon), \partial_\mu \psi(t, \mu)(\xi) \rangle] dr + \int_t^{t+\varepsilon} (\partial_t \psi(r, \mathbb{P}_r^{t, \mu, \theta^\varepsilon}) - \partial_t \psi(t, \mu)) dr \right]$$

e $\rho(\varepsilon) \rightarrow 0$ per $\varepsilon \rightarrow 0$. Consideriamo infine

$$0 \geq \frac{1}{\varepsilon} \left[\int_t^{t+\varepsilon} \inf_{\vartheta \in \Theta} \left(L(\vartheta) + \partial_t \psi(t, \mu)(\xi) + \mathbb{E}[\langle \bar{f}(\xi, \vartheta), \partial_\mu \psi(t, \mu)(\xi) \rangle] \right) dr \right] - \varepsilon + \rho(\varepsilon),$$

da cui, facendo il limite per $\varepsilon \rightarrow 0$, segue la proprietà di sopra-soluzione di viscosità

$$\partial_t \psi(t, \mu) + \inf_{\vartheta \in \Theta} \left\{ L(\vartheta) + \mathbb{E}[\langle \bar{f}(\xi, \vartheta), \partial_\mu \psi(t, \mu)(\xi) \rangle] \right\} \leq 0.$$

□

Per ottenere l'unicità della soluzione di viscosità, dobbiamo assumere, come ipotesi aggiuntiva, che i coefficienti L e $\bar{\Phi}$ siano limitati. Inoltre ci serviremo del Teorema del Confronto 5.1 in [2].

Proposizione 4.4.2. Nelle ipotesi **(H)** e supponendo che i coefficienti L e $\bar{\Phi}$ siano limitati, la funzione valore V è l'unica soluzione di viscosità limitata e continua dell'equazione di Hamilton-Jacobi-Bellman (4.3).

Dimostrazione. La funzione valore V è limitata e continua per ipotesi ed è soluzione di viscosità di (4.3) per la Proposizione 4.4.1. Sia $u : [0, T] \times \mathcal{P}_2(\mathbb{R}^{d+l}) \rightarrow \mathbb{R}$ un'altra soluzione di viscosità di (4.3) limitata e continua. Per definizione u e V sono sotto e sopra-soluzioni di viscosità di (4.3), dunque applicando il Teorema del confronto si ottengono le due disuguaglianze $u \leq V$ e $u \geq V$, da cui $u = V$. Ne segue che V è l'unica soluzione di viscosità di (4.3). □

In conclusione possiamo affermare che i risultati proposti forniscono una caratterizzazione del controllo ottimo e quindi dei coefficienti θ che rappresentano i pesi della rete. Abbiamo infatti mostrato che l'equazione di Hamilton-Jacobi-Bellman è ben posta in senso viscoso e che sotto opportune ipotesi la funzione valore risulta essere l'unica soluzione di viscosità. In questo senso, l'equazione di Hamilton-Jacobi-Bellman fornisce condizioni necessarie e sufficienti per l'esistenza di parametri ottimali, senza però costruirli esplicitamente.

Una direzione che si potrebbe percorrere per ottenere una soluzione da cui la rete possa trarre un vantaggio è quella di risolvere il problema di controllo attraverso tecniche numeriche già sviluppate. In questo modo si otterrebbe un'approssimazione dei pesi ottimali per la rete.

Bibliografia

- [1] R. Carmona, F. Delarue, *Probabilistic theory of mean field games with applications-I*, volume 83 of Probabilistic Theory and Stochastic Modelling, Springer, 2018.
- [2] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham, M. Rosestolato, *Master Bellman equation in the Wasserstein space: Uniqueness of viscosity solutions*, Preprint arXiv: 2107.10535, 2021.
- [3] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham, M. Rosestolato, *Optimal control of path-dependent McKean-Vlasov SDEs in infinite dimension*, Preprint arXiv: 2012.14772, 2021.
- [4] W. E, J. Han, Q. Li, *A mean-field optimal control formulation of deep learning*, Research in Mathematical Sciences, 6 (1), 2019.
- [5] K. He, X. Zhang, S. Ren, J. Sun, *Deep Residual Learning for Image Recognition*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778, 2016.
- [6] K. He, X. Zhang, S. Ren, J. Sun, *Identity Mappings in Deep Residual Networks*, Proceedings of the European Conference on Computer Vision (ECCV), 630-645, 2016.
- [7] C. F. Higham, D. J. Higham, *Deep learning: An introduction for applied mathematicians*, SIAM Review, 61(4), 860-891, 2019.
- [8] O. Kallenberg, *Foundations of modern probability*, Probability and its Applications (New York), Springer-Verlag, New York, second edition, 2002.
- [9] Y. LeCun, Y. Bengio, G. Hinton, *Deep learning*, Nature, 521(7553), 436-444, 2015.
- [10] A. Pascucci, *Teoria della probabilità. Variabili aleatorie e distribuzioni*, Springer, 2020.

-
- [11] Wikipedia, *Discesa stocastica del gradiente*
https://it.wikipedia.org/wiki/Discesa_stocastica_del_gradiente
- [12] Wikipedia, *Problema della scomparsa del gradiente*
https://it.wikipedia.org/wiki/Problema_della_scomparsa_del_gradiente
- [13] Wikipedia, *Training, validation, and test sets*
https://en.wikipedia.org/wiki/Training,_validation,_and_test_sets