Università di Genova

Dipartimento di Matematica

Corso di Laurea in

Matematica

Anno accademico 2026/2027

Tesi di Laurea Magistrale

Analisi compartimentale del tracciante ^{18}F-FDOPA: problemi inversi e identificabilità nella PET cerebellare con stima della funzione di ingresso dalla regione carotidea

Candidato

Umberto Beardo

Relatore Prof.ssa Cristina Campi

Correlatore Dott.ssa Chiara Razzetta

Indice

Introduzione **4**

1 Analisi compartimentale **6**

1.1 Fondamenti dell'analisi compartimentale 6

1.1.1 Analisi compartimentale e cinetica dei traccianti 6

1.1.2 Il concetto di compartimento 7

1.1.3 Quantità di sostanza e concentrazione 8

1.1.4 Modelli compartimentali lineari 8

1.1.5 Il modello a un compartimento 9

1.1.6 Un modello a due compartimenti 9

1.1.7 Il modello a N compartimenti 10

1.2 Il problema diretto 13

1.2.1 Formulazione del problema 13

1.2.2 Esistenza e unicità della soluzione 14

1.2.3 Esponenziale di matrice 14

1.2.4 Soluzione del problema con ingresso 15

1.2.5 Non negatività della soluzione 15

1.3 Il problema inverso e l'identificabilità a priori 17

1.3.1 Il problema inverso 17

1.3.2 Parametri del modello e parametri osservazionali 18

1.3.3 Exhaustive summary 19

2 PET e FDOPA **21**

2.1 Introduzione alla PET 21

2.2 Principi fisici della PET 22

2.2.1 Decadimento beta positivo 24

2.2.2 Annichilazione 24

2.2.3 Rivelazione in coincidenza 25

2.2.4 Formazione dell'immagine PET 27

2.3 PET statica e PET dinamica . . . 28
2.3.1 Regione di interesse . . . 29
2.4 FDOPA-PET . . . 29
2.5 Ipotesi del modello cerebellare . . . 30
2.6 I compartimenti . . . 31
2.7 Input function . . . 33
2.7.1 Campionamento arterioso . . . 33
2.7.2 Stima della input function dalle carotidi . . . 34
2.7.3 L'approccio di Schiepers . . . 40
2.7.4 Uno studio sul metabolismo dell'FDOPA . . . 41
2.7.5 Input function senza campionamento arterioso . . . 42

3 Il problema inverso **43**
3.1 Formulazione del problema . . . 43
3.2 Il modello matematico . . . 44
3.2.1 Il problema in notazione matriciale . . . 44
3.2.2 La trasformata di Laplace . . . 46
3.2.3 Funzioni razionali e irriducibilità . . . 50
3.2.4 Unicità dei coefficienti . . . 51
3.2.5 Identificabilità locale . . . 63
3.2.6 Identificabilità globale . . . 65
3.3 Ipotesi fisiologiche di modello . . . 75
3.3.1 Ipotesi di Huang . . . 76
3.3.2 Ipotesi di Kumakura . . . 76
3.3.3 Identificabilità del modello vincolato . . . 77
3.3.4 I casi $q = 1$ e $0 < q < 1$ . . . 81

4 Trattazione numerica del problema inverso **84**
4.1 Introduzione . . . 84
4.2 Formulazione del problema ai minimi quadrati . . . 85
4.2.1 Il modello diretto . . . 85
4.2.2 Dati sperimentali e problema discreto . . . 86
4.2.3 La funzione obiettivo . . . 87
4.2.4 Parametrizzazione logaritmica dei coefficienti . . . 88
4.3 Derivate della funzione obiettivo . . . 89
4.3.1 Calcolo della matrice Jacobiana . . . 90
4.3.2 Rango della matrice Jacobiana . . . 92
4.3.3 Gradiente . . . 94

4.3.4 Matrice Hessiana 94
4.4 Il metodo di Gauss-Newton 94
4.4.1 Direzione di discesa 96
4.4.2 Scelta della lunghezza del passo 96
4.4.3 Convergenza del metodo di Gauss-Newton 97
4.4.4 Limiti del metodo 98
4.5 Il metodo di Levenberg-Marquardt 99
4.5.1 La regione di confidenza 99
4.5.2 Il parametro di smorzamento 100
4.5.3 Convergenza del metodo di Levenberg-Marquardt 102
4.5.4 Relazione tra convergenza numerica e identificabilità 103
4.5.5 Schema iterativo 105

Conclusioni **107**

Bibliografia **109**

Introduzione

La tomografia a emissione di positroni, generalmente indicata con l'acronimo PET (*Positron Emission Tomography*), è una tecnica di imaging che consente di osservare la distribuzione di sostanze marcate con radionuclidi emettitori di positroni. In un'acquisizione PET dinamica si ottiene una successione temporale di immagini che descrive l'evoluzione della concentrazione di attività radioattiva nelle regioni considerate, fornendo informazioni di carattere funzionale e metabolico [11].

Per interpretare tali dati è necessario introdurre un modello matematico. Una delle formulazioni maggiormente utilizzate nello studio della cinetica dei traccianti è l'analisi compartimentale, nella quale il sistema biologico viene rappresentato mediante un numero finito di compartimenti cineticamente omogenei, collegati da processi di trasporto o trasformazione. L'evoluzione temporale delle quantità presenti nei compartimenti viene descritta mediante sistemi di equazioni differenziali ordinarie [2].

Una volta fissata la struttura del modello, il problema diretto consiste nel determinare l'evoluzione delle concentrazioni conoscendo i coefficienti cinetici e la funzione di ingresso. Nelle applicazioni, tuttavia, sono note le osservazioni sperimentali, mentre alcuni coefficienti cinetici devono essere determinati. Si ottiene così un problema inverso da risolvere.

Prima di procedere alla stima numerica dei parametri è necessario stabilire se essi possano essere determinati univocamente a partire dalle informazioni fornite dall'esperimento. Questa questione conduce allo studio dell'identificabilità a priori, o identificabilità strutturale, che viene effettuato in condizioni ideali assumendo nota la struttura del modello e considerando dati continui e privi di errore [2]. Tale analisi permette di verificare se differenti vettori di parametri possano produrre lo stesso comportamento a partire dagli stessi input e output vettoriali.

Il presente lavoro affronta queste problematiche con riferimento a un modello compartimentale per lo studio della cinetica del radiotracciante ^{18}F-FDOPA mediante PET. Il modello tiene conto, in particolare, della presenza di FDOPA e del suo metabolita 3-*O*-metil-FDOPA (3-OMFD) nel plasma e nel tessuto cerebellare.

Le ipotesi adottate e le relazioni tra i diversi compartimenti vengono formulate sulla base dei modelli e delle osservazioni presenti nella letteratura [5, 8, 10].

A partire dalla descrizione biologica viene costruito un sistema di equazioni differenziali lineari nelle variabili di stato, i cui coefficienti rappresentano i tassi di trasferimento tra i diversi compartimenti. Dopo aver introdotto i fondamenti dell'analisi compartimentale e il problema diretto, viene formulato il modello specifico per la cinetica della FDOPA e ne viene studiata l'identificabilità a priori.

Per l'analisi dell'identificabilità viene utilizzata una formulazione ingresso-uscita del sistema. La trasformata di Laplace permette di ricavare la funzione di trasferimento del modello e di esprimere i parametri osservazionali in funzione dei coefficienti cinetici [2]. Lo studio delle relazioni così ottenute consente di verificare se i parametri incogniti possano essere ricostruiti univocamente a partire dalle osservazioni ideali.

Successivamente il problema viene affrontato dal punto di vista numerico. Poiché i dati PET reali sono disponibili in un numero finito di istanti e sono affetti da errore, la stima dei coefficienti cinetici viene formulata come un problema di minimi quadrati non lineari. Per la sua soluzione vengono considerati il metodo di Gauss-Newton e il metodo di Levenberg-Marquardt, seguendo l'impostazione di Nocedal e Wright [9].

L'obiettivo complessivo della tesi è quindi quello di collegare la descrizione compartimentale del fenomeno biologico allo studio e alla soluzione del corrispondente problema inverso.

La tesi è quindi organizzata nel seguente modo. Nel primo capitolo vengono introdotti i fondamenti dell'analisi compartimentale e il problema diretto per un generico modello lineare. Nel secondo capitolo vengono descritti i principi della PET, la cinetica della ^{18}F-FDOPA e il modello compartimentale considerato. Il terzo capitolo è dedicato al problema inverso e allo studio dell'identificabilità a priori. Nel capitolo successivo il problema di stima viene formulato come problema di minimi quadrati non lineari e vengono analizzati i metodi di Gauss-Newton e Levenberg-Marquardt.

Capitolo 1

Analisi compartimentale

1.1 Fondamenti dell'analisi compartimentale

1.1.1 Analisi compartimentale e cinetica dei traccianti

Lo studio quantitativo di un sistema biologico presenta una difficoltà fondamentale: non è generalmente possibile descrivere e seguire nel tempo il comportamento di ogni singola molecola che partecipa ai processi considerati. L'analisi compartimentale permette di affrontare questa problematica introducendo una rappresentazione semplificata del sistema, nella quale quantità di materiale caratterizzate da un comportamento simile vengono raggruppate in un numero finito di compartimenti [2].

Le connessioni tra i compartimenti rappresentano i processi attraverso i quali il materiale viene trasferito o trasformato all'interno del sistema. La descrizione compartimentale consente quindi di ricondurre la dinamica di un sistema biologico a un insieme di variabili e di parametri cinetici che ne descrivono quantitativamente gli scambi .

Questa impostazione è particolarmente utile negli studi basati sull'impiego di traccianti. Come osservato da Cobelli, Foster e Toffolo, l'obiettivo dell'analisi cinetica consiste nel mettere in relazione la dinamica descritta dal modello con le grandezze effettivamente accessibili nell'esperimento, in modo da utilizzare le osservazioni sperimentali per ottenere informazioni sui parametri cinetici del sistema.

In questo lavoro tale impostazione viene applicata allo studio di dati ottenuti mediante tomografia a emissione di positroni. In un'acquisizione PET dinamica, la distribuzione di un radiotracciante viene osservata nel corso del tempo e le misure ottenute forniscono informazioni sull'evoluzione della concentrazione di attività radioattiva nelle regioni considerate. Il modello compartimentale permette quindi di

formalizzare matematicamente il legame tra le quantità osservabili mediante la PET e i processi di trasporto e trasformazione del tracciante che si vogliono descrivere.

Prima di considerare il modello specifico utilizzato nel nostro lavoro introduciamo quindi i principali concetti dell'analisi compartimentale e una loro formulazione matematica.

1.1.2 Il concetto di compartimento

L'analisi compartimentale consiste nel rappresentare un sistema complesso mediante un numero finito di componenti, dette *compartimenti*, tra le quali possono avvenire scambi di materiale o trasformazioni biochimiche. Le connessioni tra i compartimenti vengono generalmente rappresentate mediante frecce orientate, che indicano la direzione dei trasferimenti.

La suddivisione in compartimenti costituisce una semplificazione della realtà. Un compartimento non deve necessariamente corrispondere a una regione anatomica fisicamente separata dalle altre, ma rappresenta una quantità di materiale caratterizzata da un comportamento sufficientemente uniforme dal punto di vista cinetico.

Definizione 1.1. Un *compartimento* è una quantità di materiale che si comporta come se fosse ben miscelata e cineticamente omogenea.

Definizione 1.2. Un *modello compartimentale* è un modello costituito da un numero finito di compartimenti e da un insieme specificato di connessioni tra essi.

L'ipotesi di compartimento ben miscelato significa che la concentrazione della sostanza considerata può essere assunta uniforme all'interno del compartimento in ogni istante. Di conseguenza, due campioni prelevati contemporaneamente dallo stesso compartimento sono descritti dalla stessa concentrazione, entro il livello di accuratezza richiesto dal modello.

Un compartimento non coincide quindi necessariamente con una singola regione fisica. Quantità di materiale localizzate in regioni differenti possono essere rappresentate mediante un unico compartimento quando, sulla scala temporale dell'esperimento, il loro comportamento può essere considerato indistinguibile.

L'*omogeneità cinetica* esprime invece l'ipotesi che le particelle appartenenti allo stesso compartimento siano soggette alle medesime leggi di trasferimento.

Ogni compartimento è caratterizzato dalla quantità di sostanza presente e dai processi che ne modificano il contenuto. L'evoluzione temporale viene quindi determinata dal bilancio tra i contributi in ingresso e quelli in uscita.

1.1.3 Quantità di sostanza e concentrazione

Indichiamo con $Q_i(t)$ la quantità di sostanza presente nell'i-esimo compartimento e con V_i il volume del compartimento. La concentrazione è definita da

$$C_i(t) = \frac{Q_i(t)}{V_i}. \tag{1.1}$$

Nel seguito assumeremo che i volumi V_i siano costanti nel tempo. Di conseguenza,

$$\dot{C}_i(t) = \frac{1}{V_i}\dot{Q}_i(t). \tag{1.2}$$

Le leggi di trasferimento vengono naturalmente formulate in termini delle quantità presenti nei compartimenti. L'ipotesi di volume costante permette tuttavia di riscrivere tali equazioni direttamente in termini delle concentrazioni. Questa formulazione risulta particolarmente conveniente nel presente lavoro, poiché le grandezze ricavate dalle immagini PET sono espresse in termini di concentrazione di attività radioattiva.

1.1.4 Modelli compartimentali lineari

Nel seguito considereremo modelli compartimentali lineari a coefficienti costanti. Indichiamo con k_{ij} il coefficiente cinetico associato al trasferimento dal compartimento j al compartimento i. Adotteremo quindi la convenzione

$$k_{ij}: \qquad j \longrightarrow i. \tag{1.3}$$

Ipotesi 1.1 (Cinetica di primo ordine)**.** Il flusso di materiale dal compartimento j al compartimento i F_{ij} è proporzionale alla quantità presente nel compartimento di partenza:

$$F_{ij}(t) = k_{ij}Q_j(t), \tag{1.4}$$

dove $k_{ij} \geq 0$.

Il coefficiente k_{ij} rappresenta quindi la frazione della quantità presente nel compartimento j trasferita al compartimento i per unità di tempo e ha dimensione dell'inverso di un tempo. Analogamente, indicheremo con $k_{0i} \geq 0$ un eventuale coefficiente di rimozione dal compartimento i verso l'esterno del sistema. Il corrispondente flusso in uscita è

$$F_{0i}(t) = k_{0i}Q_i(t). \tag{1.5}$$

Ipotesi 1.2 (Coefficienti costanti)**.** I coefficienti cinetici k_{ij} e k_{0i} sono indipendenti dal tempo nell'intervallo temporale considerato.

Il sistema può inoltre ricevere contributi dall'esterno. Indichiamo con $U_i(t)$ il flusso di materiale introdotto nel compartimento i. Nella formulazione in concentrazione porremo

$$u_i(t) = \frac{U_i(t)}{V_i}, \tag{1.6}$$

in modo che $u_i(t)$ abbia le stesse unità di misura di $\dot{C}_i(t)$.

1.1.5 Il modello a un compartimento

Il modello compartimentale più semplice è costituito da un singolo compartimento, caratterizzato dalla concentrazione $C_1(t)$. Supponiamo che il compartimento riceva un ingresso assegnato $u_1(t)$ e che la sostanza possa essere rimossa dal sistema con coefficiente cinetico $k_{01} \geq 0$.

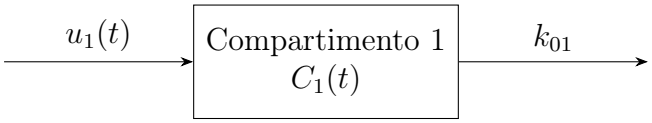

Figura 1.1: Modello compartimentale lineare a un compartimento con ingresso assegnato e processo di rimozione.

Il bilancio della quantità di sostanza è

$$\dot{Q}_1(t) = -k_{01}Q_1(t) + U_1(t). \tag{1.7}$$

Poiché $Q_1(t) = V_1C_1(t)$ e il volume V_1 è costante, dividendo per V_1 si ottiene

$$\dot{C}_1(t) = -k_{01}C_1(t) + u_1(t). \tag{1.8}$$

1.1.6 Un modello a due compartimenti

Consideriamo un sistema costituito da due compartimenti, caratterizzati dalle concentrazioni $C_1(t)$ e $C_2(t)$ e dai volumi costanti V_1 e V_2. Supponiamo che tra i due compartimenti possa avvenire uno scambio bidirezionale. Il coefficiente k_{21} descrive il trasferimento dal compartimento 1 al compartimento 2, mentre k_{12} descrive il trasferimento nella direzione opposta. Indichiamo inoltre con k_{01} e k_{02} i coefficienti associati ai processi di rimozione dal sistema e con $u_1(t)$ e $u_2(t)$ gli eventuali ingressi assegnati.

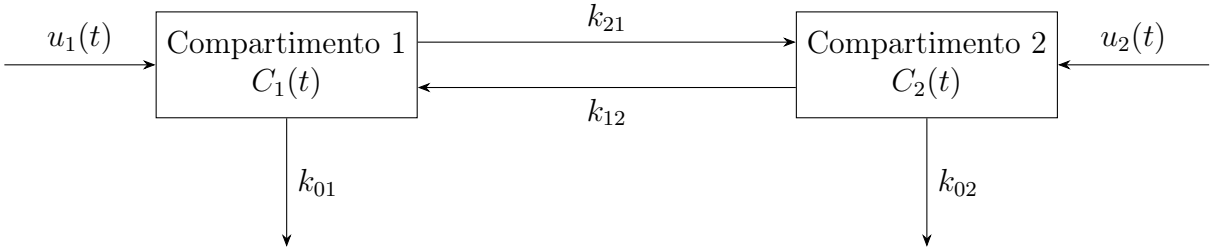

Figura 1.2: Modello compartimentale lineare a due compartimenti con scambio bidirezionale, ingressi assegnati e processi di rimozione.

Il bilancio delle quantità di sostanza è

$$\begin{cases} \dot{Q}_1(t) = -\left(k_{21} + k_{01}\right) Q_1(t) + k_{12} Q_2(t) + U_1(t), \\ \dot{Q}_2(t) = k_{21} Q_1(t) - \left(k_{12} + k_{02}\right) Q_2(t) + U_2(t). \end{cases} \tag{1.9}$$

Utilizzando

$$Q_1(t) = V_1 C_1(t), \qquad Q_2(t) = V_2 C_2(t),$$

e ricordando che i volumi sono costanti, si ottiene

$$\begin{cases} \dot{C}_1(t) = -\left(k_{21} + k_{01}\right) C_1(t) + \dfrac{V_2}{V_1} k_{12} C_2(t) + u_1(t), \\ \dot{C}_2(t) = \dfrac{V_1}{V_2} k_{21} C_1(t) - \left(k_{12} + k_{02}\right) C_2(t) + u_2(t). \end{cases} \tag{1.10}$$

I rapporti tra i volumi compaiono nei termini di trasferimento in ingresso perché una stessa quantità di materiale produce variazioni di concentrazione differenti in compartimenti di volume diverso.

Nel caso particolare $V_1 = V_2$, il sistema si riduce alla forma

$$\begin{cases} \dot{C}_1(t) = -\left(k_{21} + k_{01}\right) C_1(t) + k_{12} C_2(t) + u_1(t), \\ \dot{C}_2(t) = k_{21} C_1(t) - \left(k_{12} + k_{02}\right) C_2(t) + u_2(t). \end{cases} \tag{1.11}$$

1.1.7 Il modello a N compartimenti

Consideriamo ora un sistema costituito da N compartimenti. Per ciascun compartimento $i = 1, \ldots, N$, indichiamo con

$$Q_i(t)$$

la quantità di sostanza presente, con $V_i > 0$ il volume costante e con

$$C_i(t) = \frac{Q_i(t)}{V_i}$$

la corrispondente concentrazione.

Per ogni coppia di compartimenti distinti i e j, il coefficiente $k_{ij} \geq 0$ descrive il trasferimento dal compartimento j al compartimento i. Indichiamo inoltre con $k_{0i} \geq 0$ il coefficiente di rimozione dal compartimento i verso l'esterno del sistema. Il bilancio della quantità presente nel compartimento i è quindi

$$\dot{Q}_i(t) = \sum_{\substack{j=1 \\ j\neq i}}^{N} k_{ij} Q_j(t) - \left(k_{0i} + \sum_{\substack{j=1 \\ j\neq i}}^{N} k_{ji} \right) Q_i(t) + U_i(t). \tag{1.12}$$

La prima sommatoria raccoglie tutti i trasferimenti provenienti dagli altri compartimenti. La seconda raccoglie invece tutti i processi che rimuovono materiale dal compartimento i: i trasferimenti verso gli altri compartimenti e l'eventuale uscita dal sistema. Poiché

$$Q_i(t) = V_i C_i(t),$$

e i volumi sono costanti, dividendo (1.12) per V_i si ottiene

$$\dot{C}_i(t) = \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{V_j}{V_i} k_{ij} C_j(t) - \left(k_{0i} + \sum_{\substack{j=1 \\ j\neq i}}^{N} k_{ji} \right) C_i(t) + u_i(t), \tag{1.13}$$

dove

$$u_i(t) = \frac{U_i(t)}{V_i}. \tag{1.14}$$

Definiamo ora il vettore delle concentrazioni

$$\mathbf{C}(t) = \begin{pmatrix} C_1(t) \\ \vdots \\ C_N(t) \end{pmatrix} \in \mathbb{R}^N \tag{1.15}$$

e il vettore degli ingressi

$$\mathbf{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_N(t) \end{pmatrix}. \tag{1.16}$$

Introduciamo quindi la matrice $K \in \mathbb{R}^{N\times N}$ definita da

$$K_{ij} = \begin{cases} \dfrac{V_j}{V_i} k_{ij}, & i \neq j, \\ -\left(k_{0i} + \displaystyle\sum_{\substack{\ell=1 \\ \ell\neq i}}^{N} k_{\ell i} \right), & i = j. \end{cases} \tag{1.17}$$

È importante distinguere i coefficienti cinetici k_{ij}, definiti a partire dai flussi di quantità, dagli elementi della matrice compartimentale nella formulazione in concentrazione. Infatti, per $i \neq j$, il coefficiente che moltiplica $C_j(t)$ nell'equazione relativa a $C_i(t)$ è

$$K_{ij} = \frac{V_j}{V_i} k_{ij}. \tag{1.18}$$

Pertanto, quando i compartimenti hanno volumi differenti, il coefficiente associato al trasferimento di quantità e il coefficiente che compare nell'equazione della concentrazione del compartimento di arrivo non coincidono in generale.

Nel seguito, qualora un modello venga formulato direttamente in termini di concentrazioni, verrà specificato esplicitamente il significato dei coefficienti cinetici utilizzati nelle corrispondenti equazioni.

Il sistema può allora essere scritto nella forma matriciale

$$\boxed{\dot{\mathbf{C}}(t) = K\mathbf{C}(t) + \mathbf{u}(t)}. \tag{1.19}$$

La struttura della matrice K riflette direttamente la struttura del modello. Gli elementi fuori dalla diagonale sono non negativi,

$$K_{ij} \geq 0, \qquad i \neq j, \tag{1.20}$$

mentre gli elementi diagonali sono non positivi,

$$K_{ii} \leq 0. \tag{1.21}$$

A differenza della formulazione basata direttamente sulle quantità, nella formulazione in concentrazione la somma degli elementi di una colonna della matrice K non ha, in generale, un significato di conservazione, poiché i compartimenti possono avere volumi differenti. Il corrispondente bilancio si ottiene invece pesando ciascuna

concentrazione con il volume del compartimento. In particolare,

$$Q_{\text{tot}}(t) = \sum_{i=1}^{N} V_i C_i(t) \tag{1.22}$$

rappresenta la quantità totale presente nel sistema. Moltiplicando le equazioni (1.13) per i rispettivi volumi e sommando su tutti i compartimenti, i trasferimenti interni si cancellano a coppie e si ottiene

$$\frac{d}{dt} Q_{\text{tot}}(t) = \sum_{i=1}^{N} V_i u_i(t) - \sum_{i=1}^{N} k_{0i} V_i C_i(t). \tag{1.23}$$

La variazione della quantità totale è quindi determinata soltanto dagli ingressi esterni e dai processi di rimozione dal sistema. In particolare, in assenza di ingressi e di uscite, ossia con le condizioni

$$\mathbf{u}(t) = \mathbf{0}, \qquad k_{0i} = 0 \quad \forall i, \tag{1.24}$$

si ha

$$\frac{d}{dt} Q_{\text{tot}}(t) = 0, \tag{1.25}$$

e la quantità totale di sostanza presente nel sistema si conserva.

1.2 Il problema diretto

1.2.1 Formulazione del problema

Una volta specificata la struttura del modello, il problema diretto consiste nel determinare l'evoluzione temporale delle concentrazioni quando siano noti i coefficienti cinetici, la funzione di ingresso e lo stato iniziale.

Consideriamo quindi il problema di Cauchy in notazione matriciale

$$\begin{cases} \dot{\mathbf{C}}(t) = K\mathbf{C}(t) + \mathbf{u}(t), \\ \mathbf{C}(t_0) = \mathbf{C}_0. \end{cases} \tag{1.26}$$

dove K è una matrice costante, $\mathbf{u}(t)$ è una funzione di ingresso assegnata e $\mathbf{C}_0$ rappresenta lo stato iniziale del sistema.

1.2.2 Esistenza e unicità della soluzione

Il sistema (1.26) costituisce un problema di Cauchy lineare.

Proposizione 1.1. Sia $K \in \mathbb{R}^{N\times N}$ costante e sia

$$\mathbf{u} : [t_0, T] \longrightarrow \mathbb{R}^N$$

continua. Per ogni $\mathbf{C}_0 \in \mathbb{R}^N$, il problema (1.26) ammette una e una sola soluzione

$$\mathbf{C} \in C^1([t_0, T]; \mathbb{R}^N).$$

Dimostrazione. Definiamo

$$f(t, \mathbf{C}) = K\mathbf{C} + \mathbf{u}(t). \tag{1.27}$$

Poiché $\mathbf{u}$ è continua, la funzione

$$f : [t_0, T] \times \mathbb{R}^N \longrightarrow \mathbb{R}^N$$

è continua.

Indichiamo con $\|\cdot\|_2$ la norma euclidea in $\mathbb{R}^N$ e con la stessa notazione la corrispondente norma matriciale indotta. Per ogni $\mathbf{C}_1, \mathbf{C}_2 \in \mathbb{R}^N$ si ha

$$\begin{aligned} \|f(t, \mathbf{C}_1) - f(t, \mathbf{C}_2)\|_2 &= \|K(\mathbf{C}_1 - \mathbf{C}_2)\|_2 \\ &\leq \|K\|_2 \|\mathbf{C}_1 - \mathbf{C}_2\|_2 . \end{aligned} \tag{1.28}$$

Poiché f è continua e globalmente lipschitziana rispetto alla variabile di stato, il teorema di esistenza e unicità per i problemi di Cauchy garantisce che, per ogni dato iniziale $\mathbf{C}_0$, esiste un'unica soluzione del sistema su tutto l'intervallo $[t_0, T]$. □

1.2.3 Esponenziale di matrice

Per descrivere la soluzione del sistema introduciamo l'esponenziale di matrice.

Definizione 1.3. Sia $M \in \mathbb{R}^{N\times N}$. L'esponenziale della matrice M è definito da

$$e^M = \sum_{n=0}^{\infty} \frac{M^n}{n!}. \tag{1.29}$$

In particolare,

$$e^{tK} = I + tK + \frac{t^2K^2}{2!} + \frac{t^3K^3}{3!} + \cdots . \tag{1.30}$$

Si ha

$$\frac{d}{dt}e^{tK} = Ke^{tK} = e^{tK}K \tag{1.31}$$

e

$$e^{0K} = I. \tag{1.32}$$

Nel caso omogeneo

$$\dot{\mathbf{C}}(t) = K\mathbf{C}(t), \tag{1.33}$$

la soluzione associata alla condizione iniziale $\mathbf{C}(t_0) = \mathbf{C}_0$ è

$$\mathbf{C}(t) = e^{(t-t_0)K}\mathbf{C}_0. \tag{1.34}$$

1.2.4 Soluzione del problema con ingresso

Mediante il metodo di variazione delle costanti, la soluzione del problema (1.26) è

$$\mathbf{C}(t) = e^{(t-t_0)K}\mathbf{C}_0 + \int_{t_0}^{t} e^{(t-\tau)K}\mathbf{u}(\tau)\, d\tau. \tag{1.35}$$

Il primo termine descrive l'evoluzione dello stato iniziale in assenza di ingresso, mentre il secondo rappresenta la risposta del sistema alla funzione di ingresso.

La formula (1.35) non richiede che la matrice K sia diagonalizzabile.

Nel caso di un esperimento con tracciante si considera generalmente $t_0 = 0$. Se il tracciante non è presente nei compartimenti prima della somministrazione, si assume

$$\mathbf{C}(0) = \mathbf{0}. \tag{1.36}$$

La soluzione diventa quindi

$$\mathbf{C}(t) = \int_{0}^{t} e^{(t-\tau)K}\mathbf{u}(\tau)\, d\tau. \tag{1.37}$$

1.2.5 Non negatività della soluzione

Una proprietà fondamentale dei modelli compartimentali è la non negatività delle soluzioni. Dal punto di vista fisico, infatti, le quantità di sostanza e le corrispondenti concentrazioni non possono assumere valori negativi.

La matrice compartimentale K definita in precedenza soddisfa

$$K_{ij} \geq 0, \qquad i \neq j. \tag{1.38}$$

Definizione 1.4. Una matrice con elementi fuori dalla diagonale non negativi viene detta *matrice di Metzler*.

Proposizione 1.2. Sia K una matrice di Metzler. Se

$$\mathbf{C}(0) \geq \mathbf{0} \tag{1.39}$$

e

$$\mathbf{u}(t) \geq \mathbf{0} \tag{1.40}$$

componente per componente per ogni $t \geq 0$, allora la soluzione del sistema

$$\dot{\mathbf{C}}(t) = K\mathbf{C}(t) + \mathbf{u}(t) \tag{1.41}$$

soddisfa

$$\mathbf{C}(t) \geq \mathbf{0} \tag{1.42}$$

per ogni $t \geq 0$.

Dimostrazione. Poiché K è una matrice di Metzler, è possibile scegliere $\mu \geq 0$ sufficientemente grande affinché

$$A = K + \mu I \tag{1.43}$$

abbia tutti gli elementi non negativi.

Si ha quindi

$$e^{tK} = e^{-\mu t} e^{tA}. \tag{1.44}$$

Poiché A ha elementi non negativi, anche ogni potenza A^n, $n \geq 1$, ha elementi non negativi. Inoltre $A^0 = I$ ha elementi non negativi. Dalla definizione dell'esponenziale di matrice,

$$e^{tA} = \sum_{n=0}^{\infty} \frac{t^n A^n}{n!}, \tag{1.45}$$

e poiché per $t \geq 0$ tutti i coefficienti $t^n/n!$ sono non negativi, segue che e^{tA} ha elementi non negativi.

Poiché $e^{-\mu t} > 0$, anche e^{tK} ha elementi non negativi per ogni $t \geq 0$.

Utilizzando la formula della soluzione,

$$\mathbf{C}(t) = e^{tK}\mathbf{C}(0) + \int_0^t e^{(t-\tau)K}\mathbf{u}(\tau)\,d\tau, \tag{1.46}$$

entrambi i termini al secondo membro sono vettori non negativi. Ne segue che

$$\mathbf{C}(t) \geq \mathbf{0} \tag{1.47}$$

per ogni $t \geq 0$. □

1.3 Il problema inverso e l'identificabilità a priori

1.3.1 Il problema inverso

Nel problema diretto si assumono noti la struttura del modello, i coefficienti cinetici, la funzione di ingresso e lo stato iniziale, e si determina l'evoluzione temporale delle concentrazioni nei diversi compartimenti.

Nel problema inverso, invece, la struttura del modello viene assunta nota, mentre alcuni dei coefficienti cinetici che descrivono gli scambi tra i compartimenti sono incogniti. L'obiettivo consiste nel determinare tali coefficienti a partire dalla funzione di ingresso e dalle quantità osservabili sperimentalmente.

Indichiamo con

$$\mathbf{k} = (k_1, \ldots, k_p)^T \tag{1.48}$$

il vettore dei parametri cinetici incogniti. La matrice compartimentale dipende quindi dai parametri e verrà indicata con

$$K = K(\mathbf{k}). \tag{1.49}$$

Il sistema assume la forma

$$\begin{cases} \dot{\mathbf{C}}(t) = K(\mathbf{k})\mathbf{C}(t) + \mathbf{u}(t), \\ \mathbf{C}(0) = \mathbf{C}_0. \end{cases} \tag{1.50}$$

In generale non tutte le componenti del vettore di stato $\mathbf{C}(t)$ sono direttamente accessibili alla misura. L'esperimento fornisce invece una o più quantità dipendenti dallo stato del sistema, che costituiscono l'uscita del modello. Il problema inverso consiste quindi nel ricostruire il vettore $\mathbf{k}$ a partire dall'ingresso assegnato e dall'uscita osservata.

Prima di affrontare la stima numerica dei parametri è però necessario stabilire se le informazioni contenute nell'esperimento siano, almeno idealmente, sufficienti a determinarli. Seguendo l'impostazione di Cobelli, Foster e Toffolo [2], questo problema viene studiato assumendo nota esattamente la struttura del modello e considerando ingresso e uscita disponibili in modo continuo e privi di errori di misura.

In questo contesto, l'uscita rappresenta il dato ideale prodotto dal modello e il problema dell'identificabilità consiste nello stabilire se due vettori di parametri ammissibili che, a parità di ingresso, producono la stessa uscita debbano necessariamente coincidere. In particolare, dati due vettori

$$\mathbf{k}, \qquad \mathbf{k}', \tag{1.51}$$

si vuole stabilire se l'uguaglianza delle corrispondenti uscite per ogni istante di tempo implichi

$$\mathbf{k} = \mathbf{k}'. \tag{1.52}$$

Lo studio dell'identificabilità precede quindi la stima numerica dei parametri e riguarda l'informazione che, almeno idealmente, può essere ricavata dalla configurazione ingresso-uscita dell'esperimento.

1.3.2 Parametri del modello e parametri osservazionali

Per formulare il problema dell'identificabilità, Cobelli distingue i parametri propri del modello dalle quantità determinabili attraverso l'esperimento ingresso-uscita [2].

I coefficienti

$$\mathbf{k} = (k_1, \ldots, k_p)^T \tag{1.53}$$

sono i parametri del modello. Essi descrivono la dinamica interna del sistema, ma non sono necessariamente osservabili direttamente.

L'esperimento ideale permette invece di determinare un insieme di quantità, dette *parametri osservazionali*. Indichiamo tali parametri con

$$\mathbf{c} = (c_1, \ldots, c_r)^T. \tag{1.54}$$

I parametri osservazionali sono funzioni dei parametri del modello e possiamo

quindi scrivere

$$\begin{cases} c_1 = f_1(\mathbf{k}), \\ c_2 = f_2(\mathbf{k}), \\ \quad \vdots \\ c_r = f_r(\mathbf{k}). \end{cases} \tag{1.55}$$

Il problema dell'identificabilità consiste quindi nello stabilire se l'informazione contenuta nel vettore $\mathbf{c}$ permetta di determinare univocamente il vettore dei parametri $\mathbf{k}$.

1.3.3 Exhaustive summary

L'insieme completo delle relazioni tra i parametri osservazionali e i parametri del modello viene indicato da Cobelli con il termine *exhaustive summary* [2].

Una volta determinato il vettore dei parametri osservazionali $\mathbf{c}$, lo studio dell'identificabilità conduce quindi al sistema algebrico

$$\begin{cases} f_1(\mathbf{k}) = c_1, \\ f_2(\mathbf{k}) = c_2, \\ \quad \vdots \\ f_r(\mathbf{k}) = c_r. \end{cases} \tag{1.56}$$

La classificazione dell'identificabilità dipende dal numero di soluzioni ammissibili di questo sistema. Per quasi ogni scelta dei parametri, il modello è detto *globalmente identificabile* se la soluzione è unica, *localmente identificabile* se esiste un numero finito di soluzioni, maggiore di uno, e *non identificabile* se esistono infinite soluzioni compatibili con gli stessi parametri osservazionali [2].

Nel caso dei modelli compartimentali è inoltre necessario tenere conto dei vincoli imposti dalla natura fisica del sistema. Le soluzioni del sistema algebrico possono infatti appartenere allo spazio complesso o essere reali ma non fisicamente ammissibili. Ai fini dell'identificabilità devono quindi essere considerate le sole soluzioni appartenenti allo spazio dei parametri compartimentali, e in particolare compatibili con i vincoli di realtà e positività dei coefficienti cinetici.

Cobelli introduce inoltre il concetto di *identificabilità su intervallo*: un parametro non identificabile è detto identificabile su intervallo quando è comunque possibile determinarne limiti inferiore e superiore finiti.

La costruzione dell'exhaustive summary può essere effettuata direttamente a partire dall'espressione dell'uscita del modello. Questo procedimento diventa tuttavia rapidamente complesso all'aumentare del numero di compartimenti.

Per i modelli compartimentali lineari, Cobelli considera anche un procedimento basato sulla trasformata di Laplace, noto come *transfer function method*. Il metodo permette di ricavare le relazioni tra parametri osservazionali e parametri del modello senza dover determinare esplicitamente la soluzione temporale completa del sistema differenziale [2].

Questo approccio verrà utilizzato nel seguito per lo studio dell'identificabilità del modello considerato in questo lavoro.

Capitolo 2

PET e FDOPA

2.1 Introduzione alla PET

La tomografia a emissione di positroni, comunemente indicata con l'acronimo PET (*Positron Emission Tomography*), è una tecnica di imaging che consente di studiare in vivo la distribuzione di sostanze marcate con radionuclidi emettitori di positroni.

Per ripercorrere brevemente le principali fasi che hanno portato allo sviluppo della PET, faremo riferimento al lavoro di Wagner [11].

Le basi fisiche della PET risalgono agli anni Trenta. Nel 1930 P. Dirac postulò l'esistenza del positrone, particella con la stessa massa dell'elettrone ma carica positiva, la cui esistenza fu confermata sperimentalmente da C. D. Anderson nel 1932. Quando un positrone attraversa la materia perde progressivamente la propria energia e, incontrando un elettrone, va incontro ad annichilazione. L'energia delle due particelle viene allora convertita in due fotoni, emessi approssimativamente in direzioni opposte. La rivelazione di questi fotoni costituisce il principio fisico alla base della PET.

Un secondo passaggio fondamentale fu lo sviluppo del ciclotrone. Negli anni Trenta Ernest Lawrence e i suoi collaboratori realizzarono strumenti capaci di produrre artificialmente radionuclidi quali carbonio-11, azoto-13, ossigeno-15 e fluoro-18, destinati ad assumere un ruolo centrale nella ricerca biomedica. La possibilità di associare tali radionuclidi a molecole di interesse biologico rese possibile applicare il principio del tracciante allo studio dei processi fisiologici e metabolici.

Nel corso dei decenni successivi, le tecniche di medicina nucleare passarono dalla semplice rilevazione esterna della radioattività alla formazione di immagini della sua distribuzione. Un passaggio decisivo fu l'introduzione della tomografia e degli algoritmi di ricostruzione, che permisero di ottenere sezioni trasversali della distribuzione del tracciante. Nel 1975 Ter-Pogossian, Phelps e Hoffman descrissero uno dei primi

sistemi PET tomografici basati sulla rivelazione in coincidenza e sulla ricostruzione delle immagini.

La PET si distingue quindi dalle tecniche prevalentemente anatomiche per la natura funzionale dell'informazione fornita. Il segnale misurato riflette la distribuzione temporale e spaziale del radiofarmaco all'interno dell'organismo e può essere utilizzato per studiare processi quali il flusso sanguigno e il metabolismo. Lo sviluppo di traccianti come ^{18}F-FDG ha contribuito in modo decisivo alla diffusione della PET, rendendo possibile lo studio quantitativo di processi metabolici regionali, in particolare a livello cerebrale.

2.2 Principi fisici della PET

Per l'introduzione ai principi fisici della tomografia a emissione di positroni si farà principalmente riferimento al manuale *Nuclear Medicine Physics: A Handbook for Teachers and Students*, pubblicato dalla International Atomic Energy Agency (IAEA) [1].

Il principio alla base della PET (e di altre tecniche diagnostiche come la SPECT) è il decadimento radioattivo. Il decadimento radioattivo è un processo spontaneo attraverso il quale un nucleo padre instabile emette una particella o radiazione elettromagnetica e si trasforma in un nucleo figlio più stabile, che può essere a sua volta stabile oppure instabile. Il decadimento radioattivo è accompagnato dall'emissione di particelle energetiche, di fotoni γ , oppure di entrambi. Per descrivere un generico processo di decadimento introduciamo la seguente notazione: indichiamo con $A(t)$ l'attività in funzione del tempo e con λ la costante di decadimento radioattivo, che è un parametro caratteristico di ciascun processo di decadimento.

Definizione 2.1. La costante di decadimento radioattivo λ, moltiplicata per un intervallo di tempo molto più piccolo di $1/\lambda$, rappresenta la probabilità che un particolare atomo di una sostanza radioattiva, contenente un grande numero $N(t)$ di atomi radioattivi identici, decada durante tale intervallo di tempo. Si assume che λ sia indipendente dall'ambiente fisico in cui si trova il singolo atomo. L'unità di misura di λ è s^{-1}

Definizione 2.2. L'attività $A(t)$ di una sostanza radioattiva contenente un grande numero N(t) di atomi radioattivi identici rappresenta il numero totale di decadimenti (disintegrazioni) per unità di tempo ed è definita come il prodotto tra $N(t)$ e λ, ossia:

$$A(t) = \lambda N(t).$$

Da questa relazione si deduce come dimensionalmente si abbia che $|A(t)| = T^{-1}$. L'unità di misura di $A(t)$ nel SI è il becquerel (Bq), definito come $1\,\mathrm{Bq} = 1\,\mathrm{s}^{-1}$.

Definizione 2.3. L'attività specifica a è definita come l'attività A per unità di massa m:

$$a = \frac{A}{m} = \frac{\lambda N}{m}.$$

Nel caso di un radionuclide puro, indicando con M la sua massa molare e con N_A il numero di Avogadro, vale $N = (m/M)N_A$ e quindi

$$a = \frac{\lambda N_A}{M}.$$

L'unità di misura di a nel SI è Bq/kg.

Il decadimento può avvenire in modi diversi. Lo scenario più semplice è quello in cui un nucleo instabile padre P decade con costante di decadimento λ_P in un nucleo figlio F, che può essere a sua volta stabile o instabile. Il tasso di diminuzione del numero di nuclei radioattivi parentali $N_P(t)$ è uguale all'attività $A_P(t)$ al tempo t, definita nella definizione di attività come il prodotto $\lambda N(t)$. Se deriviamo nel tempo, otteniamo la seguente equazione differenziale:

$$\frac{dN_P(t)}{dt} = -A_P(t) = -\lambda_P N_P(t) \tag{2.1}$$

che possiamo riscrivere in forma integrale:

$$\int_{N_P(0)}^{N_P(t)} \frac{dN_P}{N_P} = -\int_0^t \lambda_P \, dt \tag{2.2}$$

con $N_P(0)$ la condizione iniziale che esprime il numero di nuclei radioattivi al tempo $t = 0$.

Dal momento che λ_P è costante, possiamo risolvere l'equazione differenziale in questo modo:

$$\ln\left(\frac{N_P(t)}{N_P(0)}\right) = -\lambda_P t$$

da cui:

$$N_P(t) = N_P(0)e^{-\lambda_P t}$$

Allora possiamo riscrivere l'attività come

$$A_P(t) = \lambda_P N_P(t) = \lambda_P N_P(0)e^{-\lambda_P t} = A_P(0)e^{-\lambda_P t}$$

con $A_P(0)$ l'attività iniziale della sostanza radioattiva.

Definizione 2.4. L'emivita (o tempo di dimezzamento) $(T_{1/2})_P$ di una sostanza radioattiva P è il tempo necessario affinché una data quantità di un elemento radioattivo si riduca alla metà della quantità iniziale $N_P(0)$.

$$N_P\big[t = (T_{1/2})_P\big] = \frac{N_P(0)}{2} = N_P(0)e^{-\lambda_P(T_{1/2})_P}$$

In termini di attività, si può definire come il tempo necessario affinché l'attività della sostanza considerata si riduca a metà della radioattività iniziale, ossia di $A_P(0) = \lambda_P N_P(0)$.

$$A_P\big[t = (T_{1/2})_P\big] = \frac{1}{2}A_P(0) = A_P(0)e^{-\lambda_P(T_{1/2})_P}$$

Si può verificare facilmente come, combinando le due espressioni di sopra, si abbia che:

$$\lambda_P = \frac{\ln 2}{(T_{1/2})_P}$$

2.2.1 Decadimento beta positivo

Vi sono numerosi tipi di decadimento, caratterizzati dal tipo di particella (o onda) emessa durante il procedimento. Per la nostra trattazione risulta particolarmente interessante il decadimento beta positivo (β^+).

In questo procedimento, un nucleo padre P, ricco di protoni, trasforma un protone in un neutrone ed emette un positrone e^+ e un neutrino elettronico ν_e. Con questo decadimento si ottiene che il numero atomico del nucleo figlio diminuisce di uno, ossia $Z_F = Z_P - 1$. Il numero di massa atomica rimane invece costante, ossia $A_F = A_P$. Si può esprimere la generica reazione in questo modo:

$$^A_Z P \rightarrow {}^A_{Z-1}F + e^+ + \nu_e$$

I radionuclidi soggetti a decadimento β^+ sono spesso chiamati emettitori di positroni e vengono utilizzati in medicina per l'imaging funzionale mediante la tecnica PET.

Il tracciante più comunemente utilizzato negli studi PET è il fluorodeossiglucosio (FDG) marcato con ^{18}F, che costituisce un buon esempio di decadimento β^+.

2.2.2 Annichilazione

Quando un nucleo padre va incontro a decadimento β^+, oppure quando un fotone ad alta energia interagisce con il campo elettrico del nucleo o con quello di un elettrone orbitale, può essere prodotto un positrone energetico.

Muovendosi attraverso un mezzo assorbente, il positrone perde la maggior parte della propria energia cinetica a causa delle interazioni coulombiane con gli atomi del mezzo. Queste interazioni determinano una perdita di energia per collisione quando l'interazione avviene con un elettrone orbitale dell'atomo assorbente, e una perdita per irraggiamento quando l'interazione avviene con il nucleo dell'atomo assorbente.

In generale, dopo aver perso tutta la propria energia cinetica attraverso perdite per collisione e per irraggiamento, il positrone subisce una collisione finale con un elettrone orbitale disponibile, a causa della forza attrattiva coulombiana tra il positrone, carico positivamente, e l'elettrone locale, carico negativamente. Questo processo prende il nome di annichilazione del positrone.

Durante l'annichilazione, il positrone e l'elettrone scompaiono e vengono sostituiti da due quanti di annichilazione emessi in direzioni opposte, ciascuno con energia pari a 0.511 MeV

Questo processo rispetta alcune leggi di conservazione:

- conservazione della carica elettrica;
- conservazione della quantità di moto lineare;
- conservazione del momento angolare;
- conservazione dell'energia totale.

Una certa percentuale delle annichilazioni avviene prima che il positrone abbia perso completamente la propria energia cinetica; in questo caso il processo prende il nome di annichilazione in volo.

I due quanti emessi durante un'annichilazione in volo non hanno necessariamente la stessa energia e non vengono necessariamente emessi in direzioni esattamente opposte.

2.2.3 Rivelazione in coincidenza

L'imaging PET si basa sul processo di emissioni di positroni a seguito del decadimento radioattivo. I positroni vengono emessi dal nucleo durante il decadimento radioattivo di alcuni isotopi instabili ricchi di protoni. Questi isotopi raggiungono una configurazione più stabile attraverso un processo di decadimento, che trasforma un protone in un neutrone ed è associato alla creazione di un positrone.

Il positrone è l'antiparticella dell'elettrone: possiede la stessa massa dell'elettrone, ma ha carica positiva. Come nel decadimento β, i positroni vengono emessi dal nucleo con energie differenti. Tali energie presentano uno spettro continuo e un valore massimo caratteristico dell'isotopo padre.

Una volta emesso dal nucleo, il positrone si muove attraverso il materiale circostante e subisce interazioni di scattering, modificando la propria direzione e perdendo energia cinetica. Dopo aver percorso una breve distanza, il positrone si arresta e interagisce con un elettrone presente nella materia circostante. La distanza percorsa prima dell'annichilazione dipende dall'energia del positrone, dal radionuclide padre e dalle caratteristiche del mezzo attraversato, e può quindi variare sensibilmente nei diversi casi.

L'interazione di un positrone e di un elettrone determina l'annichilazione di entrambi e la produzione di due fotoni, ciascuno con energia pari a 511 keV, corrispondente all'energia a riposo delle due particelle originarie.

La quantità di moto complessiva delle due particelle al momento dell'interazione è approssimabile a zero. Questo avviene perché l'energia cinetica del positrone si è dispersa quasi del tutto durante il moto nel tessuto, mentre quella dell'elettrone è trascurabile rispetto all'energia associata ai fotoni emessi. La conservazione della quantità di moto implica che i due fotoni vengano emessi in direzioni quasi esattamente opposte, formando un angolo di circa π radianti.

Queste emissioni caratteristiche, note come radiazione di annichilazione, sempre con energia pari a 511 keV, emesse simultaneamente e in direzioni quasi esattamente opposte, costituiscono il principio fisico alla base della PET e offrono vantaggi specifici rispetto all'imaging a singolo fotone nella definizione della Line of Response (LOR).

La rivelazione in coincidenza consiste nel fatto che coppie di fotoni di annichilazione correlati, ciascuno con energia pari a 511 keV, possono essere associate dal sistema di rivelazione sulla base dei rispettivi tempi di misura. Due fotoni rilevati entro un breve intervallo temporale vengono considerati come prodotti dalla stessa annichilazione positrone-elettrone e viene quindi registrato un evento di coincidenza.

L'intervallo temporale entro il quale due eventi vengono considerati coincidenti è indicato con 2τ ed è un parametro del sistema che, normalmente, non può essere modificato . Per minimizzare le coincidenze casuali, questo intervallo deve essere mantenuto il più breve possibile. Il valore di 2τ dipende dalle caratteristiche dello scanner e del sistema di rivelazione ed è generalmente dell'ordine di alcuni nanosecondi. Finestre temporali più brevi sono rese possibili da materiali rivelatori, come LSO, caratterizzati da tempi di decadimento della scintillazione più rapidi e quindi da una migliore risoluzione temporale.

Una ulteriore riduzione della finestra di coincidenza è però limitata dalle differenze nei tempi di arrivo dei due fotoni. Se un'annichilazione positrone-elettrone avviene al margine del campo di vista trasversale, uno dei due fotoni deve percor-

rere solo una breve distanza, mentre l'altro può dover attraversare quasi l'intero diametro dell'anello di rivelatori. Assumendo un diametro dell'anello pari a 90 cm e una velocità della luce pari a 3×10^8 m/s, si può stimare una differenza massima nei tempi di arrivo dell'ordine di 3 ns.

Quando un fotone impatta su un rivelatore PET, viene generato un impulso elettrico. Un discriminatore a frazione costante produce quindi un impulso logico digitale quando il segnale del rivelatore raggiunge una frazione prefissata dell'altezza massima dell'impulso. Questo impulso logico digitale ha durata τ e viene inviato a un modulo di coincidenza, che stabilisce se esso forma un evento di coincidenza con uno dei segnali provenienti dagli altri rivelatori dell'anello.

Un evento di coincidenza viene identificato quando vi è una sovrapposizione temporale tra impulsi logici provenienti da due rivelatori differenti. In altre parole, una coincidenza viene riconosciuta se due eventi di rivelazione sono registrati entro un intervallo temporale non superiore a τ.

Un fotone rilevato all'istante t può quindi formare una coincidenza con un secondo fotone rilevato nell'intervallo fino a $t + \tau$, oppure con un fotone precedentemente rilevato a partire da $t - \tau$. Per questo motivo, l'intervallo complessivo 2τ viene comunemente definito finestra temporale di coincidenza.

È essenziale che i segnali provenienti da tutti i rivelatori del sistema siano temporalmente allineati in modo coerente, così da garantire che le vere coincidenze vengano effettivamente riconosciute all'interno della finestra temporale. Differenze nelle prestazioni dei vari componenti dei rivelatori o nella lunghezza dei cavi possono introdurre sfasamenti temporali variabili. Per questo motivo vengono eseguite correzioni di allineamento temporale, con lo scopo di caratterizzare e compensare le differenze di temporizzazione tra i diversi rivelatori dell'anello.

2.2.4 Formazione dell'immagine PET

La formazione dell'immagine PET parte dalla rivelazione delle coppie di fotoni emessi durante il processo di annichilazione. Quando due fotoni vengono rilevati in coincidenza da due rivelatori differenti, si assume che l'evento di annichilazione sia avvenuto lungo la linea che congiunge i due rivelatori. Tale linea prende il nome di *Line of Response* (LOR).

Durante un'acquisizione PET viene registrato un elevato numero di eventi di coincidenza e, di conseguenza, un insieme di LOR. Queste costituiscono i dati a partire dai quali viene ricostruita la distribuzione spaziale dell'attività radioattiva all'interno del volume esaminato.

I dati acquisiti vengono generalmente organizzati in matrici bidimensionali chiamate *sinogrammi*. In un sinogramma gli eventi vengono raggruppati in funzione della posizione e dell'orientazione delle corrispondenti LOR. Una possibile rappresentazione è

$$p(x, \theta),$$

dove x identifica la posizione della linea di risposta e θ il suo angolo di orientazione.

Il nome *sinogramma* deriva dal fatto che il contributo di una sorgente puntiforme, osservato al variare dell'angolo di proiezione, descrive nel sinogramma una traiettoria sinusoidale.

A partire dai dati contenuti nei sinogrammi, mediante opportuni algoritmi di ricostruzione tomografica, vengono ricostruite le immagini PET. L'insieme delle immagini ottenute descrive la distribuzione spaziale dell'attività radioattiva nel volume esaminato.

2.3 PET statica e PET dinamica

In un'acquisizione statica viene ricostruita una singola immagine, che descrive la distribuzione dell'attività radioattiva nel volume esaminato in corrispondenza dell'intervallo temporale considerato. L'informazione cumulativa ottenuta è quindi di tipo solo spaziale.

In un'acquisizione dinamica viene invece raccolta una sequenza di immagini nel tempo, che permette di osservare e misurare l'evoluzione della distribuzione del radiofarmaco nel tempo. Ciascuna immagine della sequenza prende il nome di *frame*.

L'intervallo complessivo di acquisizione viene pertanto suddiviso in una successione di intervalli temporali, a ciascuno dei quali corrisponde un frame. Indicando con

$$I_1, I_2, \ldots, I_N$$

le immagini ricostruite, la sequenza ottenuta contiene, oltre all'informazione spaziale, anche l'informazione relativa all'evoluzione temporale dell'attività.

La durata dei frame non deve necessariamente essere costante durante l'intera acquisizione. Gli intervalli temporali possono essere uniformi oppure avere durate differenti, scelte in funzione della velocità con cui varia la distribuzione dell'attività. Queste scelte vengono prese basandosi sulla natura fisica del problema.

Un esempio di acquisizione con intervalli non uniformi è riportato da Huang et al. [5]. In tale studio, una scansione PET dinamica della durata complessiva di 125 minuti viene suddivisa in 6 frame da 0.5 minuti, 4 frame da 3 minuti, 5

frame da 10 minuti e 3 frame da 20 minuti. La durata dei frame aumenta quindi progressivamente nel corso dell'acquisizione.

Le acquisizioni dinamiche consentono quindi di visualizzare e misurare come la distribuzione del radiofarmaco varia nel tempo. Considerando una determinata regione di interesse nelle immagini della sequenza, è possibile ricavare una curva tempo-attività che descrive l'andamento temporale dell'attività misurata in tale regione.

Nel nostro lavoro siamo pertanto interessati alle acquisizioni PET dinamiche, poiché le curve tempo-attività costituiscono i dati sperimentali utilizzati per descrivere la cinetica del radiofarmaco mediante modelli compartimentali.

2.3.1 Regione di interesse

Una *Region of Interest* (ROI), o regione di interesse, è una porzione delimitata di un'immagine o di un volume selezionata per analizzare quantitativamente il segnale proveniente da una specifica regione anatomica o funzionale.

Nel caso di un'acquisizione PET dinamica, una stessa ROI può essere applicata ai diversi frame della sequenza. Per ciascun intervallo temporale è quindi possibile calcolare, ad esempio, l'attività media misurata all'interno della regione considerata.

In questo modo si ottiene una *curva tempo-attività* (*Time Activity Curve*, TAC), che descrive l'evoluzione temporale del segnale PET nella regione selezionata.

Nel presente lavoro le ROI assumono un ruolo particolare nella costruzione dei dati utilizzati dal modello: una ROI cerebellare viene utilizzata per ottenere la curva tempo-attività del cervelletto, mentre regioni poste sulle carotidi permettono di ricavare una curva vascolare utilizzata nella determinazione della funzione di ingresso.

2.4 FDOPA-PET

La [^{18}F]fluoro-3,4-diidrossifenil-L-alanina ([^{18}F]FDOPA) è un analogo radioattivo della L-DOPA, precursore della dopamina, utilizzato negli studi PET cerebrali per analizzare la funzione del sistema dopaminergico.

Una volta raggiunto il tessuto cerebrale, la FDOPA può essere convertita dalla decarbossilasi degli amminoacidi aromatici (AADC) nella corrispondente [^{18}F]fluorodopamina. Tale processo assume particolare rilevanza nelle regioni caratterizzate da un'elevata attività dopaminergica, come lo striato, dove i prodotti radioattivi derivati dalla FDOPA possono essere trattenuti nel tessuto.

L'interpretazione quantitativa delle acquisizioni FDOPA-PET non è tuttavia immediata. La PET misura infatti la concentrazione complessiva di radioattività presente in una determinata regione, ma non consente di distinguere direttamente le diverse specie chimiche contenenti il radionuclide. Il segnale osservato può quindi comprendere contributi provenienti sia dalla FDOPA sia dai suoi metaboliti radioattivi.

Un aspetto particolarmente importante è rappresentato dal metabolismo periferico della FDOPA. Nei protocolli PET cerebrali i soggetti possono essere pretrattati con carbidopa, un inibitore periferico dell'AADC, con lo scopo di ridurre la decarbossilazione della FDOPA al di fuori del sistema nervoso centrale. In queste condizioni assume particolare rilevanza la formazione periferica della 3-O-metil-^{18}F-FDOPA (3-OMFD), prodotta attraverso l'azione della catecol-O-metiltransferasi (COMT).

Con il trascorrere del tempo dall'iniezione, la 3-OMFD può costituire una componente significativa della radioattività plasmatica e può contribuire anche al segnale misurato a livello cerebrale. La presenza di questo metabolita deve pertanto essere tenuta in considerazione nella costruzione di un modello quantitativo della cinetica della FDOPA.

Nel presente lavoro l'attenzione sarà rivolta in particolare al cervelletto. In questa regione il segnale PET verrà descritto mediante un modello che considera separatamente il contributo della FDOPA e quello della 3-OMFD. Tale scelta permette di formulare un modello compartimentale sufficientemente semplice da essere analizzato matematicamente, mantenendo allo stesso tempo le principali componenti radioattive rilevanti per il segnale osservato.

2.5 Ipotesi del modello cerebellare

Nel presente lavoro ci concentriamo sulla cinetica della FDOPA e della 3-OMFD nel cervelletto. La scelta di questa regione consente di adottare una struttura compartimentale più semplice rispetto a quella necessaria per descrivere regioni ad elevata attività dopaminergica, come lo striato.

Sulla base del modello proposto da Huang et al., assumiamo infatti che nel cervelletto la conversione della FDOPA in [^{18}F]fluorodopamina (FDA) possa essere trascurata. Consideriamo quindi nel tessuto cerebellare soltanto due componenti radioattive: la FDOPA e la 3-OMFD.

Entrambe le sostanze possono attraversare la barriera ematoencefalica. Il modello comprende pertanto lo scambio bidirezionale della FDOPA tra plasma e cervelletto e, separatamente, lo scambio bidirezionale della 3-OMFD tra plasma e cervelletto.

Per quanto riguarda il metabolismo periferico, assumiamo che i soggetti siano pretrattati con carbidopa. In queste condizioni consideriamo trascurabile, ai fini del modello, il contributo degli altri metaboliti radioattivi periferici e descriviamo la componente metabolica plasmatica mediante la 3-OMFD.

La formazione della 3-OMFD a partire dalla FDOPA è un processo periferico che può avvenire in diversi tessuti dell'organismo. Nel modello tale processo non viene localizzato fisicamente nel plasma, ma viene rappresentato in maniera aggregata mediante un trasferimento dalla componente plasmatica di FDOPA alla componente plasmatica di 3-OMFD.

Non consideriamo inoltre una conversione diretta della FDOPA in 3-OMFD all'interno del compartimento cerebellare. Di conseguenza, le due specie presenti nel tessuto cerebellare vengono collegate alle corrispondenti componenti plasmatiche esclusivamente mediante processi di trasporto.

Infine, assumiamo che immediatamente prima della somministrazione non sia presente tracciante radioattivo nei compartimenti considerati. Le condizioni iniziali sono quindi nulle. Si assume inoltre che la sostanza somministrata sia FDOPA, mentre la 3-OMFD compare successivamente come conseguenza del metabolismo periferico della FDOPA.

2.6 I compartimenti

Sulla base delle ipotesi introdotte nella sezione precedente, consideriamo quattro componenti, due associate al plasma e due al tessuto cerebellare. Indichiamo con

$$C_p(t), \qquad C_m(t) \tag{2.3}$$

le concentrazioni rispettivamente di FDOPA e 3-OMFD nel plasma, e con

$$C_f(t), \qquad C_o(t) \tag{2.4}$$

le corrispondenti concentrazioni nel cervelletto.

Nel seguito utilizzeremo quindi i pedici

- p: FDOPA nel plasma;
- m: 3-OMFD nel plasma;
- f: FDOPA nel cervelletto;
- o: 3-OMFD nel cervelletto.

Le acquisizioni PET dinamiche forniscono, attraverso opportune regioni di interesse, una curva vascolare ricavata dalle carotidi e una curva tempo-attività cerebellare.

Indichiamo con

$$C_{wb}(t) \tag{2.5}$$

la curva vascolare misurata nelle carotidi. Essa rappresenta la concentrazione di attività radioattiva nel sangue intero (*whole blood*) e non coincide, in generale, con la concentrazione di attività presente nel plasma.

Nel modello compartimentale utilizzeremo invece come funzione di ingresso la concentrazione plasmatica totale

$$y_b(t) = C_p(t) + C_m(t). \tag{2.6}$$

La funzione $y_b(t)$ non è quindi direttamente osservata mediante PET, ma deve essere ricavata dalla curva vascolare $C_{wb}(t)$ attraverso una correzione che tenga conto della distribuzione della radioattività tra plasma ed elementi cellulari del sangue. La costruzione di tale funzione di ingresso verrà discussa nelle sezioni successive.

Per quanto riguarda il cervelletto, indichiamo con $y_c(t)$ la curva tempo-attività ottenuta dalla corrispondente regione di interesse. Trascurando il contributo vascolare presente nella ROI cerebellare, il segnale osservato viene descritto come somma dei contributi della FDOPA e della 3-OMFD:

$$y_c(t) = C_f(t) + C_o(t). \tag{2.7}$$

Le quattro concentrazioni interne al modello possono quindi essere raccolte nel vettore di stato

$$\mathbf{C}(t) = \begin{pmatrix} C_p(t) \\ C_m(t) \\ C_f(t) \\ C_o(t) \end{pmatrix}. \tag{2.8}$$

Le singole componenti di $\mathbf{C}(t)$ non sono direttamente distinguibili dalle misure PET. Le quantità disponibili per il modello sono infatti la funzione di ingresso plasmatica totale $y_b(t)$ e la curva cerebellare $y_c(t)$, entrambe costituite dalla somma dei contributi delle specie radioattive considerate.

2.7 Input function

Una componente essenziale delle ipotesi di modello è la definizione della *input function*, che descrive l'andamento temporale dell'attività radioattiva fornita al tessuto attraverso la circolazione arteriosa.

Nel nostro studio la determinazione della *input function* richiede di affrontare due aspetti principali. Il primo riguarda la distinzione tra l'attività misurata nel sangue intero e quella presente nel plasma, che costituisce la quantità necessaria alla descrizione degli scambi con il tessuto. Il secondo riguarda l'assenza di un campionamento arterioso diretto: non disponiamo infatti di prelievi ematici e di successive analisi che permettano di determinare direttamente le concentrazioni plasmatiche di FDOPA e 3-OMFD. L'informazione vascolare viene invece ricavata dalle immagini PET, utilizzando regioni di interesse localizzate sulle carotidi.

Nel seguito confronteremo quindi il caso in cui l'input function viene determinata mediante campionamento arterioso con quello in cui essa viene stimata indirettamente a partire dalle immagini PET.

2.7.1 Campionamento arterioso

In entrambi gli scenari considerati (prelievo arterioso o PET carotidea) il dato che si ottiene fornisce informazioni sulla concentrazione delle specie radioattive FDOPA e 3-OMFD nel sangue intero arterioso. Quello a cui siamo interessati è però la concentrazione delle sostanze a livello plasmatico. Ai nostri scopi, possiamo vedere il sangue come suddiviso in plasma e in eritrociti, la cui frazione sul totale sanguigno è rappresentata dall'ematocrito.

Nello studio di Huang del 1991 [5] gli autori monitorano la concentrazione nel sangue intero e nel plasma del F^{18} mediante un campionamento arterioso continuo. Le due concentrazioni sono legate in questa maniera:

$$C_{wb}(t) = \rho_p(1 - Hct)C_p(t) + \rho_m C_m(t)$$

con

- $C_{wb}(t)$ la concentrazione totale dei due traccianti nel sangue intero
- $C_p(t)$ la concentrazione del FDOPA nel plasma
- $C_m(t)$ la concentrazione del 3-OMFD nel plasma
- Hct l'ematocrito misurato nel sangue intero periferico

- $\rho_p(1 - Hct)$ il coefficiente di ripartizione nel sangue intero rispetto alla concentrazione plasmatica per l'FDOPA

- ρ_m il coefficiente di ripartizione nel sangue intero rispetto alla concentrazione plasmatica per 3-OMFD

Poiché la FDOPA penetra solo in misura limitata negli eritrociti, essa rimane prevalentemente nella componente plasmatica del sangue. Per questo motivo, nel passaggio dalla concentrazione plasmatica alla concentrazione nel sangue intero compare il fattore $1 - Hct$, che rappresenta approssimativamente la frazione volumetrica occupata dal plasma.

La sola misura di $C_{wb}(t)$ non è tuttavia sufficiente a determinare separatamente $C_p(t)$ e $C_m(t)$. Infatti, ad ogni istante, la relazione precedente costituisce una singola equazione nelle due quantità incognite.

Nel lavoro di Huang [5] questa problematica viene risolta utilizzando campioni ematici e analisi chimiche della frazione di 3-OMFD presente nel plasma. Tali informazioni consentono agli autori di ricostruire separatamente le curve temporali plasmatiche di FDOPA e 3-OMFD, che vengono successivamente utilizzate come due distinte funzioni di ingresso del modello compartimentale.

Questo procedimento richiede tuttavia il prelievo di campioni di sangue durante l'acquisizione PET. Questa pratica risulta particolarmente gravosa per il paziente, che deve subire il prelievo per tutta la durata dell'esame. Nel nostro lavoro si vuole invece studiare la possibilità di ricavare l'informazione relativa all'attività ematica direttamente dalle immagini PET, evitando il campionamento arterioso. Ciò introduce un'ulteriore difficoltà: l'attività misurata all'interno di una regione vascolare non permette di distinguere direttamente il contributo della FDOPA da quello della 3-OMFD.

2.7.2 Stima della input function dalle carotidi

Un'alternativa al campionamento arterioso diretto consiste nello stimare la funzione di ingresso a partire dalle immagini PET dinamiche, sfruttando il segnale radioattivo misurato in particolari strutture vascolari. La curva tempo-attività così ottenuta prende il nome di image-derived input function (IDIF).

Nel caso degli studi PET cerebrali, una possibile strategia consiste nel ricavare la funzione di ingresso direttamente da strutture vascolari comprese nel campo di vista dell'acquisizione. In particolare, la IDIF viene tipicamente estratta dalle arterie carotidi interne, e più specificamente dai sifoni carotidei; con gli scanner più recenti tuttavia anche le arterie carotidi comuni possono costituire una possibile regione di

interesse [3]. Un esempio di questo approccio è fornito da De Francisci et al. nel toolbox EMATA.

In questo lavoro viene considerato un tracciante diverso, l'[^{18}F]FDG. Pur trattandosi di una differenza rilevante, il metodo proposto fornisce alcuni elementi utili per descrivere l'estrazione di una curva arteriosa direttamente dalle immagini PET.

Nel lavoro vengono considerate come possibili regioni di estrazione sia le arterie carotidi interne (ICA), generalmente comprese nel campo di vista delle acquisizioni PET cerebrali, sia le arterie carotidi comuni (CCA), nel caso l'estensione del FOV ne consenta l'osservazione.

Una volta individuata la regione vascolare, dalle immagini acquisite nei diversi istanti è possibile ricavare la corrispondente curva tempo-attività. Nel metodo EMATA la regione delle carotidi viene dapprima segmentata e, al suo interno, vengono selezionati i voxel la cui evoluzione temporale presenta le caratteristiche maggiormente compatibili con un segnale arterioso. La media delle curve associate a tali voxel fornisce una prima stima della IDIF.

EMATA

Il metodo EMATA consiste in una successione di sei moduli, ciascuno dei quali interviene su uno specifico aspetto dell'estrazione e dell'elaborazione della input function, dedotta a partire da immagini. I moduli previsti sono:

1. correzione del movimento
2. segmentazione delle carotidi
3. selezione dei voxel arteriosi
4. denoising della curva ottenuta
5. correzione dello spillover
6. analisi grafica di Patlak.

Non tutti i passaggi sono necessari per ottenere una prima stima della funzione di ingresso: la segmentazione delle carotidi e la successiva selezione dei voxel sono sufficienti per costruire una IDIF grezza, mentre gli altri moduli permettono di correggere o elaborare ulteriormente la curva ottenuta. Gli input fondamentali richiesti dal metodo sono costituiti dalla sequenza di immagini PET dinamiche e dalla corrispondente griglia temporale.

Correzione del movimento

Nel corso di un'acquisizione PET dinamica, eventuali movimenti del paziente possono modificare la corrispondenza tra le strutture vascolari e i voxel dell'immagine. Questo aspetto è particolarmente rilevante nel caso delle arterie carotidi. Un disallineamento tra i diversi frame può infatti alterare la forma della IDIF e propagarsi quindi nella successiva stima dei parametri cinetici. Per limitare questo effetto, EMATA prevede un modulo opzionale di motion correction, nel quale i volumi della sequenza PET vengono riallineati rispetto a un frame di riferimento.

Segmentazione delle carotidi

Uno dei passaggi principali nella stima della IDIF consiste nell'individuazione delle regioni dell'immagine corrispondenti ai vasi da cui dedurre la concentrazione arteriosa. Una particolarità del metodo EMATA è che tale operazione viene eseguita automaticamente, utilizzando esclusivamente le immagini PET, senza richiedere esami aggiuntivi (ad esempio MRI).

La procedura prende in considerazione tre differenti strutture vascolari: le arterie carotidi interne (ICA), le arterie carotidi comuni (CCA) e il seno sagittale superiore (SSS). Le prime due sono i possibili siti di estrazione della IDIF, mentre il seno sagittale superiore viene utilizzato per ottimizzare la segmentazione delle carotidi.

Come primo passaggio viene utilizzato un template PET di riferimento, nel quale sono state disegnate delle regioni contenenti le tre strutture vascolari. Il template viene registrato sull'immagine PET statica del soggetto, ottenuta a partire dai frame tardivi dell'acquisizione dinamica. In questo modo le regioni definite sul template vengono riportate nello spazio individuale del soggetto e forniscono delle aree all'interno delle quali viene successivamente ricercata la struttura vascolare.

L'identificazione dei vasi sfrutta anche l'informazione temporale contenuta nell'acquisizione PET dinamica. Nei primi istanti successivi all'iniezione del radiofarmaco una parte rilevante della sostanza radioattiva è ancora concentrata nel sistema vascolare, mentre l'assorbimento nei tessuti è limitato. Sommando un opportuno numero di frame iniziali è quindi possibile ottenere un'immagine nella quale i vasi risultano maggiormente evidenziati. Tale immagine viene indicata dagli autori come *pseudo-angiografia* PET.

Una difficoltà è rappresentata dalla possibile sovrapposizione spaziale tra le regioni arteriose e alcune strutture venose. In particolare, le regioni contenenti le CCA possono comprendere anche porzioni delle vene giugulari, mentre in prossimità delle ICA possono essere presenti contributi provenienti dai seni venosi. Per limitare ta-

le contaminazione EMATA sfrutta la diversa dinamica temporale del passaggio del bolo radioattivo nel sistema arterioso e in quello venoso.

A questo scopo viene inizialmente segmentato il seno sagittale superiore. Dalle curve tempo-attività dei voxel appartenenti a tale regione vengono determinati i rispettivi tempi di picco e la moda di questi valori, che viene utilizzata come stima del tempo del picco venoso. Poiché il picco venoso segue quello arterioso, questa informazione permette di individuare una finestra temporale maggiormente rappresentativa della componente arteriosa.

Viene quindi costruita una seconda pseudo-angiografia utilizzando soltanto i frame compresi nella finestra temporale così determinata. Sulla pseudo-angiografia arteriosa viene infine applicato, all'interno delle regioni corrispondenti alle ICA o alle CCA, un algoritmo di *vessel enhancement* basato sulla matrice Hessiana. Questo consente di evidenziare le strutture con geometria compatibile con quella dei vasi e di ottenere la maschera finale delle carotidi.

Selezione dei voxel arteriosi

Una volta ottenuta la maschera delle carotidi, non tutti i voxel che vi appartengono vengono utilizzati direttamente per costruire la funzione di ingresso. A causa della limitata risoluzione spaziale delle immagini PET, infatti, alcuni voxel inclusi nella maschera possono essere influenzati dall'attività dei tessuti circostanti. L'utilizzo dell'intera regione segmentata potrebbe quindi determinare una sottostima dell'ampiezza del picco arterioso e, allo stesso tempo, una sovrastima dell'attività nella parte finale della curva.

Per ridurre questo effetto, EMATA introduce una procedura di selezione dei *hot voxels*, finalizzata a individuare, all'interno della maschera carotidea, i voxel la cui evoluzione temporale risulta maggiormente compatibile con quella attesa per un segnale arterioso [3].

Per ciascun voxel viene innanzitutto estratta la corrispondente curva tempo-attività. La forma di ogni curva viene quindi caratterizzata mediante un insieme di parametri, tra cui il tempo e l'ampiezza del picco, il valore medio della parte finale della curva, la pendenza nella fase crescente e in quella finale, l'area sottesa alla curva prima e dopo il picco e la deviazione standard.

Una prima selezione viene effettuata sulla base del tempo del picco. Il tempo di picco arterioso globale viene stimato come la moda dei tempi di picco delle curve considerate; i voxel il cui massimo non risulta sufficientemente sincronizzato con tale valore vengono esclusi dalle elaborazioni successive. In questo modo vengono

eliminati i voxel la cui dinamica temporale risulta poco compatibile con il passaggio del bolo nel compartimento arterioso.

Sui voxel rimanenti viene successivamente applicato un algoritmo di clustering k-means con $K = 2$, utilizzando come variabili le caratteristiche estratte dalle rispettive curve. Tra i due gruppi ottenuti viene selezionato quello il cui centroide presenta la maggiore ampiezza del picco, poiché maggiormente compatibile con un segnale proveniente dall'interno del vaso arterioso.

I voxel appartenenti al gruppo selezionato vengono infine ordinati in base all'ampiezza del rispettivo picco. Tra questi vengono mantenuti i primi N voxel, cioè quelli caratterizzati dai valori di picco più elevati, e la media delle loro curve viene utilizzata per ottenere una prima stima dell'input function, indicata come *raw IDIF*. Il numero N non è fissato in maniera univoca, ma dipende dal sito di estrazione considerato, ICA o CCA, e dalle dimensioni dei voxel dell'immagine PET. Viene inoltre permesso all'utente di modificare tale parametro.

La selezione degli hot voxels permette quindi di utilizzare soltanto la parte della regione segmentata che presenta una dinamica maggiormente rappresentativa del segnale arterioso. Oltre a limitare gli effetti dovuti alla contaminazione da parte dei tessuti circostanti, questo passaggio può compensare in parte eventuali imprecisioni nella precedente segmentazione delle carotidi.

Denoising della IDIF

La IDIF ottenuta dalla selezione dei voxel arteriosi è ancora soggetta al rumore tipico delle acquisizioni PET dinamiche. Questo risulta particolarmente problematico poiché eventuali oscillazioni dovute al rumore tendono a propagarsi alla soluzione del problema inverso. Per questo motivo EMATA introduce una fase di denoising basata su un modello parametrico della concentrazione plasmatica del tracciante. La curva viene descritta distinguendo una fase iniziale precedente all'iniezione, una rapida fase crescente approssimata mediante una funzione lineare e una fase successiva al picco rappresentata tramite una combinazione di tre termini esponenziali. I parametri della parte lineare vengono stimati mediante minimi quadrati lineari, mentre quelli della componente multiesponenziale sono determinati mediante minimi quadrati non lineari pesati. Il risultato è una curva regolarizzata che conserva l'andamento temporale della IDIF riducendo le fluttuazioni associate al rumore dovuto alla misurazione.

Correzione dello spillover

A causa della limitata risoluzione spaziale delle immagini PET, l'attività misurata nei voxel appartenenti alle carotidi può essere influenzata dal contributo dei tessuti circostanti. Questo fenomeno, riconducibile all'effetto di volume parziale, è particolarmente rilevante quando le dimensioni del vaso sono dello stesso ordine della risoluzione dello scanner. In EMATA il segnale misurato nella regione carotidea viene modellato come

$$C_{\text{meas}}(t) = RC\, C_{\text{pl}}(t) + SP\, C_t(t),$$

dove $C_{\text{pl}}(t)$ rappresenta l'attività plasmatica che si vuole ricostruire, $C_t(t)$ l'attività del tessuto circostante, RC il recovery coefficient e SP il coefficiente di spillover. Il metodo stima quindi il contributo del tessuto di background e i coefficienti RC e SP, così da correggere la curva misurata e ottenere una stima più accurata della IDIF.

Analisi grafica di Patlak

L'ultimo modulo di EMATA consente di utilizzare la IDIF ottenuta nei passaggi precedenti per effettuare l'analisi grafica di Patlak e ricavare una stima a livello di voxel del parametro di uptake K_i. Questo metodo è applicabile poiché il sistema presenta un solo compartimento irreversibile, a fronte di uno o più compartimenti reversibili. Gli autori considerano questa condizione come verificata nel contesto del loro problema. L'analisi di Patlak trasforma il problema cinetico in una relazione lineare tra opportune combinazioni dell'attività tissutale e della concentrazione plasmatica: per ogni voxel, i parametri della relazione vengono stimati mediante minimi quadrati lineari pesati. In EMATA questo modulo rappresenta quindi un'applicazione della funzione di ingresso precedentemente ricostruita, permettendo di valutare come la IDIF estratta dalle immagini possa essere utilizzata per ottenere parametri quantitativi della cinetica del tracciante.

Limiti del metodo

Nonostante EMATA permetta di ottenere automaticamente una IDIF dalle immagini PET dinamiche, il metodo presenta alcuni limiti. La qualità della funzione di ingresso dipende innanzitutto dalla corretta segmentazione delle strutture vascolari e dalla risoluzione temporale dell'acquisizione, che deve essere sufficientemente elevata nei primi istanti per descrivere in modo accurato il picco arterioso e distinguerlo da quello venoso. Inoltre, la correzione dell'effetto di volume parziale può introdurre

una certa variabilità, in particolare nel caso delle carotidi comuni, per le quali gli autori osservano che il metodo di correzione dello spillover risulta meno affidabile rispetto alle carotidi interne. Anche piccole differenze nella forma della IDIF rispetto alla funzione arteriosa di riferimento possono propagarsi alla stima dei parametri cinetici, soprattutto nel caso dei microparametri dei modelli compartimentali.

Il modello EMATA ha inoltre una caratteristica che rende la sua applicazione al caso FDOPA non immediata. Per l'FDG gli autori possono assumere trascurabile il contributo dei radiometaboliti e, pertanto, possono utilizzare il segnale vascolare come surrogato della concentrazione plasmatica del tracciante. Nel caso della FDOPA invece la radioattività misurata nelle carotidi può contenere contributi provenienti sia dalla FDOPA sia dai suoi radiometaboliti. La IDIF ottenuta dalle immagini non può quindi essere identificata direttamente con la concentrazione plasmatica della sola FDOPA.

2.7.3 L'approccio di Schiepers

Nel paper di Schiepers ed altri del 2007 [10], gli autori forniscono un interessante spunto su come affrontare la nostra questione. A differenza del lavoro di Huang e altri del 1991 [5], gli autori si trovano a dover stimare la suddivisione del tracciante in FDOPA e 3-OMFD senza avere un campionamento sanguigno. La loro idea è di stimare le concentrazioni con informazioni ottenute con studi precedenti sul metabolismo di queste sostanze.

Considerano trascurabile il contributo degli altri metaboliti periferici e approssimano la frazione di 3-OMFD presente nel sangue mediante la funzione

$$f_m(t) = 0.84\left(1 - e^{-0.028t}\right), \tag{2.9}$$

dove t è espresso in minuti dall'iniezione del tracciante.

Assumendo che l'attività ematica sia costituita dalle sole componenti FDOPA e 3-OMFD, tale frazione permette di scrivere

$$C_m(t) = f_m(t)\, y_b(t), \tag{2.10}$$

e quindi

$$C_p(t) = [1 - f_m(t)]\, y_b(t). \tag{2.11}$$

Gli autori stessi però suggeriscono cautela in un simile approccio: la funzione che descrive la frazione di 3-OMFD sul totale non rappresenta una legge generale della conversione della FDOPA, ma una correzione empirica ottenuta a partire da

dati metabolici di una popolazione di riferimento. Essa costituisce pertanto una possibile modalità per stimare le due componenti della curva ematica nel caso in cui non siano disponibili misure metaboliche individuali.

2.7.4 Uno studio sul metabolismo dell'FDOPA

Un differente approccio per la stima del partizionamento dei metaboliti viene fornito in un paper di Melega et al. [8]. In questo lavoro gli autori approfondiscono il metabolismo cerebrale e periferico dell'FDOPA, con lo scopo di fornire una base biochimica per studiare la cinetica del tracciante. Lo studio comprende esperimenti condotti sia su soggetti umani che su primati.

Per la caratterizzazione del metabolismo periferico nell'uomo vengono considerati sei volontari sani, pretrattati con 200 mg di carbidopa 90 minuti prima della somministrazione della FDOPA. Dopo l'iniezione del tracciante vengono effettuati prelievi dall'arteria radiale all'interno di una finestra temporale di 90 minuti. Il plasma ottenuto viene analizzato mediante cromatografia liquida ad alte prestazioni (HPLC), che permette di separare e identificare le diverse specie radioattive presenti.

L'analisi mostra che, nelle condizioni sperimentali considerate, la carbidopa inibisce sufficientemente la decarbossilazione periferica della FDOPA: nel plasma non vengono infatti rilevati FDA e i suoi successivi metaboliti. Il metabolismo periferico osservabile risulta pertanto limitato alla formazione di 3-OMFD.

Di conseguenza, nel caso di soggetti pretrattati con carbidopa, la radioattività plasmatica può essere descritta considerando principalmente due componenti: $C_p(t)$ e $C_m(t)$. La loro importanza relativa varia nel tempo: immediatamente dopo la somministrazione prevale, come prevedibile, la FDOPA, mentre la frazione di 3-OMFD aumenta progressivamente, fino a rappresentare circa il 60% della radioattività plasmatica totale a 60 minuti dall'iniezione.

La componente animale dello studio permette inoltre di analizzare direttamente la composizione della radioattività nei diversi tessuti cerebrali. A 60 minuti dalla somministrazione della FDOPA viene analizzato, tra gli altri, il cervelletto. In esso, al contrario di altre regioni, l'FDA non viene rilevata e la radioattività è costituita prevalentemente dal 3-OMFD, insieme a una componente residua di FDOPA.

Questi risultati forniscono due indicazioni particolarmente rilevanti per la costruzione del modello considerato nel nostro lavoro. Si conferma che il pretrattamento con carbidopa consente di approssimare il contributo radioattivo ematico mediante le sole componenti FDOPA e 3-OMFD. Inoltre l'assenza di una quantitá significativa di FDA nel cervelletto supporta una descrizione semplificata della regione cerebellare basata sui contributi della FDOPA e della 3-OMFD.

2.7.5 Input function senza campionamento arterioso

Uno studio che ha ipotesi modellistiche simili alle nostre è quello di Zaragori del 2021 [12]. In questo articolo non si ha accesso al campionamento arterioso, la curva di attività ematica viene ottenuta direttamente dalle immagini PET mediante regioni di interesse posizionate sulle carotidi interne.

Dal momento che la misurazione ottenuta contiene entrambe le concentrazioni, gli autori effettuano una correzione metabolica della input function utilizzando informazioni disponibili in letteratura sull'ematocrito e sulle suddivisioni temporali delle diverse specie radioattive presenti nel plasma.

Vengono in particolare introdotte le funzioni $f^p_{\mathrm{DOPA}}(t)$, $f^p_{\mathrm{OMFD}}(t)$ e $f^p_{\mathrm{METS}}(t)$, che descrivono rispettivamente le frazioni plasmatiche di FDOPA, 3-OMFD e degli altri metaboliti radioattivi. Tali funzioni vengono ottenute interpolando valori precedentemente riportati in letteratura, e consentono di ricavare dalla curva ematica una stima della input function plasmatica della FDOPA.

Questo approccio consente quindi di evitare il campionamento arterioso, ma introduce una dipendenza da informazioni metaboliche ottenute su popolazioni di riferimento. Le frazioni utilizzate non sono infatti determinate individualmente per ciascun paziente, e questo viene indicato dagli autori come una possibile fonte di errore nella descrizione dell'input function.

Capitolo 3

Il problema inverso

3.1 Formulazione del problema

Nel primo capitolo abbiamo introdotto il problema inverso per un generico modello compartimentale e il concetto di identificabilità a priori. Applichiamo ora tali strumenti al modello cerebellare della FDOPA e della 3-OMFD descritto nel capitolo precedente.

Nel nostro caso le quantità disponibili sono la funzione di ingresso plasmatica totale

$$y_b(t) = C_p(t) + C_m(t) \tag{3.1}$$

e la curva tempo-attività cerebellare

$$y_c(t) = C_f(t) + C_o(t). \tag{3.2}$$

Come discusso nel capitolo precedente, nel problema inverso $y_b(t)$ e $y_c(t)$ vengono considerate note, mentre le singole concentrazioni

$$C_p(t), \qquad C_m(t), \qquad C_f(t), \qquad C_o(t) \tag{3.3}$$

non sono osservabili separatamente.

I coefficienti cinetici incogniti sono rappresentati dal vettore

$$\mathbf{k} = \begin{pmatrix} k_{mp} \\ k_{fp} \\ k_{pf} \\ k_{om} \\ k_{mo} \end{pmatrix}. \tag{3.4}$$

Il problema inverso consiste nello stabilire se il comportamento ingresso-uscita descritto dalla coppia

$$(y_b(t), y_c(t)) \tag{3.5}$$

sia sufficiente a determinare univocamente il vettore $\mathbf{k}$.

In accordo con la formulazione dell'identificabilità a priori introdotta nel primo capitolo, considereremo quindi due vettori di parametri ammissibili

$$\mathbf{k}, \qquad \mathbf{k}', \tag{3.6}$$

e studieremo se, a parità della funzione di ingresso $y_b(t)$, l'uguaglianza delle corrispondenti uscite

$$y_c(t; \mathbf{k}) = y_c(t; \mathbf{k}') \qquad \text{per ogni } t \tag{3.7}$$

implichi necessariamente

$$\mathbf{k} = \mathbf{k}'. \tag{3.8}$$

La successiva analisi verrà svolta dapprima sul modello generale, considerando i cinque coefficienti cinetici indipendenti. In seguito verranno introdotte ulteriori relazioni motivate fisiologicamente, ottenendo una parametrizzazione ridotta del modello.

3.2 Il modello matematico

3.2.1 Il problema in notazione matriciale

Abbiamo visto come il modello catenario possa essere rappresentato mediante un sistema di equazioni differenziali ordinarie con variabile temporale. Nel caso del nostro problema cerebellare, il metabolismo dell'FDOPA e del 3-OMFD può essere rappresentato con questo sistema:

$$\begin{cases} \dot{C}_m(t) = -\left(k_{mp} + k_{om}\right) C_m(t) + k_{mo} C_o(t) + k_{mp} y_b(t), \\ \dot{C}_f(t) = -k_{fp} C_m(t) - k_{pf} C_f(t) + k_{fp} y_b(t), \\ \dot{C}_o(t) = k_{om} C_m(t) - k_{mo} C_o(t). \end{cases} \tag{3.9}$$

con le condizioni iniziali

$$\begin{cases} C_m(0) = 0, \\ C_f(0) = 0, \\ C_o(0) = 0. \end{cases} \tag{3.10}$$

che rappresentano l'assenza di tracciante nel corpo del paziente prima dell'iniezione. I dati osservati sono

$$y_b(t) = C_p(t) + C_m(t), \qquad y_c(t) = C_f(t) + C_o(t). \tag{3.11}$$

La curva cerebellare $y_c(t)$ viene ricavata dalle immagini PET mediante una ROI sul cervelletto. La funzione di ingresso $y_b(t)$ è invece la concentrazione plasmatica totale ottenuta, dopo opportuni accorgimenti, a partire dalla curva vascolare $C_{wb}(t)$ misurata mediante una ROI sulle carotidi.

Nel problema inverso entrambe le funzioni $y_b(t)$ e $y_c(t)$ vengono considerate note, mentre le loro componenti $C_p(t)$, $C_m(t)$, $C_f(t)$ e $C_o(t)$ non sono osservabili separatamente.

Utilizziamo questo fatto per legare due concentrazioni ignote, stabilendo la seguente relazione:

$$C_p(t) = y_b(t) - C_m(t). \tag{3.12}$$

Vogliamo adesso riscrivere il sistema di equazioni differenziali in notazione matriciale. Otteniamo quindi

$$\begin{cases} \dot{\mathbf{C}}(t) = M\mathbf{C}(t) + B\, y_b(t), \\ \mathbf{C}(0) = 0. \end{cases} \tag{3.13}$$

con

$$\mathbf{C}(t) = \begin{pmatrix} C_m(t) \\ C_f(t) \\ C_o(t) \end{pmatrix} \tag{3.14}$$

$$M = \begin{pmatrix} -(k_{mp} + k_{om}) & 0 & k_{mo} \\ -k_{fp} & -k_{pf} & 0 \\ k_{om} & 0 & -k_{mo} \end{pmatrix} \qquad B = \begin{pmatrix} k_{mp} \\ k_{fp} \\ 0 \end{pmatrix} \tag{3.15}$$

$$\mathbf{C}(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{3.16}$$

$$y_c(t) = \alpha^T \mathbf{C}(t), \qquad \alpha = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \tag{3.17}$$

Ricordiamo inoltre il vettore dei coefficienti

$$\mathbf{k} = \begin{pmatrix} k_{mp} \\ k_{fp} \\ k_{pf} \\ k_{om} \\ k_{mo} \end{pmatrix}. \tag{3.18}$$

che è l'incognita del problema inverso. Come visto nel primo capitolo, la soluzione del problema diretto è

$$\mathbf{C}(t) = \int_0^t e^{(t-\tau)M(\mathbf{k})} B(\mathbf{k}) y_b(\tau)\, d\tau. \tag{3.19}$$

Dalla quantità osservata nel cervelletto si ottiene che

$$y_c(t) = \alpha^T \int_0^t e^{(t-\tau)M(\mathbf{k})} B(\mathbf{k}) y_b(\tau)\, d\tau. \tag{3.20}$$

Il problema inverso consiste pertanto nel determinare il vettore dei parametri a partire dalle quantitá note $y_b(t)$ e $y_c(t)$.

3.2.2 La trasformata di Laplace

Definizione e proprietà

Per procedere nella nostra dimostrazione dobbiamo prima introdurre un noto strumento di analisi matematica, la trasformata di Laplace

Definizione 3.1. Una funzione $f : \mathbb{R}^+ \to \mathbb{R}$ si dice di ordine esponenziale se esistono $c > 0$, $\beta \in \mathbb{R}$ e $t_0 \geq 0$ tali che

$$|f(t)| \leq ce^{\beta t} \quad \forall\, t \geq t_0.$$

Definizione 3.2 (Trasformata di Laplace)**.** Sia

$$f : \mathbb{R}_+ \longrightarrow \mathcal{B}$$

una funzione continua a valori in uno spazio di Banach $\mathcal{B}$. Si definisce *trasformata di Laplace* di f la funzione

$$\widetilde{f}(s) = \mathcal{L}\{f\}(s) = \int_0^{+\infty} e^{-st} f(t)\, dt, \qquad s \in \mathbb{C},$$

per tutti i valori di s per i quali l'integrale converge.

Nel caso di una funzione vettoriale

$$\mathbf{C}(t) = \begin{pmatrix} C_1(t) \\ \vdots \\ C_n(t) \end{pmatrix},$$

la trasformata di Laplace è definita componente per componente:

$$\widetilde{\mathbf{C}}(s) = \mathcal{L}\{\mathbf{C}\}(s) = \begin{pmatrix} \widetilde{C}_1(s) \\ \vdots \\ \widetilde{C}_n(s) \end{pmatrix}.$$

Enunciamo alcune proprietà utili di questa trasformata:

1. **Linearità**: date $f, g \in C^0(\mathbb{R}^+, \mathbb{R})$ di ordine esponenziale e $a, b \in \mathbb{R}$, allora

$$\mathcal{L}\{af + bg\}(s) = a\mathcal{L}\{f\}(s) + b\mathcal{L}\{g\}(s).$$

2. **Derivazione**: data $f \in C^1(\mathbb{R}^+, \mathbb{R})$ di ordine esponenziale, con f' di ordine esponenziale, allora

$$\mathcal{L}\{f'(t)\}(s) = s\mathcal{L}\{\{\}(s) - f(0).$$

3. **Convoluzione**: date $f, g \in C^0(\mathbb{R}^+, \mathbb{R})$ di ordine esponenziale, allora

$$\mathcal{L}\{(f * g)(t)\}(s) = \widetilde{f}(s)\widetilde{g}(s),$$

dove

$$(f * g)(t) = \int_0^t f(\tau) g(t - \tau)\, d\tau.$$

4. **Spostamento in frequenza**: data $f \in C^0(\mathbb{R}^+, \mathbb{R})$ di ordine esponenziale e $a \in \mathbb{R}$, allora

$$\mathcal{L}\{e^{at} f(t)\}(s) = \widetilde{f}(s - a).$$

5. **Scalamento temporale**: data $f \in C^0(\mathbb{R}^+, \mathbb{R})$ di ordine esponenziale e $a > 0$, allora
$$\mathcal{L}\{f(at)\}(s) = \frac{1}{a}\widetilde{f}\left(\frac{s}{a}\right).$$

Per poter applicare la trasformata di Laplace al sistema assumiamo che la funzione di ingresso y_b sia continua e di ordine esponenziale.

Tale ipotesi è sufficiente a garantire l'esistenza delle trasformate delle altre funzioni che compaiono nel modello. Infatti, essendo M una matrice costante, dalla definizione dell'esponenziale di matrice si ha

$$e^{tM} = \sum_{n=0}^{\infty} \frac{t^n M^n}{n!}. \tag{3.21}$$

Considerando una norma matriciale submoltiplicativa (come la norma 1,2 o infinito),

$$\left\|e^{tM}\right\| \leq \sum_{n=0}^{\infty} \frac{t^n \|M^n\|}{n!} \tag{3.22}$$

$$\leq \sum_{n=0}^{\infty} \frac{t^n \|M\|^n}{n!} = e^{t\|M\|}, \qquad t \geq 0. \tag{3.23}$$

Pertanto e^{tM} cresce al più esponenzialmente.

Poiché la soluzione del problema diretto è

$$\mathbf{C}(t) = \int_0^t e^{(t-\tau)M} B\, y_b(\tau)\, d\tau, \tag{3.24}$$

se esistono $c > 0$ e $\beta \in \mathbb{R}$ tali che

$$|y_b(t)| \leq ce^{\beta t}, \tag{3.25}$$

si ha

$$\|\mathbf{C}(t)\| \leq \int_0^t \left\|e^{(t-\tau)M}\right\| \|B\| |y_b(\tau)|\, d\tau \tag{3.26}$$

$$\leq Ac\|B\| \int_0^t e^{\omega(t-\tau)} e^{\beta\tau}\, d\tau. \tag{3.27}$$

Ne segue che anche $\mathbf{C}(t)$ è di ordine esponenziale. Di conseguenza, essendo

$$y_c(t) = \alpha^T \mathbf{C}(t), \tag{3.28}$$

anche $y_c(t)$ è di ordine esponenziale.

Infine, dalla relazione

$$\dot{\mathbf{C}}(t) = M\mathbf{C}(t) + B\, y_b(t) \tag{3.29}$$

segue che anche $\dot{\mathbf{C}}(t)$ è di ordine esponenziale.

Supponendo quindi che la funzione di ingresso $y_b(t)$ sia di ordine esponenziale, anche la soluzione $\mathbf{C}(t)$ del sistema lineare, la sua derivata $\dot{\mathbf{C}}(t)$ e l'uscita $y_c(t) = \alpha^T \mathbf{C}(t)$ risultano di ordine esponenziale. Esiste pertanto un valore reale σ_0 tale che le trasformate di Laplace di y_b, $\mathbf{C}$, $\dot{\mathbf{C}}$ e y_c siano definite per

$$\mathrm{Re}(s) > \sigma_0.$$

Utilizzo della trasformata di Laplace

Andiamo adesso ad applicare la trasformata di Laplace a tutti i termini del problema, e otteniamo la seguente formulazione:

$$\widetilde{\mathbf{C}}(s) = \mathcal{L}\{\mathbf{C}(t)\}(s), \qquad \widetilde{y_b}(s) = \mathcal{L}\{y_b(t)\}(s), \qquad \widetilde{y_c}(s) = \mathcal{L}\{y_c(t)\}(s). \tag{3.30}$$

Il sistema diventa pertanto, trasformando e usando la seconda proprietà,

$$s\widetilde{\mathbf{C}}(s) = M\widetilde{\mathbf{C}}(s) + B\,\widetilde{y_b}(s). \tag{3.31}$$

Otteniamo raccogliendo

$$(sI - M)\widetilde{\mathbf{C}}(s) = B\,\widetilde{y_b}(s), \tag{3.32}$$

Consideriamo la matrice

$$sI - M = \begin{pmatrix} s + k_{mp} + k_{om} & 0 & -k_{mo} \\ k_{fp} & s + k_{pf} & 0 \\ -k_{om} & 0 & s + k_{mo} \end{pmatrix}. \tag{3.33}$$

Invertibilità di $sI - M$

Per ricavare la relazione ingresso-uscita nel dominio di Laplace è necessario stabilire per quali valori di s la matrice $sI - M$ sia invertibile. Nel modello generale si ha

$$sI - M = \begin{pmatrix} s + k_{mp} + k_{om} & 0 & -k_{mo} \\ k_{fp} & s + k_{pf} & 0 \\ -k_{om} & 0 & s + k_{mo} \end{pmatrix}. \tag{3.34}$$

Sviluppando il determinante rispetto alla seconda colonna si ottiene

$$\det(sI - M) = (s + k_{pf})\left[(s + k_{mp} + k_{om})(s + k_{mo}) - k_{om}k_{mo}\right] \tag{3.35}$$

$$= (s + k_{pf})\left[s^2 + (k_{mp} + k_{om} + k_{mo})s + k_{mp}k_{mo}\right]. \tag{3.36}$$

Il primo fattore si annulla per

$$s_1 = -k_{pf}, \tag{3.37}$$

che è reale e negativo, poiché $k_{pf} > 0$.

Consideriamo il fattore quadratico

$$s^2 + as + b, \tag{3.38}$$

dove

$$a = k_{mp} + k_{om} + k_{mo} > 0, \qquad b = k_{mp}k_{mo} > 0.$$

Il discriminante soddisfa

$$a^2 - 4b = (k_{mp} - k_{mo})^2 \tag{3.39}$$

$$+ 2k_{om}(k_{mp} + k_{mo}) + k_{om}^2 > 0. \tag{3.40}$$

Le due radici sono pertanto reali e distinte. Poiché la loro somma è $-a < 0$ e il loro prodotto è $b > 0$, entrambe sono strettamente negative.

Anche la radice $s = -k_{pf}$ è strettamente negativa. Di conseguenza,

$$\det(sI - M) \neq 0 \qquad \text{per } \mathrm{Re}(s) \geq 0. \tag{3.41}$$

3.2.3 Funzioni razionali e irriducibilità

Per dimostrare l'unicità andiamo prima a considerare un risultato algebrico.

Definizione 3.3 (Funzione razionale irriducibile)**.** Siano $Q, P \in \mathbb{R}[s]$, con $P \not\equiv 0$. La funzione razionale

$$f(s) = \frac{Q(s)}{P(s)} \tag{3.42}$$

si dice *irriducibile* se i polinomi Q e P non hanno radici comuni, ossia se non esiste $s_0 \in \mathbb{C}$ tale che

$$Q(s_0) = 0 \qquad \text{e} \qquad P(s_0) = 0. \tag{3.43}$$

Teorema 1 (Uguaglianza di funzioni razionali irriducibili)**.** *Siano*

$$f_1(s) = \frac{Q_1(s)}{P_1(s)}, \qquad f_2(s) = \frac{Q_2(s)}{P_2(s)}$$

due funzioni razionali irriducibili, con P_1 e P_2 monici e dello stesso grado. Se

$$f_1(s) = f_2(s),$$

allora

$$P_1(s) = P_2(s) \qquad e \qquad Q_1(s) = Q_2(s).$$

Dimostrazione. Dall'uguaglianza delle due funzioni razionali segue

$$Q_1(s)P_2(s) = Q_2(s)P_1(s).$$

Poiché P_1 divide Q_1P_2 e Q_1 e P_1 sono coprimi, per il lemma di Euclide segue che

$$P_1 \mid P_2.$$

Analogamente, poiché P_2 divide Q_2P_1 e Q_2 e P_2 sono coprimi, si ottiene

$$P_2 \mid P_1.$$

Pertanto P_1 e P_2 differiscono al più per una costante moltiplicativa non nulla. Essendo entrambi monici, tale costante è uguale a 1, e quindi

$$P_1(s) = P_2(s).$$

Sostituendo nella precedente uguaglianza si ottiene infine

$$Q_1(s) = Q_2(s).$$

□

Per una trattazione generale dell'unicità della rappresentazione coprima di una funzione razionale si veda, ad esempio, Fuhrmann [4].

3.2.4 Unicità dei coefficienti

Lo scopo di questa sezione è studiare l'identificabilità strutturale del modello nella sua forma generale, senza imporre ulteriori relazioni tra i coefficienti cinetici. In

particolare, vogliamo stabilire se due diversi vettori di parametri possano produrre, a partire dallo stesso ingresso $y_b(t)$, la medesima uscita osservata $y_c(t)$.

Consideriamo quindi

$$\mathbf{k} = \begin{pmatrix} k_{mp} \\ k_{fp} \\ k_{pf} \\ k_{om} \\ k_{mo} \end{pmatrix}, \qquad \mathbf{k}' = \begin{pmatrix} k'_{mp} \\ k'_{fp} \\ k'_{pf} \\ k'_{om} \\ k'_{mo} \end{pmatrix}, \tag{3.44}$$

e supponiamo che, a partire dallo stesso ingresso $y_b(t)$, essi producano la medesima uscita $y_c(t)$. Il problema di unicità consiste nel determinare sotto quali condizioni tale uguaglianza implichi

$$\mathbf{k} = \mathbf{k}'. \tag{3.45}$$

In questa prima fase ricaviamo la funzione di trasferimento del modello generale e determiniamo le condizioni sotto le quali essa possiede una rappresentazione irriducibile. Tale risultato sarà successivamente utilizzato per il confronto tra i due vettori di parametri.

Funzione di trasferimento del modello generale

Per studiare il problema di unicità trasformiamo la relazione ingresso-uscita nel dominio di Laplace. In questo modo il confronto tra due vettori di parametri viene ricondotto al confronto delle corrispondenti funzioni di trasferimento.

Indicando con M, B e con M', B' le matrici associate rispettivamente a $\mathbf{k}$ e $\mathbf{k}'$, si ha

$$\widetilde{y}_c(s) = \alpha^T (sI - M)^{-1} B \, \widetilde{y}_b(s) \tag{3.46}$$

e

$$\widetilde{y}_c(s) = \alpha^T (sI - M')^{-1} B' \, \widetilde{y}_b(s). \tag{3.47}$$

Supponiamo che l'ingresso y_b non sia identicamente nullo. Dalla uguaglianza delle uscite segue

$$\left[\alpha^T (sI - M)^{-1} B - \alpha^T (sI - M')^{-1} B'\right] \widetilde{y}_b(s) = 0. \tag{3.48}$$

La trasformata $\widetilde{y}_b$ è analitica nel proprio semipiano di convergenza e, poiché y_b non è identicamente nulla, $\widetilde{y}_b$ non è identicamente nulla. I suoi eventuali zeri sono quindi isolati. Esiste pertanto un insieme con punti di accumulazione nel semipiano

di convergenza per il quale

$$\widetilde{y}_b(s) \neq 0.$$

Per tali valori di s possiamo dividere per $\widetilde{y}_b(s)$ e ottenere

$$\alpha^T(sI - M)^{-1}B = \alpha^T(sI - M')^{-1}B'. \tag{3.49}$$

I due membri sono funzioni razionali di s. Poiché coincidono su un insieme con un punto di accumulazione appartenente a una regione in cui sono entrambe definite, essi coincidono identicamente come funzioni razionali.

Poniamo

$$\alpha^T(sI - M)^{-1}B = \frac{Q(s)}{P(s)}, \qquad \alpha^T(sI - M')^{-1}B' = \frac{Q'(s)}{P'(s)}. \tag{3.50}$$

Si ha pertanto

$$\frac{Q(s)}{P(s)} = \frac{Q'(s)}{P'(s)}. \tag{3.51}$$

Determiniamo ora esplicitamente i polinomi P e Q associati al vettore $\mathbf{k}$. Nel modello generale

$$M = \begin{pmatrix} -(k_{mp} + k_{om}) & 0 & k_{mo} \\ -k_{fp} & -k_{pf} & 0 \\ k_{om} & 0 & -k_{mo} \end{pmatrix}, \tag{3.52}$$

per cui

$$sI - M = \begin{pmatrix} s + k_{mp} + k_{om} & 0 & -k_{mo} \\ k_{fp} & s + k_{pf} & 0 \\ -k_{om} & 0 & s + k_{mo} \end{pmatrix}. \tag{3.53}$$

Ponendo

$$P(s) = \det(sI - M), \tag{3.54}$$

e sviluppando il determinante rispetto alla seconda colonna, si ottiene

$$P(s) = (s + k_{pf})\left[(s + k_{mp} + k_{om})(s + k_{mo}) - k_{om}k_{mo}\right] \tag{3.55}$$

$$= (s + k_{pf})\left[s^2 + (k_{mp} + k_{om} + k_{mo})s + k_{mp}k_{mo}\right]. \tag{3.56}$$

Introduciamo quindi

$$R(s) = s^2 + (k_{mp} + k_{om} + k_{mo})s + k_{mp}k_{mo}, \tag{3.57}$$

e quindi

$$P(s) = (s + k_{pf})R(s). \tag{3.58}$$

La forma sviluppata di P è

$$P(s) = s^3 + (k_{mp} + k_{om} + k_{mo} + k_{pf})s^2 \tag{3.59}$$

$$+ \left[k_{mp}k_{mo} + k_{pf}(k_{mp} + k_{om} + k_{mo})\right] s + k_{mp}k_{mo}k_{pf}. \tag{3.60}$$

Per determinare il numeratore consideriamo il sistema trasformato

$$(sI - M)\widetilde{\mathbf{C}}(s) = B\,\widetilde{y}_b(s), \tag{3.61}$$

ossia

$$\begin{cases} (s + k_{mp} + k_{om})\widetilde{C}_m - k_{mo}\widetilde{C}_o = k_{mp}\widetilde{y}_b, \\ k_{fp}\widetilde{C}_m + (s + k_{pf})\widetilde{C}_f = k_{fp}\widetilde{y}_b, \\ -k_{om}\widetilde{C}_m + (s + k_{mo})\widetilde{C}_o = 0. \end{cases} \tag{3.62}$$

Dalla terza equazione segue

$$\widetilde{C}_o = \frac{k_{om}}{s + k_{mo}}\widetilde{C}_m. \tag{3.63}$$

Sostituendo nella prima equazione e utilizzando la definizione di $R(s)$ si ricava

$$\widetilde{C}_m = \frac{k_{mp}(s + k_{mo})}{R(s)}\widetilde{y}_b \tag{3.64}$$

e quindi

$$\widetilde{C}_o = \frac{k_{mp}k_{om}}{R(s)}\widetilde{y}_b. \tag{3.65}$$

Dalla seconda equazione si ottiene invece

$$\widetilde{C}_f = \frac{k_{fp}}{s + k_{pf}}\left(\widetilde{y}_b - \widetilde{C}_m\right) \tag{3.66}$$

$$= \frac{k_{fp}s(s + k_{om} + k_{mo})}{(s + k_{pf})R(s)}\widetilde{y}_b. \tag{3.67}$$

Poiché

$$\widetilde{y}_c = \widetilde{C}_f + \widetilde{C}_o, \tag{3.68}$$

utilizzando (3.58) si ottiene

$$\widetilde{y}_c(s) = \frac{k_{fp}s(s + k_{om} + k_{mo}) + k_{mp}k_{om}(s + k_{pf})}{P(s)}\widetilde{y}_b(s). \tag{3.69}$$

Il numeratore della funzione di trasferimento è pertanto

$$Q(s) = k_{fp}s(s + k_{om} + k_{mo}) + k_{mp}k_{om}(s + k_{pf}). \tag{3.70}$$

Sviluppando,

$$Q(s) = k_{fp}s^2 + [k_{fp}(k_{om} + k_{mo}) + k_{mp}k_{om}]\, s + k_{mp}k_{om}k_{pf}. \tag{3.71}$$

La funzione di trasferimento del modello generale è quindi

$$\frac{\widetilde{y}_c(s)}{\widetilde{y}_b(s)} = \frac{k_{fp}s^2 + [k_{fp}(k_{om} + k_{mo}) + k_{mp}k_{om}]\, s + k_{mp}k_{om}k_{pf}}{(s + k_{pf})\,[s^2 + (k_{mp} + k_{om} + k_{mo})s + k_{mp}k_{mo}]}. \tag{3.72}$$

Abbiamo quindi ricondotto l'ipotesi di uguaglianza delle uscite all'uguaglianza delle corrispondenti funzioni di trasferimento. Per poter applicare il teorema sulle funzioni razionali all'unicità del vettore dei coefficienti occorre studiare le radici di P e di Q. Vogliamo infatti verificare quando essi siano coprimi.

Irriducibilità della funzione di trasferimento

Dalla (3.58) si ha

$$P(s) = (s + k_{pf})R(s). \tag{3.73}$$

Per stabilire quando P e Q siano coprimi consideriamo separatamente il fattore lineare $s + k_{pf}$ e il fattore quadratico $R(s)$.

Il fattore lineare si annulla per

$$s = -k_{pf}. \tag{3.74}$$

Valutando (3.71) in tale punto si ottiene

$$Q(-k_{pf}) = k_{fp}k_{pf}^2 - [k_{fp}(k_{om} + k_{mo}) + k_{mp}k_{om}]\, k_{pf} + k_{mp}k_{om}k_{pf} \tag{3.75}$$

$$= k_{fp}k_{pf}\,(k_{pf} - k_{om} - k_{mo})\,. \tag{3.76}$$

Poiché $k_{fp} > 0$ e $k_{pf} > 0$, si ha

$$Q(-k_{pf}) = 0 \tag{3.77}$$

se e solo se

$$k_{pf} = k_{om} + k_{mo}. \tag{3.78}$$

Pertanto il fattore $s + k_{pf}$ non è comune a P e Q se e solo se

$$k_{pf} \neq k_{om} + k_{mo}. \tag{3.79}$$

Consideriamo ora il fattore quadratico $R(s)$. Dalle (3.57) e (3.71) si ha

$$\begin{aligned} Q(s) - k_{fp}R(s) &= k_{mp}(k_{om} - k_{fp})s + k_{mp}\left(k_{om}k_{pf} - k_{fp}k_{mo}\right) && (3.80)\\ &= k_{mp}\left[(k_{om} - k_{fp})s + k_{om}k_{pf} - k_{fp}k_{mo}\right]. && (3.81) \end{aligned}$$

Supponiamo che s_0 sia una radice comune di Q e R. Allora

$$Q(s_0) = R(s_0) = 0 \tag{3.82}$$

e dalla (3.81) segue necessariamente

$$(k_{om} - k_{fp})s_0 + k_{om}k_{pf} - k_{fp}k_{mo} = 0. \tag{3.83}$$

Consideriamo dapprima il caso

$$k_{om} \neq k_{fp}. \tag{3.84}$$

Dalla (3.83) l'unica possibile radice comune è

$$s_0 = \frac{k_{fp}k_{mo} - k_{om}k_{pf}}{k_{om} - k_{fp}}. \tag{3.85}$$

Affinché tale valore sia effettivamente una radice comune deve inoltre soddisfare

$$R(s_0) = 0. \tag{3.86}$$

Sostituendo la (3.85) in R e semplificando si ottiene

$$R(s_0) = \frac{k_{om}\Gamma}{(k_{om} - k_{fp})^2}, \tag{3.87}$$

dove

$$\begin{aligned} \Gamma &= k_{mp}(k_{fp} - k_{om})(k_{pf} - k_{mo}) && (3.88)\\ &\quad - (k_{fp}k_{mo} - k_{pf}k_{om})(k_{fp} + k_{pf} - k_{om} - k_{mo}). && (3.89) \end{aligned}$$

Poiché $k_{om} > 0$ e $k_{om} \neq k_{fp}$,

$$R(s_0) = 0 \tag{3.90}$$

se e solo se

$$\Gamma = 0. \tag{3.91}$$

Rimane da considerare il caso

$$k_{om} = k_{fp}. \tag{3.92}$$

In tal caso la (3.81) diventa

$$Q(s) - k_{fp}R(s) = k_{mp}k_{fp}(k_{pf} - k_{mo}). \tag{3.93}$$

Se

$$k_{pf} \neq k_{mo}, \tag{3.94}$$

la differenza $Q - k_{fp}R$ è una costante non nulla e quindi Q e R non possono avere radici comuni. Se invece

$$k_{pf} = k_{mo}, \tag{3.95}$$

si ha

$$Q(s) = k_{fp}R(s), \tag{3.96}$$

e pertanto R è un fattore comune di P e Q.

Anche quest'ultimo caso è contenuto nella condizione (3.91). Infatti, ponendo $k_{om} = k_{fp}$ nella (3.89), si ottiene

$$\Gamma = k_{fp}(k_{pf} - k_{mo})^2. \tag{3.97}$$

Poiché $k_{fp} > 0$, segue che

$$\Gamma = 0 \tag{3.98}$$

se e solo se

$$k_{pf} = k_{mo}. \tag{3.99}$$

Possiamo pertanto concludere che, nel modello generale a cinque coefficienti e senza introdurre ulteriori relazioni tra i parametri, la funzione di trasferimento è irriducibile se e solo se

$$k_{pf} \neq k_{om} + k_{mo} \tag{3.100}$$

e

$$\Gamma \neq 0, \tag{3.101}$$

dove Γ è definito dalla (3.89).

Mappa dei coefficienti della funzione di trasferimento

Andiamo a sviluppare i polinomi P e Q:

$$P(s) = s^3 + a_2 s^2 + a_1 s + a_0, \qquad Q(s) = b_2 s^2 + b_1 s + b_0. \tag{3.102}$$

Confrontando i termini di pari grado di P e di Q (3.60) e (3.71) otteniamo che

$$\begin{cases} a_2 = k_{mp} + k_{om} + k_{mo} + k_{pf}, \\ a_1 = k_{mp}k_{mo} + k_{pf}(k_{mp} + k_{om} + k_{mo}), \\ a_0 = k_{mp}k_{mo}k_{pf}, \\ b_2 = k_{fp}, \\ b_1 = k_{fp}(k_{om} + k_{mo}) + k_{mp}k_{om}, \\ b_0 = k_{mp}k_{om}k_{pf}. \end{cases} \tag{3.103}$$

Introduciamo l'insieme dei parametri ove P e Q risultano coprimi:

$$\mathcal{U} = \left\{ \mathbf{k} \in (0, +\infty)^5 : k_{pf} \neq k_{om} + k_{mo}, \quad \Gamma(\mathbf{k}) \neq 0 \right\}. \tag{3.104}$$

Possiamo quindi definire la mappa dei coefficienti

$$\Psi : \mathcal{U} \longrightarrow \mathbb{R}^6, \tag{3.105}$$

e poniamo

$$\Psi(\mathbf{k}) = (a_2, a_1, a_0, b_2, b_1, b_0)^T$$

La sola positività dei coefficienti cinetici non è sufficiente a garantire l'irriducibilità della funzione di trasferimento. Le condizioni (3.100) e (3.101) individuano la regione dello spazio dei parametri nella quale la rappresentazione razionale ottenuta è irriducibile.

Siano ora $\mathbf{k}, \mathbf{k}' \in \mathcal{U}$ due vettori di parametri che determinano la stessa funzione di trasferimento. Poiché, per definizione di $\mathcal{U}$, le corrispondenti rappresentazioni razionali sono irriducibili e i denominatori sono monici, dal teorema sull'uguaglianza delle funzioni razionali irriducibili segue

$$P(s) = P'(s), \qquad Q(s) = Q'(s).$$

Di conseguenza i sei coefficienti coincidono e quindi

$$\Psi(\mathbf{k}) = \Psi(\mathbf{k}'). \tag{3.106}$$

Pertanto, nella regione $\mathcal{U}$, due vettori di parametri determinano la stessa funzione di trasferimento se e solo se determinano gli stessi coefficienti dei polinomi P e Q, ossia se e solo se

$$\Psi(\mathbf{k}) = \Psi(\mathbf{k}').$$

Di conseguenza, lo studio dell'identificabilità del modello nella regione $\mathcal{U}$ può essere ricondotto allo studio dell'iniettività della mappa dei coefficienti Ψ.

I punti di $(0, +\infty)^5 \setminus \mathcal{U}$, nei quali la rappresentazione razionale è riducibile, dovranno invece essere considerati separatamente.

Jacobiana della mappa dei coefficienti

Per studiare l'identificabilità locale del modello consideriamo la matrice Jacobiana della mappa

$$\Psi : (0, +\infty)^5 \longrightarrow \mathbb{R}^6.$$

In questa fase manteniamo come dominio l'intero spazio dei parametri ammissibili in modo da determinare per quali valori dei parametri la Jacobiana abbia rango pieno.

Mantenendo l'ordine

$$\mathbf{k} = (k_{mp}, k_{fp}, k_{pf}, k_{om}, k_{mo})^T$$

per i parametri e

$$\Psi(\mathbf{k}) = (a_2, a_1, a_0, b_2, b_1, b_0)^T$$

per le componenti della mappa, si ottiene

$$J\Psi(\mathbf{k}) = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ k_{mo} + k_{pf} & 0 & k_{mp} + k_{om} + k_{mo} & k_{pf} & k_{mp} + k_{pf} \\ k_{mo}k_{pf} & 0 & k_{mp}k_{mo} & 0 & k_{mp}k_{pf} \\ 0 & 1 & 0 & 0 & 0 \\ k_{om} & k_{om} + k_{mo} & 0 & k_{fp} + k_{mp} & k_{fp} \\ k_{om}k_{pf} & 0 & k_{mp}k_{om} & k_{mp}k_{pf} & 0 \end{pmatrix}. \tag{3.107}$$

Essendo $J\Psi(\mathbf{k})$ una matrice 6×5, il suo rango massimo è pari a 5. Osserviamo

innanzitutto che la quarta componente della mappa Ψ è

$$b_2 = k_{fp}.$$

Pertanto la quarta riga della Jacobiana è

$$\nabla b_2 = (0, 1, 0, 0, 0).$$

La seconda colonna della Jacobiana, corrispondente al parametro k_{fp}, è quindi

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ k_{om} + k_{mo} \\ 0 \end{pmatrix}.$$

Eseguiamo la trasformazione

$$R_5 = R_5 - (k_{om} + k_{mo})R_4,$$

che non modifica il rango. Dopo tale operazione la seconda colonna diventa

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Dopo tale operazione, la seconda colonna contiene un unico elemento non nullo, uguale a 1, nella quarta riga. Questo elemento costituisce un pivot indipendente dal resto della matrice. Eliminando quindi la quarta riga e la seconda colonna, otteniamo una matrice 5×4 $A(\mathbf{k})$ tale che

$$\operatorname{rank} J\Psi(\mathbf{k}) = 1 + \operatorname{rank} A(\mathbf{k}).$$

$$A(\mathbf{k}) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ k_{mo}+k_{pf} & k_{mp}+k_{om}+k_{mo} & k_{pf} & k_{mp}+k_{pf} \\ k_{mo}k_{pf} & k_{mp}k_{mo} & 0 & k_{mp}k_{pf} \\ k_{om} & 0 & k_{fp}+k_{mp} & k_{fp} \\ k_{om}k_{pf} & k_{mp}k_{om} & k_{mp}k_{pf} & 0 \end{pmatrix}, \tag{3.108}$$

le cui colonne corrispondono, nell'ordine, ai parametri

$$k_{mp}, \qquad k_{pf}, \qquad k_{om}, \qquad k_{mo}.$$

Allora

$$\operatorname{rank} J\Psi(\mathbf{k}) = 5 \quad \Longleftrightarrow \quad \operatorname{rank} A(\mathbf{k}) = 4.$$

Per determinare quando ciò avviene consideriamo i minori di ordine 4 di A. Indichiamo con M_i il determinante della matrice ottenuta eliminando la i-esima riga di A. Due di tali minori assumono la forma

$$M_4 = k_{mp}k_{pf}\,(k_{mp} - k_{mo} - k_{om})\,\Delta(\mathbf{k}), \tag{3.109}$$

$$M_5 = k_{mp}\,(k_{mp} + k_{fp} - k_{mo} - k_{om})\,\Delta(\mathbf{k}), \tag{3.110}$$

dove

$$\Delta(\mathbf{k}) = k_{mp}k_{mo} - k_{mp}k_{pf} - k_{mo}k_{pf} - k_{om}k_{pf} + k_{pf}^2. \tag{3.111}$$

La stessa quantità può essere scritta come

$$\Delta(\mathbf{k}) = (k_{mp} - k_{pf})(k_{mo} - k_{pf}) - k_{om}k_{pf}. \tag{3.112}$$

Supponiamo inizialmente che

$$\Delta(\mathbf{k}) \neq 0.$$

Poiché $\mathbf{k} \in (0, +\infty)^5$, si ha

$$k_{mp} > 0, \qquad k_{pf} > 0, \qquad k_{fp} > 0.$$

I due fattori

$$k_{mp} - k_{mo} - k_{om}$$

e

$$k_{mp} + k_{fp} - k_{mo} - k_{om}$$

non possono annullarsi contemporaneamente, poiché la loro differenza è $k_{fp} > 0$. Di conseguenza, se $\Delta(\mathbf{k}) \neq 0$, almeno uno tra M_4 e M_5 è diverso da zero. Pertanto

$$\Delta(\mathbf{k}) \neq 0 \quad \Longrightarrow \quad \operatorname{rank} J\Psi(\mathbf{k}) = 5. \tag{3.113}$$

Resta da considerare il caso

$$\Delta(\mathbf{k}) = 0.$$

Dalla (3.111) segue

$$k_{om} = \frac{(k_{mp} - k_{pf})(k_{mo} - k_{pf})}{k_{pf}}. \tag{3.114}$$

Poiché $k_{om} > 0$ e $k_{pf} > 0$, deve valere

$$(k_{mp} - k_{pf})(k_{mo} - k_{pf}) > 0,$$

da cui, in particolare,

$$k_{mp} \neq k_{pf}.$$

Sostituendo la relazione (3.114) nei tre minori rimanenti si ottiene

$$M_1 = -k_{mp}(k_{mp} - k_{pf})\Lambda(\mathbf{k}), \tag{3.115}$$

$$M_2 = -\frac{k_{mp}(k_{mp} - k_{pf})}{k_{pf}}\Lambda(\mathbf{k}), \tag{3.116}$$

$$M_3 = -\frac{k_{mp}(k_{mp} - k_{pf})}{k_{pf}^2}\Lambda(\mathbf{k}), \tag{3.117}$$

dove

$$\begin{aligned}\Lambda(\mathbf{k}) = {} & k_{mp}(k_{pf} - k_{mo})\left(k_{mp}k_{mo} - 2k_{mp}k_{pf} + k_{pf}^2\right) \\ & + k_{fp}k_{pf}\left(k_{mp}k_{mo} - k_{mp}k_{pf} + k_{pf}^2\right).\end{aligned} \tag{3.118}$$

Nel caso $\Delta(\mathbf{k}) = 0$ i fattori

$$k_{mp}, \qquad k_{pf}, \qquad k_{mp} - k_{pf}$$

sono tutti diversi da zero. Pertanto almeno uno dei minori M_1, M_2, M_3 è diverso da zero se e solo se

$$\Lambda(\mathbf{k}) \neq 0.$$

Ne segue che

$$\Delta(\mathbf{k}) = 0, \qquad \Lambda(\mathbf{k}) \neq 0 \quad \Longrightarrow \quad \operatorname{rank} J\Psi(\mathbf{k}) = 5. \tag{3.119}$$

Se invece

$$\Delta(\mathbf{k}) = 0, \qquad \Lambda(\mathbf{k}) = 0,$$

tutti i minori di ordine 5 della Jacobiana si annullano e quindi

$$\operatorname{rank} J\Psi(\mathbf{k}) \leq 4.$$

D'altra parte, il minore di ordine 4 ottenuto selezionando le righe $1, 4, 5, 6$ e le colonne corrispondenti ai parametri $k_{mp}, k_{fp}, k_{om}, k_{mo}$ è

$$\det \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ k_{om} & k_{om} + k_{mo} & k_{fp} + k_{mp} & k_{fp} \\ k_{om}k_{pf} & 0 & k_{mp}k_{pf} & 0 \end{pmatrix} = -k_{mp}k_{fp}k_{pf}.$$

Poiché i parametri sono positivi, tale determinante non si annulla. Pertanto

$$\operatorname{rank} J\Psi(\mathbf{k}) \geq 4.$$

Possiamo quindi caratterizzare completamente il rango della Jacobiana:

$$\operatorname{rank} J\Psi(\mathbf{k}) = \begin{cases} 5, & \Delta(\mathbf{k}) \neq 0, \\ 5, & \Delta(\mathbf{k}) = 0 \text{ e } \Lambda(\mathbf{k}) \neq 0, \\ 4, & \Delta(\mathbf{k}) = 0 \text{ e } \Lambda(\mathbf{k}) = 0. \end{cases} \tag{3.120}$$

La Jacobiana ha dunque rango pieno in tutti i punti di $(0, +\infty)^5$, ad eccezione dell'insieme definito simultaneamente dalle condizioni

$$\Delta(\mathbf{k}) = 0, \qquad \Lambda(\mathbf{k}) = 0.$$

3.2.5 Identificabilità locale

Possiamo ora utilizzare il risultato sul rango della Jacobiana per ottenere una condizione sufficiente di unicità locale dei parametri.

Sia

$$\mathbf{k}^* \in \mathcal{U}$$

tale che

$$\operatorname{rank} D\Psi(\mathbf{k}^*) = 5. \tag{3.121}$$

Poiché $D\Psi(\mathbf{k}^*)$ ha rango massimo, esiste almeno un suo minore 5×5 non nullo. Esistono quindi cinque componenti della mappa Ψ, che indichiamo con

$$\Psi_{i_1}, \ldots, \Psi_{i_5},$$

tali che la mappa

$$\widehat{\Psi} = (\Psi_{i_1}, \ldots, \Psi_{i_5})$$

ha matrice Jacobiana invertibile nel punto $\mathbf{k}^*$.

L'insieme $\mathcal{U}$ è aperto, essendo definito mediante condizioni di non annullamento di funzioni continue. Possiamo quindi applicare il teorema della funzione inversa alla mappa $\widehat{\Psi}$. Esiste pertanto un intorno $V \subset \mathcal{U}$ di $\mathbf{k}^*$ nel quale $\widehat{\Psi}$ è iniettiva.

Sia ora $\mathbf{k} \in V$ un secondo vettore di parametri che determina la stessa funzione di trasferimento associata a $\mathbf{k}^*$, ossia

$$\frac{Q(s;\mathbf{k})}{P(s;\mathbf{k})} = \frac{Q(s;\mathbf{k}^*)}{P(s;\mathbf{k}^*)}.$$

Poiché

$$\mathbf{k}, \mathbf{k}^* \in \mathcal{U},$$

le corrispondenti rappresentazioni razionali sono irriducibili. Inoltre i denominatori P sono monici. Dall'unicità della rappresentazione razionale irriducibile segue quindi

$$P(s;\mathbf{k}) = P(s;\mathbf{k}^*), \qquad Q(s;\mathbf{k}) = Q(s;\mathbf{k}^*).$$

I coefficienti dei due polinomi coincidono e pertanto

$$\Psi(\mathbf{k}) = \Psi(\mathbf{k}^*).$$

In particolare,

$$\widehat{\Psi}(\mathbf{k}) = \widehat{\Psi}(\mathbf{k}^*).$$

Dall'iniettività di $\widehat{\Psi}$ nell'intorno V segue infine

$$\mathbf{k} = \mathbf{k}^*.$$

Abbiamo quindi dimostrato che, per ogni $\mathbf{k}^* \in \mathcal{U}$ tale che

$$\operatorname{rank} D\Psi(\mathbf{k}^*) = 5,$$

il vettore $\mathbf{k}^*$ è localmente univocamente determinato dalla funzione di trasferimento: esiste cioè un intorno di $\mathbf{k}^*$ nel quale non esistono altri vettori di parametri che producano la stessa funzione di trasferimento.

Dalla caratterizzazione del rango ottenuta nella sezione precedente, tale proprietà vale se

$$\Delta(\mathbf{k}^*) \neq 0,$$

oppure se

$$\Delta(\mathbf{k}^*) = 0, \qquad H(\mathbf{k}^*) \neq 0.$$

Pertanto la mappa dei coefficienti è localmente iniettiva in tutti i punti $\mathbf{k}^* \in \mathcal{U}$ nei quali

$$\operatorname{rank} D\Psi(\mathbf{k}^*) = 5,$$

che costituiscono i punti regolari della mappa Ψ.

Rimangono da considerare, all'interno di $\mathcal{U}$, i punti per i quali

$$\Delta(\mathbf{k}^*) = 0, \qquad H(\mathbf{k}^*) = 0.$$

In tali punti la Jacobiana ha rango 4. La perdita di rango non implica, tuttavia, la non identificabilità: indica soltanto che il criterio basato sulla Jacobiana non è sufficiente, da solo, a stabilire l'iniettività locale della mappa dei coefficienti.

Lo studio del rango fornisce quindi un risultato di unicità locale nei punti regolari. La classificazione dell'identificabilità del modello, nel senso introdotto precedentemente attraverso il numero delle soluzioni ammissibili dell'exhaustive summary, verrà completata nel seguito mediante lo studio diretto del sistema algebrico associato alla mappa Ψ.

3.2.6 Identificabilità globale

Passiamo ora allo studio dell'identificabilità globale del modello generale a cinque parametri. A differenza del caso locale, nel quale abbiamo studiato il rango della matrice Jacobiana della mappa dei coefficienti, consideriamo direttamente le relazioni algebriche tra i parametri cinetici e i coefficienti dei polinomi P e Q.

Siano

$$\mathbf{k}, \mathbf{k}' \in \mathcal{U}$$

due vettori di parametri che determinano la stessa funzione di trasferimento. Poiché nella regione $\mathcal{U}$ le corrispondenti rappresentazioni razionali sono irriducibili e i denominatori sono monici, l'uguaglianza delle funzioni di trasferimento implica

$$P(s;\mathbf{k}) = P(s;\mathbf{k}'), \qquad Q(s;\mathbf{k}) = Q(s;\mathbf{k}').$$

Di conseguenza, i sei coefficienti dei due polinomi coincidono e quindi

$$\Psi(\mathbf{k}) = \Psi(\mathbf{k}').$$

Nella regione $\mathcal{U}$, lo studio dell'identificabilità globale può pertanto essere ricondotto al problema di determinare quanti vettori positivi di parametri possano produrre uno stesso insieme di coefficienti

$$(a_2, a_1, a_0, b_2, b_1, b_0).$$

Ricordiamo che tali coefficienti soddisfano

$$\begin{cases} a_2 = k_{mp} + k_{om} + k_{mo} + k_{pf}, \\ a_1 = k_{mp}k_{mo} + k_{pf}(k_{mp} + k_{om} + k_{mo}), \\ a_0 = k_{mp}k_{mo}k_{pf}, \\ b_2 = k_{fp}, \\ b_1 = k_{fp}(k_{om} + k_{mo}) + k_{mp}k_{om}, \\ b_0 = k_{mp}k_{om}k_{pf}. \end{cases} \tag{3.122}$$

Alcuni parametri, o rapporti tra essi, possono essere ricavati direttamente dai coefficienti. Dalla quarta equazione di (3.122) si ottiene

$$k_{fp} = b_2. \tag{3.123}$$

Inoltre, poiché tutti i parametri sono positivi, a_0 e b_0 sono strettamente positivi e possiamo considerare il rapporto

$$\frac{a_0}{b_0} = \frac{k_{mp}k_{mo}k_{pf}}{k_{mp}k_{om}k_{pf}} = \frac{k_{mo}}{k_{om}}. \tag{3.124}$$

Poniamo quindi

$$\rho = \frac{a_0}{b_0} > 0, \tag{3.125}$$

da cui segue

$$k_{mo} = \rho k_{om}. \tag{3.126}$$

Cerchiamo ora di esprimere tutti i parametri rimanenti in funzione di una sola incognita. Poniamo

$$z = k_{pf}. \tag{3.127}$$

Poiché $k_{pf} > 0$, deve essere $z > 0$.

Dalla relazione

$$b_0 = k_{mp} k_{om} k_{pf}$$

segue

$$k_{mp} k_{om} = \frac{b_0}{z}. \tag{3.128}$$

Utilizzando

$$k_{fp} = b_2, \qquad k_{mo} = \rho k_{om},$$

la relazione relativa al coefficiente b_1 diventa

$$b_1 = b_2(1+\rho)k_{om} + k_{mp}k_{om}.$$

Sostituendo (3.128), otteniamo

$$b_1 = b_2(1+\rho)k_{om} + \frac{b_0}{z},$$

da cui

$$k_{om} = \frac{b_1 - \dfrac{b_0}{z}}{b_2(1+\rho)}.$$

Utilizzando

$$\rho = \frac{a_0}{b_0},$$

possiamo riscrivere tale espressione come

$$k_{om}(z) = \frac{b_0(b_1 z - b_0)}{z\, b_2(a_0 + b_0)}. \tag{3.129}$$

Dalla relazione $k_{mo} = \rho k_{om}$ segue

$$k_{mo}(z) = \frac{a_0(b_1 z - b_0)}{z\, b_2(a_0 + b_0)}. \tag{3.130}$$

Infine, tornando alla relazione

$$b_0 = k_{mp} k_{om} z,$$

si ricava

$$k_{mp}(z) = \frac{b_2(a_0 + b_0)}{b_1 z - b_0}. \tag{3.131}$$

Pertanto, una volta fissato $z = k_{pf}$, tutti gli altri parametri sono univocamente determinati:

$$\begin{cases} k_{fp} = b_2, \\ k_{pf} = z, \\ k_{mp} = \dfrac{b_2(a_0 + b_0)}{b_1 z - b_0}, \\ k_{om} = \dfrac{b_0(b_1 z - b_0)}{z\, b_2(a_0 + b_0)}, \\ k_{mo} = \dfrac{a_0(b_1 z - b_0)}{z\, b_2(a_0 + b_0)}. \end{cases} \tag{3.132}$$

Poiché a_0, b_0, b_1 e b_2 sono positivi, la positività dei parametri richiede

$$b_1 z - b_0 > 0.$$

I valori ammissibili dell'incognita devono quindi soddisfare

$$z > \frac{b_0}{b_1}. \tag{3.133}$$

Rimane da determinare quali valori di z siano compatibili anche con i coefficienti a_2 e a_1. Consideriamo inizialmente a_2. Sostituendo nella prima equazione di (3.122) le espressioni (3.132), si ottiene

$$a_2 = z + \frac{b_2(a_0 + b_0)}{b_1 z - b_0} + \frac{b_1 z - b_0}{b_2 z}. \tag{3.134}$$

Portando tutti i termini al primo membro e moltiplicando per

$$z b_2 (b_1 z - b_0),$$

che è strettamente positivo nell'intervallo ammissibile (3.133), otteniamo l'equazione polinomiale

$$G(z) = 0, \tag{3.135}$$

dove

$$\begin{aligned} G(z) = {} & (z - a_2) z b_2 (b_1 z - b_0) \\ & + z b_2^2 (a_0 + b_0) + (b_1 z - b_0)^2. \end{aligned} \tag{3.136}$$

Il polinomio G ha grado 3.

Consideriamo ora il coefficiente a_1. Poiché

$$a_0 = k_{mp}k_{mo}z,$$

si ha

$$k_{mp}k_{mo} = \frac{a_0}{z}.$$

Dalla relazione relativa ad a_2 segue inoltre

$$k_{mp} + k_{om} + k_{mo} = a_2 - z.$$

Pertanto

$$a_1 = \frac{a_0}{z} + z(a_2 - z).$$

Moltiplicando per $z > 0$ otteniamo

$$C(z) = z^3 - a_2z^2 + a_1z - a_0 = 0. \tag{3.137}$$

Osserviamo che

$$P(-z) = -z^3 + a_2z^2 - a_1z + a_0 = -C(z),$$

e quindi

$$C(z) = 0 \quad \Longleftrightarrow \quad P(-z) = 0.$$

Ciò è coerente con la fattorizzazione

$$P(s) = (s + k_{pf})\left[s^2 + (k_{mp} + k_{om} + k_{mo})s + k_{mp}k_{mo}\right],$$

poiché $z = k_{pf}$.

Abbiamo quindi ricondotto il problema alla ricerca dei valori reali z che soddisfano

$$\begin{cases} C(z) = 0, \\ G(z) = 0, \\ z > \dfrac{b_0}{b_1}. \end{cases} \tag{3.138}$$

Per ogni soluzione ammissibile z del sistema (3.138), le formule (3.132) determinano un unico vettore positivo di parametri. Viceversa, ogni vettore positivo che produce i coefficienti fissati determina, attraverso $z = k_{pf}$, una soluzione ammissibile dello stesso sistema.

Esiste pertanto una corrispondenza biunivoca tra i vettori positivi di parametri compatibili con il sestetto di coefficienti fissato e le radici comuni ammissibili dei polinomi C e G.

Poiché C è un polinomio monico di grado 3, tali radici sono in numero finito e non possono essere più di tre. Uno stesso sestetto di coefficienti può quindi essere prodotto da al più tre vettori positivi di parametri.

Studio delle radici comuni

Per stabilire se il valore $z = k_{pf}$ sia unico, studiamo più nel dettaglio i polinomi C e G.

Per alleggerire temporaneamente la notazione poniamo

$$x = k_{mp}, \qquad f = k_{fp}, \qquad p = k_{pf}, \qquad o = k_{om}, \qquad m = k_{mo}.$$

I coefficienti della funzione di trasferimento assumono quindi la forma

$$\begin{cases} a_2 = x + o + m + p, \\ a_1 = xm + p(x + o + m), \\ a_0 = xmp, \\ b_2 = f, \\ b_1 = f(o + m) + xo, \\ b_0 = xop. \end{cases} \tag{3.139}$$

Sostituendo tali espressioni nel polinomio C otteniamo

$$C(z) = (z - p)\left[z^2 - (x + o + m)z + xm\right]. \tag{3.140}$$

Poniamo

$$S(z) = z^2 - (x + o + m)z + xm, \tag{3.141}$$

così che

$$C(z) = (z - p)S(z).$$

La presenza del fattore $z - p$ è coerente con

$$p = k_{pf}, \qquad P(-p) = 0.$$

Sostituendo analogamente (3.139) nell'espressione di G e raccogliendo i termini,

si ottiene

$$G(z) = (z - p)R(z), \tag{3.142}$$

dove

$$\begin{aligned} R(z) = {} & f\left[f(m+o) + ox\right] z^2 \\ & - x\left[f^2(m+o) - fo(m+o) + fo(p+x) - o^2x\right] z \\ & - o^2px^2. \end{aligned} \tag{3.143}$$

Il valore

$$z = p = k_{pf}$$

è quindi sempre una radice comune di C e G. Eventuali ulteriori valori compatibili con lo stesso sestetto di coefficienti devono essere radici comuni dei due polinomi quadratici S e R.

Se S e R non hanno radici comuni, l'unica radice comune di C e G è pertanto

$$z = p = k_{pf}.$$

Per la (3.132), l'unicità di z implica l'unicità dell'intero vettore dei parametri.

Risultante dei polinomi quadratici

Per stabilire quando S e R possiedano una radice comune, consideriamo il loro risultante rispetto alla variabile z:

$$\mathcal{R}(\mathbf{k}) = \mathrm{Res}_z(S, R). \tag{3.144}$$

Per le proprietà del risultante,

$$\mathcal{R}(\mathbf{k}) \neq 0 \quad \Longleftrightarrow \quad S \text{ e } R \text{ non hanno radici comuni.} \tag{3.145}$$

Pertanto

$$\mathcal{R}(\mathbf{k}) \neq 0 \quad \Longrightarrow \quad z = k_{pf} \text{ è l'unica radice comune di } C \text{ e } G. \tag{3.146}$$

Ne segue che, per ogni

$$\mathbf{k} \in \mathcal{U}$$

tale che

$$\mathcal{R}(\mathbf{k}) \neq 0,$$

non esiste un secondo vettore appartenente a $\mathcal{U}$ che produca la stessa funzione di trasferimento.

Mostriamo che tale unicità vale anche rispetto a eventuali vettori appartenenti a $(0,+\infty)^5 \setminus \mathcal{U}$. Supponiamo infatti che $\mathbf{k} \in \mathcal{U}$ e che $\mathbf{k}' \in (0,+\infty)^5$ determini la stessa funzione di trasferimento.

La funzione associata a $\mathbf{k}$ possiede una rappresentazione irriducibile con denominatore monico di grado 3. Se la rappresentazione associata a $\mathbf{k}'$ fosse riducibile, dopo la cancellazione di un fattore comune non costante si otterrebbe una rappresentazione irriducibile della stessa funzione razionale con denominatore di grado strettamente minore di 3. Ciò contraddice l'unicità, a meno di un fattore moltiplicativo costante, della rappresentazione razionale irriducibile associata a $\mathbf{k}$.

Pertanto anche $\mathbf{k}'$ deve appartenere a $\mathcal{U}$. Ne segue che, se

$$\mathbf{k} \in \mathcal{U}, \qquad \mathcal{R}(\mathbf{k}) \neq 0,$$

non esiste alcun secondo vettore

$$\mathbf{k}' \in (0,+\infty)^5$$

che produca la stessa funzione di trasferimento.

Il risultante $\mathcal{R}$ è un polinomio nei parametri cinetici e non è identicamente nullo. Per verificarlo è sufficiente considerare, ad esempio,

$$(k_{mp}, k_{fp}, k_{pf}, k_{om}, k_{mo}) = (2,1,1,1,2),$$

per il quale

$$\mathcal{R}(2,1,1,1,2) = -68 \neq 0. \tag{3.147}$$

L'insieme

$$\left\{\mathbf{k} \in (0,+\infty)^5 : \mathcal{R}(\mathbf{k}) = 0\right\}$$

è quindi un sottoinsieme algebrico proprio dello spazio dei parametri.

Un caso eccezionale

L'annullamento del risultante può effettivamente corrispondere a una perdita di unicità. Consideriamo i vettori

$$\mathbf{k}^{(1)} = \left(2,\ 1,\ \frac{5}{6},\ 1,\ 2\right)^T \tag{3.148}$$

e

$$\mathbf{k}^{(2)} = \left(\frac{3}{2},\, 1,\, 1,\, \frac{10}{9},\, \frac{20}{9}\right)^T. \tag{3.149}$$

Le componenti sono ordinate come

$$(k_{mp}, k_{fp}, k_{pf}, k_{om}, k_{mo}).$$

Sostituendo i due vettori nelle espressioni dei coefficienti si ottiene, in entrambi i casi,

$$a_2 = \frac{35}{6}, \qquad a_1 = \frac{49}{6}, \qquad a_0 = \frac{10}{3}, \tag{3.150}$$

e

$$b_2 = 1, \qquad b_1 = 5, \qquad b_0 = \frac{5}{3}. \tag{3.151}$$

Essi determinano quindi gli stessi polinomi

$$P(s) = s^3 + \frac{35}{6}s^2 + \frac{49}{6}s + \frac{10}{3}, \tag{3.152}$$

e

$$Q(s) = s^2 + 5s + \frac{5}{3}. \tag{3.153}$$

In questo caso

$$C(z) = \frac{1}{6}(z-4)(z-1)(6z-5), \tag{3.154}$$

e

$$G(z) = \frac{5}{18}(z-1)(3z+2)(6z-5). \tag{3.155}$$

Le due radici comuni positive sono

$$z = \frac{5}{6} \qquad \text{e} \qquad z = 1,$$

e corrispondono precisamente ai due diversi valori di k_{pf} dei vettori $\mathbf{k}^{(1)}$ e $\mathbf{k}^{(2)}$.

Un calcolo esplicito mostra inoltre che i corrispondenti polinomi P e Q sono coprimi. Entrambi i vettori appartengono quindi alla regione $\mathcal{U}$ e la non unicità non è dovuta a una cancellazione di fattori comuni.

Conclusione sull'identificabilità globale

Lo studio precedente mostra che il vettore dei parametri è univocamente determinato per ogni

$$\mathbf{k} \in \mathcal{U}$$

tale che

$$\mathcal{R}(\mathbf{k}) \neq 0.$$

Indichiamo con λ^5 la misura di Lebesgue su $\mathbb{R}^5$. Quando consideriamo un sottoinsieme di $(0,\infty)^5$, la misura è quella indotta dalla restrizione di λ^5 a $(0,+\infty)^5$.

Ricordiamo ora un risultato della teoria delle funzioni reali analitiche.

Lemma 3.1. Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione reale analitica non identicamente nulla. Allora il suo insieme degli zeri

$$Z(f) = \{x \in \mathbb{R}^n : f(x) = 0\}$$

ha misura di Lebesgue nulla in $\mathbb{R}^n$.

Il risultato segue dal fatto che una funzione reale analitica che si annulla su un insieme di misura positiva è identicamente nulla [6, p. 67].

Questo lemma è utile nel nostro caso poiché si applica a ogni polinomio reale non identicamente nullo, dal momento che ogni funzione polinomiale è reale analitica.

Dalla definizione della regione di irriducibilità $\mathcal{U}$ si ha

$$\begin{aligned}(0,+\infty)^5 \setminus \mathcal{U} = {} & \left\{\mathbf{k} \in (0,+\infty)^5 : k_{pf} - k_{om} - k_{mo} = 0\right\} \\ & \cup \left\{\mathbf{k} \in (0,+\infty)^5 : \Gamma(\mathbf{k}) = 0\right\}.\end{aligned} \tag{3.156}$$

Le funzioni

$$k_{pf} - k_{om} - k_{mo}$$

e

$$\Gamma(\mathbf{k})$$

sono polinomi nelle componenti di $\mathbf{k}$ e non sono identicamente nulli. In particolare, per il secondo si verifica, ad esempio,

$$\Gamma(2,1,1,2,1) = -1 \neq 0.$$

Per il lemma precedente, i rispettivi insiemi degli zeri hanno misura λ^5 nulla. Poiché l'unione finita di insiemi di misura nulla ha ancora misura nulla, segue che

$$\lambda^5((0,+\infty)^5 \setminus \mathcal{U}) = 0. \tag{3.157}$$

Analogamente, il risultante

$$\mathcal{R}(\mathbf{k})$$

è un polinomio nei parametri cinetici. Poiché

$$\mathcal{R}(2,1,1,1,2) = -68 \neq 0,$$

esso non è identicamente nullo. Applicando nuovamente il lemma, l'insieme

$$\left\{\mathbf{k} \in (0,+\infty)^5 : \mathcal{R}(\mathbf{k}) = 0\right\}$$

ha misura λ^5 nulla.

Pertanto anche l'insieme

$$((0,+\infty)^5 \setminus \mathcal{U}) \cup \left\{\mathbf{k} \in (0,+\infty)^5 : \mathcal{R}(\mathbf{k}) = 0\right\}$$

ha misura λ^5 nulla.

Ne segue che, per quasi ogni

$$\mathbf{k} \in (0,\infty)^5,$$

la funzione di trasferimento determina univocamente il vettore dei parametri cinetici.

Secondo la definizione generica di identificabilità adottata precedentemente, il modello generale a cinque parametri risulta pertanto globalmente identificabile.

Il caso eccezionale appena mostrato evidenzia tuttavia che la mappa dai parametri alla funzione di trasferimento non è iniettiva in ogni punto di $(0,+\infty)^5$: esistono configurazioni appartenenti a un insieme algebrico di misura nulla per le quali vettori positivi distinti producono la stessa funzione di trasferimento.

3.3 Ipotesi fisiologiche di modello

Il lavoro precedente mostra che il modello generale a cinque parametri può essere analizzato direttamente dal punto di vista dell'identificabilità, senza introdurre ulteriori relazioni tra i coefficienti cinetici. Le ipotesi fisiologiche considerate nel seguito non vengono quindi introdotte allo scopo di rendere identificabile il modello, ma per ottenere una parametrizzazione più vincolata e coerente con alcune relazioni sperimentali proposte in letteratura.

In particolare, tali ipotesi permettono di ridurre il numero dei parametri indipendenti, imponendo relazioni tra i coefficienti che descrivono il trasporto della FDOPA e della 3-OMFD. Il modello ridotto che ne deriva costituisce quindi un caso particolare del modello generale studiato nelle sezioni precedenti. Nel capitolo successivo,

dedicato alla formulazione numerica del problema inverso, considereremo invece il modello generale a cinque parametri indipendenti.

Le relazioni adottate sono motivate dai risultati riportati da Huang et al. [5] e Kumakura et al. [7], che forniscono indicazioni sperimentali sui rapporti tra alcuni coefficienti di trasporto e distribuzione della FDOPA e della 3-OMFD.

3.3.1 Ipotesi di Huang

Huang et al. [5] osservano che il rapporto tra le efficienze di trasporto attraverso la barriera emato-encefalica della 3-OMFD e della FDOPA può essere descritto mediante una costante q. Nel loro studio viene ottenuto un valore prossimo a 1.7. Kumakura et al. [7] utilizzano la stessa ipotesi, fissando invece $q = 1.5$ sulla base di precedenti determinazioni sperimentali.

Assumiamo quindi

$$\frac{k_{om}}{k_{fp}} = q, \tag{3.158}$$

ossia

$$k_{om} = q\, k_{fp}. \tag{3.159}$$

Nel seguito della trattazione matematica non sarà necessario fissare uno specifico valore numerico di q: sarà sufficiente assumere

$$q > 1. \tag{3.160}$$

3.3.2 Ipotesi di Kumakura

Una seconda relazione tra i coefficienti cinetici viene ricavata dall'impostazione adottata da Kumakura et al. [7].

Nella notazione utilizzata dagli autori, K_1^D e K_1^M indicano rispettivamente i coefficienti di trasporto della FDOPA e della 3-OMFD dal sangue al tessuto cerebrale. I coefficienti k_2^D e k_2^M descrivono invece il corrispondente trasporto in direzione opposta, dal tessuto cerebrale al sangue.

Per una regione nella quale non si considera una componente significativa di legame specifico o di intrappolamento metabolico del tracciante, il volume di distribuzione all'equilibrio viene definito come

$$V_e = \frac{K_1}{k_2}.$$

Kumakura et al. assumono che FDOPA e 3-OMFD presentino lo stesso rapporto di distribuzione tra tessuto cerebrale e sangue all'equilibrio, espresso nel modello dal rapporto tra il coefficiente di ingresso nel tessuto e il corrispondente coefficiente di ritorno. Imponiamo quindi che

$$V_e = \frac{K_1^D}{k_2^D} = \frac{K_1^M}{k_2^M}. \tag{3.161}$$

Nel modello considerato nel presente lavoro le corrispondenze sono

$$K_1^D \longleftrightarrow k_{fp}, \qquad k_2^D \longleftrightarrow k_{pf}, \tag{3.162}$$

e

$$K_1^M \longleftrightarrow k_{om}, \qquad k_2^M \longleftrightarrow k_{mo}. \tag{3.163}$$

L'ipotesi di uguaglianza dei volumi di distribuzione viene quindi trasposta nel nostro modello nella forma

$$\frac{k_{fp}}{k_{pf}} = \frac{k_{om}}{k_{mo}}. \tag{3.164}$$

Questa ipotesi non fissa separatamente i coefficienti di ingresso e di uscita, ma impone una relazione tra essi, richiedendo che FDOPA e 3-OMFD presentino lo stesso volume di distribuzione all'equilibrio.

Combinando tale relazione con l'ipotesi precedentemente introdotta,

$$k_{om} = q\, k_{fp}, \tag{3.165}$$

si ottiene

$$\frac{k_{fp}}{k_{pf}} = \frac{q\, k_{fp}}{k_{mo}}. \tag{3.166}$$

Poiché $k_{fp} > 0$, segue

$$k_{mo} = q\, k_{pf}. \tag{3.167}$$

Le due relazioni utilizzate nella successiva analisi di identificabilità sono pertanto

$$\begin{cases} k_{om} = q\, k_{fp}, \\ k_{mo} = q\, k_{pf}, \end{cases} \qquad q > 1. \tag{3.168}$$

3.3.3 Identificabilità del modello vincolato

Introduciamo ora le relazioni tra i coefficienti cinetici ricavate dalle ipotesi di Huang e Kumakura e studiamo l'identificabilità del modello all'interno della corrispondente

famiglia vincolata. In particolare, assumiamo

$$k_{om} = qk_{fp}, \qquad k_{mo} = qk_{pf}, \qquad q > 1, \tag{3.169}$$

con q noto e fissato.

Sotto tali ipotesi il numero di coefficienti cinetici indipendenti si riduce da cinque a tre, e il vettore dei parametri liberi può essere considerato nella forma

$$\mathbf{k}_r = \begin{pmatrix} k_{mp} \\ k_{fp} \\ k_{pf} \end{pmatrix}. \tag{3.170}$$

Vogliamo verificare se la relazione ingresso-uscita determini univocamente questi tre coefficienti e, di conseguenza, anche k_{om} e k_{mo}.

Riscriviamo i polinomi associati al vettore $\mathbf{k}$

$$P(s) = (s + k_{pf}) \left[s^2 + (k_{mp} + qk_{fp} + qk_{pf})\, s + qk_{mp}k_{pf} \right], \tag{3.171}$$

e

$$Q(s) = k_{fp} \left[s^2 + q\,(k_{mp} + k_{fp} + k_{pf})\, s + qk_{mp}k_{pf} \right]. \tag{3.172}$$

Mostriamo ora che P e Q sono coprimi. Poniamo

$$R(s) = s^2 + (k_{mp} + qk_{fp} + qk_{pf})\, s + qk_{mp}k_{pf}. \tag{3.173}$$

Si ha quindi

$$P(s) = (s + k_{pf})R(s). \tag{3.174}$$

Inoltre,

$$\frac{Q(s)}{k_{fp}} = R(s) + (q-1)k_{mp}s. \tag{3.175}$$

Supponiamo che s_0 sia una radice di R. Allora

$$R(s_0) = 0 \tag{3.176}$$

e pertanto

$$\frac{Q(s_0)}{k_{fp}} = (q-1)k_{mp}s_0. \tag{3.177}$$

Poiché

$$q > 1, \qquad k_{mp} > 0,$$

affinché si abbia anche $Q(s_0) = 0$ dovrebbe essere

$$s_0 = 0.$$

Tuttavia

$$R(0) = qk_{mp}k_{pf} > 0, \tag{3.178}$$

per cui $s_0 = 0$ non può essere una radice di R. Ne segue che R e Q non hanno radici comuni.

Resta da verificare che la radice del fattore $s + k_{pf}$ non sia anche una radice di Q. Calcolando Q in $s = -k_{pf}$ si ottiene

$$Q(-k_{pf}) = k_{fp}k_{pf}\left[(1-q)k_{pf} - qk_{fp}\right] \tag{3.179}$$

$$= -k_{fp}k_{pf}\left[(q-1)k_{pf} + qk_{fp}\right]. \tag{3.180}$$

Poiché

$$k_{fp} > 0, \qquad k_{pf} > 0, \qquad q > 1,$$

si ha

$$Q(-k_{pf}) \neq 0. \tag{3.181}$$

Pertanto P e Q non hanno radici comuni e la funzione razionale

$$\frac{Q(s)}{P(s)}$$

è irriducibile.

Lo stesso argomento si applica ai polinomi P' e Q' associati al vettore $\mathbf{k}'$, poiché anche i coefficienti primati sono strettamente positivi e soddisfano le medesime ipotesi. Di conseguenza, anche

$$\frac{Q'(s)}{P'(s)}$$

è irriducibile.

Inoltre P e P' sono polinomi monici dello stesso grado. Per il teorema sull'uguaglianza delle funzioni razionali irriducibili segue quindi

$$P(s) = P'(s), \qquad Q(s) = Q'(s). \tag{3.182}$$

Confrontando i coefficienti dei polinomi P e P', e quelli dei polinomi Q e Q', si

ottiene il sistema

$$\begin{cases} k_{mp} + qk_{fp} + (q+1)k_{pf} = k'_{mp} + qk'_{fp} + (q+1)k'_{pf}, \\ k_{pf}\left[(q+1)k_{mp} + qk_{fp} + qk_{pf}\right] = k'_{pf}\left[(q+1)k'_{mp} + qk'_{fp} + qk'_{pf}\right], \\ qk_{mp}k_{pf}^2 = qk'_{mp}(k'_{pf})^2, \\ k_{fp} = k'_{fp}, \\ qk_{fp}\left(k_{mp} + k_{fp} + k_{pf}\right) = qk'_{fp}\left(k'_{mp} + k'_{fp} + k'_{pf}\right), \\ qk_{fp}k_{mp}k_{pf} = qk'_{fp}k'_{mp}k'_{pf}. \end{cases} \tag{3.183}$$

La quarta equazione del sistema fornisce direttamente

$$k_{fp} = k'_{fp}. \tag{3.184}$$

Utilizzando questa uguaglianza nella quinta equazione e ricordando che $q > 0$ e $k_{fp} > 0$, si ottiene

$$k_{mp} + k_{pf} = k'_{mp} + k'_{pf}. \tag{3.185}$$

Analogamente, dalla prima equazione del sistema segue

$$k_{mp} + (q+1)k_{pf} = k'_{mp} + (q+1)k'_{pf}. \tag{3.186}$$

Sottraendo le due relazioni si ricava

$$qk_{pf} = qk'_{pf}. \tag{3.187}$$

Poiché $q > 0$, segue

$$k_{pf} = k'_{pf}. \tag{3.188}$$

Sostituendo tale uguaglianza nella relazione

$$k_{mp} + k_{pf} = k'_{mp} + k'_{pf}, \tag{3.189}$$

si ottiene

$$k_{mp} = k'_{mp}. \tag{3.190}$$

Infine, dalle ipotesi di modello,

$$k_{om} = qk_{fp} = qk'_{fp} = k'_{om}, \tag{3.191}$$

e

$$k_{mo} = qk_{pf} = qk'_{pf} = k'_{mo}. \tag{3.192}$$

Pertanto

$$k_{mp} = k'_{mp}, \qquad k_{fp} = k'_{fp}, \qquad k_{pf} = k'_{pf}, \qquad k_{om} = k'_{om}, \qquad k_{mo} = k'_{mo}, \tag{3.193}$$

e dunque

$$\mathbf{k} = \mathbf{k}'. \tag{3.194}$$

Ne segue che, sotto le ipotesi considerate, il vettore dei coefficienti cinetici è univocamente determinato dalle funzioni osservate $y_b(t)$ e $y_c(t)$.

3.3.4 I casi $q = 1$ e $0 < q < 1$

Il parametro q è stato stimato in maniera sperimentale in entrambi gli studi citati come $q > 1$. Andiamo però adesso a indagare cosa possa comportare il fatto che il parametro q assuma valori diversi, in particolare nel caso dell'identificabilità. Ricordiamo che

$$k_{om} = qk_{fp}, \qquad k_{mo} = qk_{pf}, \tag{3.195}$$

con $q > 0$ noto e fissato. Introducendo

$$R_q(s) = s^2 + (k_{mp} + qk_{fp} + qk_{pf})s + qk_{mp}k_{pf}, \tag{3.196}$$

si ha

$$P(s) = (s + k_{pf})R_q(s) \tag{3.197}$$

e

$$Q(s) = k_{fp}\left[R_q(s) + (q-1)k_{mp}s\right]. \tag{3.198}$$

Nel caso $q = 1$ segue immediatamente che

$$Q(s) = k_{fp}R_1(s), \tag{3.199}$$

e pertanto

$$\frac{Q(s)}{P(s)} = \frac{k_{fp}}{s + k_{pf}}. \tag{3.200}$$

Il coefficiente k_{mp} non compare quindi nella funzione di trasferimento. Ne segue che valori diversi di k_{mp} producono la medesima relazione ingresso-uscita, a parità di k_{fp} e k_{pf}. Il modello non è pertanto strutturalmente identificabile per $q = 1$.

Consideriamo ora il caso $0 < q < 1$. Sostituendo le relazioni del modello vincolato nella condizione generale associata al fattore quadratico si ottiene

$$\Gamma = k_{mp} k_{fp} k_{pf} (q-1)^2. \tag{3.201}$$

Poiché i coefficienti cinetici sono strettamente positivi e $q \neq 1$, si ha

$$\Gamma > 0. \tag{3.202}$$

Il numeratore Q e il fattore quadratico R_q non hanno quindi radici comuni.

Rimane da esaminare il fattore lineare $s + k_{pf}$. Si ha

$$Q(-k_{pf}) = k_{fp} k_{pf} \left[(1-q) k_{pf} - q k_{fp} \right]. \tag{3.203}$$

Pertanto

$$Q(-k_{pf}) = 0 \tag{3.204}$$

se e solo se

$$q = \frac{k_{pf}}{k_{fp} + k_{pf}}. \tag{3.205}$$

Se la condizione (3.205) non è soddisfatta, P e Q sono coprimi e l'analisi di identificabilità può essere condotta mediante il confronto delle rappresentazioni razionali irriducibili.

Se invece vale (3.205), il fattore $s + k_{pf}$ è comune a numeratore e denominatore. In questo caso

$$Q(s) = k_{fp} (s + k_{pf})(s + q k_{mp}), \tag{3.206}$$

mentre

$$P(s) = (s + k_{pf}) \left[s^2 + (k_{mp} + k_{pf}) s + q k_{mp} k_{pf} \right]. \tag{3.207}$$

La funzione di trasferimento ridotta è quindi

$$\frac{Q(s)}{P(s)} = \frac{k_{fp} (s + q k_{mp})}{s^2 + (k_{mp} + k_{pf}) s + q k_{mp} k_{pf}}. \tag{3.208}$$

Anche in questo caso, tuttavia, i parametri rimangono determinati univocamente all'interno della famiglia vincolata con q noto. Il coefficiente del termine di primo grado al numeratore determina k_{fp}; dalla (3.205) si ricava quindi

$$k_{pf} = \frac{q}{1-q} k_{fp}, \tag{3.209}$$

mentre il termine noto del numeratore determina

$$qk_{fp}k_{mp}, \tag{3.210}$$

e quindi k_{mp}.

Il caso $q = 1$ rappresenta pertanto una situazione in cui non si ha identificabilità mentre per $0 < q < 1$ il modello rimane identificabile. La condizione $q > 1$ è quindi sufficiente a garantire automaticamente l'irriducibilità della funzione di trasferimento, ma risulta più restrittiva di quanto richiesto dall'identificabilità del modello.

Capitolo 4

Trattazione numerica del problema inverso

4.1 Introduzione

Nel capitolo precedente abbiamo affrontato il problema inverso dal punto di vista teorico, studiandone l'identificabilità strutturale. Abbiamo dimostrato che il modello generale a cinque parametri è genericamente globalmente identificabile.

Abbiamo inoltre introdotto alcune relazioni fisiologiche tra i coefficienti cinetici, ottenendo un modello ridotto a tre parametri. In questo capitolo consideriamo nuovamente il modello generale, senza imporre tali relazioni, e affrontiamo il problema della stima numerica dei cinque coefficienti cinetici indipendenti.

L'analisi dell'identificabilità strutturale presuppone tuttavia condizioni ideali, nelle quali la funzione di ingresso e la curva cerebellare sono note esattamente nell'intervallo temporale considerato. Nelle applicazioni reali la situazione è differente: le acquisizioni PET dinamiche forniscono un numero finito di osservazioni, inevitabilmente affette da errori sperimentali.

Non è quindi ragionevole richiedere che il modello riproduca esattamente tutti i dati osservati. Il problema viene invece formulato come un problema di approssimazione: si cercano i coefficienti cinetici per i quali la curva prevista dal modello risulta il più possibile vicina ai dati sperimentali.

Seguendo l'impostazione di Nocedal e Wright [9], il problema viene ricondotto a un problema di minimi quadrati non lineari. La particolare struttura della funzione obiettivo permette di utilizzare metodi numerici specifici, tra i quali il metodo di Gauss-Newton e il metodo di Levenberg-Marquardt.

L'identificabilità strutturale e la determinazione numerica dei parametri rappresentano quindi due aspetti distinti del problema inverso. L'identificabilità struttu-

rale generica garantisce, in condizioni ideali, l'unicità dei parametri per quasi ogni vettore ammissibile. Essa non garantisce però la stabilità della ricostruzione rispetto alle perturbazioni dei dati, né esclude la presenza di minimi locali della funzione obiettivo. Inoltre, non fornisce di per sé un procedimento per ricavare i parametri dalle osservazioni sperimentali.

4.2 Formulazione del problema ai minimi quadrati

4.2.1 Il modello diretto

Sia $T > 0$ il tempo finale dell'acquisizione PET e consideriamo l'intervallo temporale $[0, T]$.

Consideriamo il modello generale introdotto nel capitolo precedente, i cui cinque coefficienti cinetici indipendenti sono raccolti nel vettore

$$\mathbf{k} = \begin{pmatrix} k_{mp} \\ k_{fp} \\ k_{pf} \\ k_{om} \\ k_{mo} \end{pmatrix}. \tag{4.1}$$

L'insieme dei parametri considerato nell'analisi è $(0, +\infty)^5$, poiché tutti i coefficienti cinetici vengono assunti reali e strettamente positivi.

Per ogni $\mathbf{k} \in (0, +\infty)^5$, il sistema differenziale introdotto nel capitolo precedente può essere scritto nella forma

$$\begin{cases} \dot{\mathbf{C}}(t; \mathbf{k}) = M(\mathbf{k})\mathbf{C}(t; \mathbf{k}) + B(\mathbf{k})y_b(t), \\ \mathbf{C}(0; \mathbf{k}) = \mathbf{0}, \end{cases} \tag{4.2}$$

dove

$$M(\mathbf{k}) = \begin{pmatrix} -(k_{mp} + k_{om}) & 0 & k_{mo} \\ -k_{fp} & -k_{pf} & 0 \\ k_{om} & 0 & -k_{mo} \end{pmatrix} \tag{4.3}$$

e

$$B(\mathbf{k}) = \begin{pmatrix} k_{mp} \\ k_{fp} \\ 0 \end{pmatrix}. \tag{4.4}$$

La curva cerebellare prevista dal modello è

$$y_c(t;\mathbf{k}) = \alpha^T \mathbf{C}(t;\mathbf{k}), \qquad \alpha = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \tag{4.5}$$

Assumiamo che la funzione di ingresso $y_b \in C([0,T],\mathbb{R})$ sia nota, fissata e indipendente dai parametri cinetici.

Per ogni $\mathbf{k} \in (0,+\infty)^5$, la soluzione del problema diretto è quindi

$$\mathbf{C}(t;\mathbf{k}) = \int_0^t e^{(t-\tau)M(\mathbf{k})} B(\mathbf{k}) y_b(\tau)\, d\tau. \tag{4.6}$$

Possiamo pertanto definire l'operatore diretto

$$\mathcal{F} : (0,+\infty)^5 \longrightarrow C^1([0,T],\mathbb{R}) \tag{4.7}$$

mediante

$$\mathcal{F}(\mathbf{k})(t) = y_c(t;\mathbf{k}) = \alpha^T \int_0^t e^{(t-\tau)M(\mathbf{k})} B(\mathbf{k}) y_b(\tau)\, d\tau. \tag{4.8}$$

L'operatore $\mathcal{F}$ associa quindi a ogni vettore di parametri ammissibile la curva cerebellare prevista dal modello.

Benché il sistema differenziale sia lineare rispetto alle variabili di stato, l'uscita dipende in generale in modo non lineare dai coefficienti cinetici. Questi compaiono infatti sia nella matrice $M(\mathbf{k})$, e quindi nell'esponenziale di matrice, sia nel vettore $B(\mathbf{k})$.

4.2.2 Dati sperimentali e problema discreto

In un'acquisizione PET dinamica non è disponibile l'intera funzione $y_c(t)$ in modo continuo, ma un numero finito di osservazioni. Indichiamo con

$$t_1, \ldots, t_N$$

i tempi rappresentativi delle N acquisizioni e con

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} \in \mathbb{R}^N \tag{4.9}$$

il vettore dei dati cerebellari osservati.

Nel seguito utilizziamo, per semplicità, il valore della curva modellistica nei tempi t_i. Definiamo quindi

$$F : (0, +\infty)^5 \longrightarrow \mathbb{R}^N \tag{4.10}$$

mediante

$$F(\mathbf{k}) = \begin{pmatrix} y_c(t_1; \mathbf{k}) \\ \vdots \\ y_c(t_N; \mathbf{k}) \end{pmatrix}. \tag{4.11}$$

Nella formulazione adottata, ciascuna osservazione viene approssimata mediante il valore della curva modellistica in un istante rappresentativo del corrispondente frame temporale.

Una descrizione più fedele dell'acquisizione PET richiederebbe di considerare il valore medio della curva nell'intervallo temporale associato a ciascun frame, sostituendo la valutazione istantanea con la corrispondente media temporale.

Seguendo la formulazione standard dei problemi di fitting [9], definiamo il vettore dei residui come

$$\mathbf{r}(\mathbf{k}) = F(\mathbf{k}) - \mathbf{y}. \tag{4.12}$$

La sua i-esima componente è quindi

$$r_i(\mathbf{k}) = y_c(t_i; \mathbf{k}) - y_i. \tag{4.13}$$

Il residuo misura la discrepanza tra la previsione del modello e il dato sperimentale.

4.2.3 La funzione obiettivo

Nel caso ideale, supponendo che i dati siano generati esattamente dal modello, si cercherebbe un vettore $\mathbf{k}$ tale che

$$F(\mathbf{k}) = \mathbf{y}, \tag{4.14}$$

equivalentemente, $\mathbf{r}(\mathbf{k}) = \mathbf{0}$.

In presenza di errori sperimentali, non si richiede che il modello riproduca esattamente tutte le osservazioni. Si cercano invece i parametri che minimizzano la discrepanza tra le previsioni del modello e i dati osservati.

Si introduce pertanto la funzione obiettivo

$$\Phi(\mathbf{k}) = \frac{1}{2}\|\mathbf{r}(\mathbf{k})\|_2^2, \tag{4.15}$$

Possiamo esplicitare la funzione obiettivo in questo modo:

$$\Phi(\mathbf{k}) = \frac{1}{2}\sum_{i=1}^{N}\left[y_c(t_i;\mathbf{k}) - y_i\right]^2. \tag{4.16}$$

La stima dei coefficienti cinetici viene quindi ricondotta al problema della ricerca di

$$\widehat{\mathbf{k}} \in \underset{\mathbf{k}\in(0,+\infty)^5}{\arg\min}\ \Phi(\mathbf{k}). \tag{4.17}$$

Qualora il minimo venga raggiunto, indichiamo con $\widehat{\mathbf{k}}$ un vettore minimizzante, ossia

$$\widehat{\mathbf{k}} \in \underset{\mathbf{k}\in(0,+\infty)^5}{\arg\min}\ \Phi(\mathbf{k}). \tag{4.18}$$

Si tratta di un problema di minimi quadrati non lineari, poiché le componenti del vettore dei residui dipendono in modo non lineare da $\mathbf{k}$.

4.2.4 Parametrizzazione logaritmica dei coefficienti

Il problema di minimizzazione è stato formulato sull'insieme $(0,+\infty)^5$. La positività dei coefficienti cinetici costituisce quindi un vincolo che deve essere preservato durante la procedura iterativa.

Per incorporare direttamente tale vincolo introduciamo una parametrizzazione logaritmica. Poniamo

$$\theta_1 = \log k_{mp}, \qquad \theta_2 = \log k_{fp}, \qquad \theta_3 = \log k_{pf}, \qquad \theta_4 = \log k_{om}, \qquad \theta_5 = \log k_{mo}. \tag{4.19}$$

Definiamo quindi la mappa

$$E : \mathbb{R}^5 \longrightarrow (0,+\infty)^5, \qquad E(\boldsymbol{\theta}) = \begin{pmatrix} e^{\theta_1} \\ e^{\theta_2} \\ e^{\theta_3} \\ e^{\theta_4} \\ e^{\theta_5} \end{pmatrix}. \tag{4.20}$$

La mappa E è biunivoca e permette di rappresentare ogni vettore di coefficienti positivi mediante un unico $\boldsymbol{\theta} \in \mathbb{R}^5$.

Introduciamo pertanto l'operatore discreto riparametrizzato

$$\widetilde{F}(\boldsymbol{\theta}) = F\big(E(\boldsymbol{\theta})\big) \tag{4.21}$$

e il corrispondente vettore dei residui

$$\widetilde{\mathbf{r}}(\boldsymbol{\theta}) = \widetilde{F}(\boldsymbol{\theta}) - \mathbf{y} = \mathbf{r}\big(E(\boldsymbol{\theta})\big). \tag{4.22}$$

La funzione obiettivo diventa

$$\widetilde{\Phi}(\boldsymbol{\theta}) = \frac{1}{2}\,\|\widetilde{\mathbf{r}}(\boldsymbol{\theta})\|_2^2 = \Phi\big(E(\boldsymbol{\theta})\big). \tag{4.23}$$

Poiché E è biunivoca, il problema di minimizzazione può essere riformulato equivalentemente come

$$\inf_{\boldsymbol{\theta}\in\mathbb{R}^5} \widetilde{\Phi}(\boldsymbol{\theta}) = \inf_{\mathbf{k}\in(0,+\infty)^5} \Phi(\mathbf{k}). \tag{4.24}$$

Qualora il minimo venga raggiunto, indicando con $\widehat{\boldsymbol{\theta}}$ un minimizzatore, la corrispondente stima dei coefficienti cinetici è

$$\widehat{\mathbf{k}} = E(\widehat{\boldsymbol{\theta}}). \tag{4.25}$$

La parametrizzazione logaritmica permette quindi di soddisfare automaticamente il vincolo di positività dei coefficienti e di formulare il problema di ottimizzazione nello spazio non vincolato $\mathbb{R}^5$.

Essa non garantisce, tuttavia, l'esistenza di un minimo né le altre ipotesi richieste dai risultati di convergenza, quali, ad esempio, la limitatezza degli insiemi di livello.

4.3 Derivate della funzione obiettivo

Una delle caratteristiche fondamentali dei problemi di minimi quadrati è che gradiente e matrice Hessiana della funzione obiettivo possiedono una struttura particolare [9].

Definiamo innanzitutto la matrice Jacobiana del vettore dei residui rispetto ai coefficienti cinetici fisici,

$$J_{\mathbf{k}}(\mathbf{k}) = \begin{pmatrix} \dfrac{\partial r_1}{\partial k_{mp}} & \dfrac{\partial r_1}{\partial k_{fp}} & \dfrac{\partial r_1}{\partial k_{pf}} & \dfrac{\partial r_1}{\partial k_{om}} & \dfrac{\partial r_1}{\partial k_{mo}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \dfrac{\partial r_N}{\partial k_{mp}} & \dfrac{\partial r_N}{\partial k_{fp}} & \dfrac{\partial r_N}{\partial k_{pf}} & \dfrac{\partial r_N}{\partial k_{om}} & \dfrac{\partial r_N}{\partial k_{mo}} \end{pmatrix} \in \mathbb{R}^{N\times 5}. \tag{4.26}$$

Per una perturbazione $\mathbf{p} \in \mathbb{R}^5$ dei coefficienti fisici si ha, sviluppando al primo

ordine,

$$\mathbf{r}(\mathbf{k}+\mathbf{p}) = \mathbf{r}(\mathbf{k}) + J_{\mathbf{k}}(\mathbf{k})\mathbf{p} + o(\|\mathbf{p}\|). \tag{4.27}$$

Nella formulazione numerica introdotta nella sezione precedente, tuttavia, le variabili di ottimizzazione sono $\boldsymbol{\theta} \in \mathbb{R}^5$. La matrice utilizzata dai metodi iterativi sarà pertanto la Jacobiana del vettore dei residui riparametrizzato rispetto a tali variabili.

4.3.1 Calcolo della matrice Jacobiana

Nel problema considerato, ciascun residuo dipende dai cinque coefficienti cinetici attraverso la soluzione del problema diretto. Per calcolare le derivate dei residui rispetto ai parametri utilizziamo la tecnica delle equazioni di sensibilità [2].

Per ciascun parametro k_j introduciamo il vettore di sensibilità

$$\mathbf{S}_j(t) = \frac{\partial \mathbf{C}(t;\mathbf{k})}{\partial k_j}, \tag{4.28}$$

dove $j \in \{mp, fp, pf, om, mo\}$.

Derivando rispetto a k_j il sistema differenziale soddisfatto da $\mathbf{C}$, otteniamo

$$\frac{\partial}{\partial k_j}\dot{\mathbf{C}}(t;\mathbf{k}) = \frac{\partial}{\partial k_j}\left[M(\mathbf{k})\mathbf{C}(t;\mathbf{k}) + B(\mathbf{k})y_b(t)\right] \tag{4.29}$$

$$= \frac{\partial M}{\partial k_j}\mathbf{C}(t;\mathbf{k}) + M(\mathbf{k})\frac{\partial \mathbf{C}(t;\mathbf{k})}{\partial k_j} + \frac{\partial B}{\partial k_j}y_b(t), \tag{4.30}$$

poiché la funzione di ingresso $y_b(t)$ viene considerata nota e indipendente dai parametri cinetici.

Introducendo il termine non omogeneo

$$\mathbf{f}_j(t) = \frac{\partial M}{\partial k_j}\mathbf{C}(t;\mathbf{k}) + \frac{\partial B}{\partial k_j}y_b(t), \tag{4.31}$$

il vettore di sensibilità soddisfa il problema di Cauchy

$$\begin{cases} \dot{\mathbf{S}}_j(t) = M(\mathbf{k})\mathbf{S}_j(t) + \mathbf{f}_j(t), \\ \mathbf{S}_j(0) = \mathbf{0}. \end{cases} \tag{4.32}$$

La condizione iniziale segue dal fatto che $\mathbf{C}(0;\mathbf{k}) = \mathbf{0}$ non dipende dai coefficienti cinetici.

Ricordando che

$$\mathbf{C}(t) = \begin{pmatrix} C_m(t) \\ C_f(t) \\ C_o(t) \end{pmatrix},$$

e utilizzando le espressioni delle matrici $M(\mathbf{k})$ e $B(\mathbf{k})$, si ottengono i cinque termini non omogenei:

$$\mathbf{f}_{mp}(t) = \begin{pmatrix} y_b(t) - C_m(t) \\ 0 \\ 0 \end{pmatrix}, \qquad \mathbf{f}_{fp}(t) = \begin{pmatrix} 0 \\ y_b(t) - C_m(t) \\ 0 \end{pmatrix}, \tag{4.33}$$

$$\mathbf{f}_{pf}(t) = \begin{pmatrix} 0 \\ -C_f(t) \\ 0 \end{pmatrix}, \qquad \mathbf{f}_{om}(t) = \begin{pmatrix} -C_m(t) \\ 0 \\ C_m(t) \end{pmatrix}, \tag{4.34}$$

$$\mathbf{f}_{mo}(t) = \begin{pmatrix} C_o(t) \\ 0 \\ -C_o(t) \end{pmatrix}. \tag{4.35}$$

I cinque problemi di sensibilità sono quindi sistemi differenziali lineari non omogenei. Una volta nota la soluzione del problema diretto, i termini $\mathbf{f}_j(t)$ risultano noti e i sistemi possono essere risolti per ottenere le derivate delle variabili di stato rispetto ai coefficienti cinetici.

Osserviamo inoltre che le equazioni relative a C_m e C_o non dipendono dai parametri k_{fp} e k_{pf}.

Di conseguenza,

$$S_{m,fp}(t) = S_{o,fp}(t) = 0, \qquad S_{m,pf}(t) = S_{o,pf}(t) = 0. \tag{4.36}$$

I problemi di sensibilità associati a questi due parametri si riducono pertanto alle equazioni scalari

$$\begin{cases} \dot{S}_{f,fp} = -k_{pf} S_{f,fp} + y_b - C_m, \\ S_{f,fp}(0) = 0, \end{cases} \tag{4.37}$$

e

$$\begin{cases} \dot{S}_{f,pf} = -k_{pf} S_{f,pf} - C_f, \\ S_{f,pf}(0) = 0. \end{cases} \tag{4.38}$$

Ricordando che

$$y_c(t;\mathbf{k}) = \alpha^T \mathbf{C}(t;\mathbf{k}),$$

per ciascun parametro si ha

$$\frac{\partial y_c(t;\mathbf{k})}{\partial k_j} = \alpha^T \mathbf{S}_j(t) = S_{f,j}(t) + S_{o,j}(t). \tag{4.39}$$

Poiché i dati sperimentali y_i non dipendono dai parametri, gli elementi della Jacobiana rispetto ai coefficienti fisici sono

$$\boxed{[J_{\mathbf{k}}(\mathbf{k})]_{ij} = S_{f,j}(t_i) + S_{o,j}(t_i).} \tag{4.40}$$

La matrice $J_{\mathbf{k}}(\mathbf{k})$ appartiene quindi a $\mathbb{R}^{N\times 5}$.

Poiché il problema numerico viene formulato nelle variabili logaritmiche $\boldsymbol{\theta}$, la Jacobiana effettivamente utilizzata si ottiene applicando la regola della catena:

$$J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = J_{\mathbf{k}}(\mathbf{k}) \frac{\partial \mathbf{k}}{\partial \boldsymbol{\theta}}. \tag{4.41}$$

Essendo

$$\frac{\partial \mathbf{k}}{\partial \boldsymbol{\theta}} = \operatorname{diag}(k_{mp}, k_{fp}, k_{pf}, k_{om}, k_{mo}), \tag{4.42}$$

si ottiene

$$\boxed{J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = J_{\mathbf{k}}(\mathbf{k}) \operatorname{diag}(k_{mp}, k_{fp}, k_{pf}, k_{om}, k_{mo}).} \tag{4.43}$$

In particolare, la sua componente (i, j) è

$$[J_{\boldsymbol{\theta}}(\boldsymbol{\theta})]_{ij} = k_j \left[S_{f,j}(t_i) + S_{o,j}(t_i)\right]. \tag{4.44}$$

Poiché tutti i coefficienti cinetici sono strettamente positivi, la matrice diagonale nella (4.43) è invertibile.

Pertanto

$$\operatorname{rank} J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = \operatorname{rank} J_{\mathbf{k}}(\mathbf{k}). \tag{4.45}$$

La riparametrizzazione logaritmica preserva quindi il rango puntuale della matrice Jacobiana.

4.3.2 Rango della matrice Jacobiana

Nell'analisi del metodo di Gauss-Newton assumeremo, sotto opportune ipotesi, che la matrice Jacobiana abbia rango pieno rispetto alle colonne.

Poiché $J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \in \mathbb{R}^{N\times 5}$, tale condizione richiede $N \geq 5$ e

$$\operatorname{rank} J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = 5. \tag{4.46}$$

Come mostrato nella sottosezione precedente, la riparametrizzazione logaritmica preserva il rango puntuale della matrice Jacobiana.

La proprietà di rango pieno dipende non soltanto dai parametri cinetici, ma anche dalla funzione di ingresso $y_b(t)$ e dagli istanti di osservazione t_i.

Per caratterizzare questa proprietà introduciamo i valori singolari della matrice Jacobiana [9], ordinati come

$$\sigma_1(J_{\boldsymbol{\theta}}) \geq \cdots \geq \sigma_5(J_{\boldsymbol{\theta}}) \geq 0. \tag{4.47}$$

Indicando con $\sigma_{\min}(J_{\boldsymbol{\theta}}) = \sigma_5(J_{\boldsymbol{\theta}})$ il più piccolo valore singolare, si ha

$$\operatorname{rank} J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) = 5 \quad \Longleftrightarrow \quad \sigma_{\min}\left(J_{\boldsymbol{\theta}}(\boldsymbol{\theta})\right) > 0. \tag{4.48}$$

Questa proprietà deve essere distinta dall'identificabilità strutturale studiata nel capitolo precedente. Quest'ultima riguarda l'intera relazione ingresso-uscita ideale, mentre la Jacobiana del problema numerico viene costruita a partire da un numero finito di osservazioni.

L'identificabilità strutturale non garantisce quindi automaticamente il rango pieno della Jacobiana per una qualunque scelta degli istanti di campionamento.

Nell'analisi di convergenza del metodo di Gauss-Newton è inoltre utile considerare una condizione di rango pieno uniforme.

Dato un opportuno insieme $\mathcal{D} \subseteq \mathbb{R}^5$, si richiede l'esistenza di una costante $\gamma > 0$ tale che

$$\sigma_{\min}\left(J_{\boldsymbol{\theta}}(\boldsymbol{\theta})\right) \geq \gamma, \qquad \forall \boldsymbol{\theta} \in \mathcal{D}. \tag{4.49}$$

Questa condizione garantisce che la Jacobiana rimanga uniformemente lontana dalla perdita di rango nell'insieme considerato.

Valori molto piccoli del minimo valore singolare indicano la presenza di direzioni parametriche alle quali le osservazioni sono poco sensibili. Il condizionamento relativo della Jacobiana può essere valutato mediante il rapporto tra il massimo e il minimo valore singolare.

Infine, osserviamo che la parametrizzazione logaritmica preserva il rango della Jacobiana, ma non ne preserva necessariamente i valori singolari. Le proprietà numeriche vengono pertanto studiate direttamente sulla matrice $J_{\boldsymbol{\theta}}$.

4.3.3 Gradiente

Poiché

$$\widetilde{\Phi}(\boldsymbol{\theta}) = \frac{1}{2}\sum_{i=1}^{N}\widetilde{r}_i(\boldsymbol{\theta})^2, \tag{4.50}$$

derivando rispetto ai cinque parametri si ottiene

$$\nabla\widetilde{\Phi}(\boldsymbol{\theta}) = \sum_{i=1}^{N}\widetilde{r}_i(\boldsymbol{\theta})\nabla\widetilde{r}_i(\boldsymbol{\theta}). \tag{4.51}$$

Questa espressione può essere scritta in forma matriciale come

$$\nabla\widetilde{\Phi}(\boldsymbol{\theta}) = J_{\boldsymbol{\theta}}(\boldsymbol{\theta})^T\widetilde{\mathbf{r}}(\boldsymbol{\theta}). \tag{4.52}$$

4.3.4 Matrice Hessiana

Derivando ancora, mediante la regola di derivazione del prodotto, si ottiene

$$\nabla^2\widetilde{\Phi}(\boldsymbol{\theta}) = J_{\boldsymbol{\theta}}(\boldsymbol{\theta})^T J_{\boldsymbol{\theta}}(\boldsymbol{\theta}) + \sum_{i=1}^{N}\widetilde{r}_i(\boldsymbol{\theta})\nabla^2\widetilde{r}_i(\boldsymbol{\theta}). \tag{4.53}$$

La matrice Hessiana è quindi costituita da due contributi.

Il primo, $J_{\boldsymbol{\theta}}^T J_{\boldsymbol{\theta}}$, dipende soltanto dalle derivate prime dei residui ed è sempre semidefinito positivo. Quando la Jacobiana ha rango pieno rispetto alle colonne, tale matrice è definita positiva.

Il secondo contributo richiede invece la conoscenza delle derivate seconde dei residui e non è necessariamente semidefinito positivo.

Questa decomposizione costituisce la caratteristica fondamentale sfruttata dai metodi specifici per i minimi quadrati non lineari.

Quando i residui sono piccoli e le loro derivate seconde rimangono limitate, oppure quando le derivate seconde sono sufficientemente piccole, il secondo contributo può risultare trascurabile.

In tali condizioni, la matrice $J_{\boldsymbol{\theta}}^T J_{\boldsymbol{\theta}}$ costituisce un'approssimazione della Hessiana, come verrà sfruttato dal metodo di Gauss-Newton.

4.4 Il metodo di Gauss-Newton

Il metodo di Gauss-Newton sfrutta la particolare struttura della matrice Hessiana nei problemi di minimi quadrati non lineari.

Il metodo di Newton applicato a $\widetilde{\Phi}$ richiederebbe, alla ℓ-esima iterazione, la soluzione del sistema

$$\nabla^2\widetilde{\Phi}(\boldsymbol{\theta}^{(\ell)})\mathbf{p}^{(\ell)} = -\nabla\widetilde{\Phi}(\boldsymbol{\theta}^{(\ell)}), \tag{4.54}$$

supponendo che la matrice Hessiana sia invertibile.

Il metodo di Gauss-Newton sostituisce l'Hessiana completa con l'approssimazione

$$\nabla^2\widetilde{\Phi}(\boldsymbol{\theta}) \simeq J_{\boldsymbol{\theta}}(\boldsymbol{\theta})^T J_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \tag{4.55}$$

trascurando il termine che coinvolge le derivate seconde dei residui.

Alla ℓ-esima iterazione disponiamo di una stima corrente dei parametri $\boldsymbol{\theta}^{(\ell)}$. Ponendo

$$\mathbf{r}^{(\ell)} = \widetilde{\mathbf{r}}\left(\boldsymbol{\theta}^{(\ell)}\right), \qquad J^{(\ell)} = J_{\boldsymbol{\theta}}\left(\boldsymbol{\theta}^{(\ell)}\right), \tag{4.56}$$

il metodo determina una correzione $\mathbf{p}_{\mathrm{GN}}^{(\ell)} \in \mathbb{R}^5$ attraverso le equazioni normali

$$\left(J^{(\ell)}\right)^T J^{(\ell)}\mathbf{p}_{\mathrm{GN}}^{(\ell)} = -\left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)}. \tag{4.57}$$

Un'interpretazione equivalente si ottiene linearizzando il vettore dei residui in un intorno del punto corrente:

$$\widetilde{\mathbf{r}}\left(\boldsymbol{\theta}^{(\ell)} + \mathbf{p}\right) \simeq \mathbf{r}^{(\ell)} + J^{(\ell)}\mathbf{p}. \tag{4.58}$$

La correzione di Gauss-Newton viene quindi scelta come soluzione del problema lineare ai minimi quadrati

$$\min_{\mathbf{p}\in\mathbb{R}^5} \frac{1}{2}\left\|\mathbf{r}^{(\ell)} + J^{(\ell)}\mathbf{p}\right\|_2^2. \tag{4.59}$$

Le equazioni normali associate a questo problema coincidono con (4.57). Pertanto, a ogni iterazione, il problema non lineare originario viene approssimato localmente mediante un problema lineare ai minimi quadrati.

La nuova stima dei parametri logaritmici viene ottenuta mediante

$$\boldsymbol{\theta}^{(\ell+1)} = \boldsymbol{\theta}^{(\ell)} + \alpha_\ell \mathbf{p}_{\mathrm{GN}}^{(\ell)}, \tag{4.60}$$

dove $\alpha_\ell > 0$ rappresenta la lunghezza del passo.

La direzione $\mathbf{p}_{\mathrm{GN}}^{(\ell)}$ stabilisce quindi come modificare i parametri, mentre α_ℓ determina l'ampiezza dell'aggiornamento.

I coefficienti cinetici fisici vengono infine ricostruiti mediante

$$\mathbf{k}^{(\ell+1)} = E\left(\boldsymbol{\theta}^{(\ell+1)}\right). \tag{4.61}$$

Grazie alla parametrizzazione logaritmica, ogni iterazione produce automaticamente coefficienti strettamente positivi.

4.4.1 Direzione di discesa

Supponiamo che $J^{(\ell)}$ abbia rango pieno rispetto alle colonne, ossia

$$\operatorname{rank} J^{(\ell)} = 5. \tag{4.62}$$

In questo caso la matrice $\left(J^{(\ell)}\right)^T J^{(\ell)}$ è definita positiva e il passo di Gauss-Newton è univocamente determinato.

Inoltre, utilizzando (4.57), si ottiene

$$\left(\mathbf{p}_{\mathrm{GN}}^{(\ell)}\right)^T \nabla\widetilde{\Phi}\left(\boldsymbol{\theta}^{(\ell)}\right) = \left(\mathbf{p}_{\mathrm{GN}}^{(\ell)}\right)^T \left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)} \tag{4.63}$$

$$= -\left(\mathbf{p}_{\mathrm{GN}}^{(\ell)}\right)^T \left(J^{(\ell)}\right)^T J^{(\ell)} \mathbf{p}_{\mathrm{GN}}^{(\ell)} \tag{4.64}$$

$$= -\left\| J^{(\ell)} \mathbf{p}_{\mathrm{GN}}^{(\ell)} \right\|_2^2. \tag{4.65}$$

Se il punto corrente non è stazionario, il passo di Gauss-Newton è non nullo. Poiché $J^{(\ell)}$ ha rango pieno, si ha quindi

$$\left(\mathbf{p}_{\mathrm{GN}}^{(\ell)}\right)^T \nabla\widetilde{\Phi}\left(\boldsymbol{\theta}^{(\ell)}\right) < 0. \tag{4.66}$$

La direzione di Gauss-Newton è pertanto una direzione di discesa per la funzione obiettivo.

4.4.2 Scelta della lunghezza del passo

Una volta determinata la direzione $\mathbf{p}_{\mathrm{GN}}^{(\ell)}$, occorre scegliere la lunghezza del passo $\alpha_\ell > 0$.

Tale scelta deve garantire una sufficiente riduzione della funzione obiettivo, senza richiedere la determinazione esatta del minimo lungo la direzione di ricerca.

Seguendo Nocedal e Wright [9], la lunghezza del passo può essere scelta in modo da soddisfare le condizioni di Wolfe.

Date due costanti c_1 e c_2 tali che

$$0 < c_1 < c_2 < 1,$$

la prima condizione richiede

$$\begin{aligned} \widetilde{\Phi}\left(\boldsymbol{\theta}^{(\ell)} + \alpha_\ell \mathbf{p}_{\mathrm{GN}}^{(\ell)}\right) \leq \widetilde{\Phi}\left(\boldsymbol{\theta}^{(\ell)}\right) \\ + c_1 \alpha_\ell \nabla \widetilde{\Phi}\left(\boldsymbol{\theta}^{(\ell)}\right)^T \mathbf{p}_{\mathrm{GN}}^{(\ell)}. \end{aligned} \tag{4.67}$$

Questa condizione, detta di diminuzione sufficiente, impone che il nuovo valore della funzione obiettivo presenti una riduzione adeguata rispetto al punto corrente.

La seconda condizione è

$$\begin{aligned} \nabla \widetilde{\Phi}\left(\boldsymbol{\theta}^{(\ell)} + \alpha_\ell \mathbf{p}_{\mathrm{GN}}^{(\ell)}\right)^T \mathbf{p}_{\mathrm{GN}}^{(\ell)} \\ \geq c_2 \nabla \widetilde{\Phi}\left(\boldsymbol{\theta}^{(\ell)}\right)^T \mathbf{p}_{\mathrm{GN}}^{(\ell)}. \end{aligned} \tag{4.68}$$

Essa costituisce una condizione sulla variazione della derivata direzionale della funzione obiettivo e impedisce, in particolare, di accettare passi eccessivamente piccoli.

Le condizioni (4.67) e (4.68) permettono quindi di scegliere una lunghezza del passo che garantisca un'adeguata diminuzione della funzione obiettivo e un sufficiente avanzamento lungo la direzione di ricerca.

4.4.3 Convergenza del metodo di Gauss-Newton

Nocedal e Wright [9] forniscono un risultato di convergenza per il metodo di Gauss-Newton sotto opportune ipotesi di regolarità.

Consideriamo l'insieme di livello

$$\mathcal{L} = \left\{ \boldsymbol{\theta} \in \mathbb{R}^5 : \widetilde{\Phi}(\boldsymbol{\theta}) \leq \widetilde{\Phi}\left(\boldsymbol{\theta}^{(0)}\right) \right\}, \tag{4.69}$$

determinato dal punto iniziale $\boldsymbol{\theta}^{(0)}$.

Supponiamo che $\mathcal{L}$ sia limitato e che ciascuna funzione residuo $\widetilde{r}_i$ sia continuamente differenziabile, con derivata lipschitziana, in un intorno $\mathcal{N}$ di $\mathcal{L}$.

Si richiede inoltre che la matrice Jacobiana soddisfi una condizione di rango pieno uniforme: deve esistere una costante $\gamma > 0$ tale che

$$\|J_{\boldsymbol{\theta}}(\boldsymbol{\theta})\mathbf{z}\|_2 \geq \gamma \|\mathbf{z}\|_2, \tag{4.70}$$

per ogni $\boldsymbol{\theta} \in \mathcal{N}$ e per ogni $\mathbf{z} \in \mathbb{R}^5$.

Tale condizione garantisce che il minimo valore singolare della Jacobiana sia uniformemente limitato inferiormente da una costante positiva nella regione considerata.

Se le iterazioni $\boldsymbol{\theta}^{(\ell)}$ sono generate mediante il metodo di Gauss-Newton e le lunghezze del passo α_ℓ soddisfano le condizioni di Wolfe, allora

$$\lim_{\ell \to \infty} \left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)} = \mathbf{0}. \tag{4.71}$$

Poiché

$$\nabla \widetilde{\Phi}(\boldsymbol{\theta}) = J_{\boldsymbol{\theta}}(\boldsymbol{\theta})^T \widetilde{\mathbf{r}}(\boldsymbol{\theta}), \tag{4.72}$$

si ottiene

$$\lim_{\ell \to \infty} \nabla \widetilde{\Phi} \left(\boldsymbol{\theta}^{(\ell)}\right) = \mathbf{0}. \tag{4.73}$$

Il risultato garantisce quindi che, sotto le ipotesi precedenti, la norma del gradiente della funzione obiettivo tenda a zero lungo la successione delle iterazioni.

Ciò non implica necessariamente che la successione dei parametri converga a un minimizzatore locale o globale della funzione obiettivo.

Osserviamo inoltre che, nel problema considerato, la limitatezza degli insiemi di livello e la condizione di rango pieno uniforme costituiscono ipotesi aggiuntive, non automaticamente garantite dalla formulazione del modello diretto.

4.4.4 Limiti del metodo

L'efficacia del metodo di Gauss-Newton dipende dalla qualità dell'approssimazione della Hessiana mediante la matrice

$$J_{\boldsymbol{\theta}}(\boldsymbol{\theta})^T J_{\boldsymbol{\theta}}(\boldsymbol{\theta}). \tag{4.74}$$

Come osservato nella sezione precedente, tale approssimazione risulta particolarmente accurata quando il contributo delle derivate seconde dei residui è sufficientemente piccolo.

Questo può verificarsi, in particolare, quando i residui sono piccoli e le loro derivate seconde rimangono limitate, oppure quando le funzioni residuo sono prossime a funzioni affini rispetto ai parametri.

Un'ulteriore difficoltà si presenta quando la matrice Jacobiana non ha rango pieno. In questo caso $J_{\boldsymbol{\theta}}^T J_{\boldsymbol{\theta}}$ è singolare e le equazioni normali non determinano univocamente la direzione di ricerca.

Il problema lineare ai minimi quadrati associato ammette ancora soluzione, ma possono esistere più direzioni di Gauss-Newton. Inoltre, il risultato di convergenza precedentemente descritto non è più garantito.

Anche quando la Jacobiana ha rango pieno, una situazione di forte mal condizionamento può rendere problematica la determinazione numerica del passo.

Infatti, il numero di condizionamento della matrice delle equazioni normali soddisfa

$$\kappa_2\left(J_{\boldsymbol{\theta}}^T J_{\boldsymbol{\theta}}\right) = \kappa_2(J_{\boldsymbol{\theta}})^2. \tag{4.75}$$

Per questo motivo, nella risoluzione numerica del problema lineare ai minimi quadrati, può essere preferibile utilizzare fattorizzazioni QR oppure SVD, evitando di formare esplicitamente la matrice $J_{\boldsymbol{\theta}}^T J_{\boldsymbol{\theta}}$.

Le difficoltà associate alla perdita di rango o al mal condizionamento della Jacobiana motivano l'introduzione del metodo di Levenberg-Marquardt.

Nella formulazione basata su una regione di confidenza, che presenteremo nella sezione successiva, tale metodo modifica le equazioni di Gauss-Newton mediante un termine di regolarizzazione, permettendo di controllare l'ampiezza della correzione dei parametri.

4.5 Il metodo di Levenberg-Marquardt

Seguendo l'impostazione di Nocedal e Wright [9], il metodo di Levenberg-Marquardt utilizza la stessa approssimazione della Hessiana adottata da Gauss-Newton,

$$\nabla^2 \widetilde{\Phi}(\boldsymbol{\theta}) \simeq J_{\boldsymbol{\theta}}(\boldsymbol{\theta})^T J_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \tag{4.76}$$

ma controlla l'ampiezza della correzione dei parametri attraverso una strategia basata su una regione di confidenza.

4.5.1 La regione di confidenza

Il metodo di Gauss-Newton utilizza un'approssimazione lineare del vettore dei residui nell'intorno del punto corrente $\boldsymbol{\theta}^{(\ell)}$. Tale approssimazione può perdere accuratezza quando ci si allontana eccessivamente dal punto in cui viene costruita.

Il metodo di Levenberg-Marquardt considera pertanto passi $\mathbf{p} \in \mathbb{R}^5$ appartenenti alla regione

$$\left\{\mathbf{p} \in \mathbb{R}^5 : \|\mathbf{p}\|_2 \leq \Delta_\ell\right\}, \tag{4.77}$$

dove $\Delta_\ell > 0$ è il raggio della regione di confidenza.

Indicando, come nella sezione precedente, con $\mathbf{r}^{(\ell)}$ il vettore dei residui e con $J^{(\ell)}$ la Jacobiana calcolati nel punto corrente, introduciamo il modello locale

$$m_\ell(\mathbf{p}) = \frac{1}{2}\left\|\mathbf{r}^{(\ell)} + J^{(\ell)}\mathbf{p}\right\|_2^2. \tag{4.78}$$

Esso deriva dalla linearizzazione del vettore dei residui; in particolare,

$$m_\ell(\mathbf{0}) = \frac{1}{2}\|\mathbf{r}^{(\ell)}\|_2^2 = \widetilde{\Phi}(\boldsymbol{\theta}^{(\ell)}). \tag{4.79}$$

Il passo viene quindi cercato risolvendo il problema

$$\min_{\|\mathbf{p}\|_2 \le \Delta_\ell} m_\ell(\mathbf{p}) = \min_{\|\mathbf{p}\|_2 \le \Delta_\ell} \frac{1}{2}\left\|\mathbf{r}^{(\ell)} + J^{(\ell)}\mathbf{p}\right\|_2^2. \tag{4.80}$$

Il raggio Δ_ℓ viene adattato durante le iterazioni confrontando la riduzione effettiva della funzione obiettivo con quella prevista dal modello locale. Per un passo $\mathbf{p}^{(\ell)}$ che produca una diminuzione prevista strettamente positiva, si definisce

$$\rho_\ell = \frac{\widetilde{\Phi}(\boldsymbol{\theta}^{(\ell)}) - \widetilde{\Phi}(\boldsymbol{\theta}^{(\ell)} + \mathbf{p}^{(\ell)})}{m_\ell(\mathbf{0}) - m_\ell(\mathbf{p}^{(\ell)})}. \tag{4.81}$$

Il numeratore rappresenta la riduzione effettiva, mentre il denominatore rappresenta la riduzione prevista dal modello.

Se ρ_ℓ è prossimo a 1, il modello locale descrive accuratamente il comportamento della funzione obiettivo lungo il passo considerato. Se ρ_ℓ è piccolo o negativo, l'accordo è invece insufficiente e può essere opportuno ridurre la regione di confidenza. Le regole di aggiornamento del raggio e di accettazione del passo verranno precisate nello schema iterativo.

4.5.2 Il parametro di smorzamento

Consideriamo il problema con regione di confidenza (4.80). Per individuare le condizioni di ottimalità, scriviamo il vincolo nella forma

$$\frac{1}{2}\left(\|\mathbf{p}\|_2^2 - \Delta_\ell^2\right) \le 0 \tag{4.82}$$

e introduciamo la funzione lagrangiana

$$\mathcal{L}_\ell(\mathbf{p}, \lambda_\ell) = \frac{1}{2}\left\|\mathbf{r}^{(\ell)} + J^{(\ell)}\mathbf{p}\right\|_2^2 + \frac{\lambda_\ell}{2}\left(\|\mathbf{p}\|_2^2 - \Delta_\ell^2\right), \tag{4.83}$$

dove $\lambda_\ell \ge 0$ è il moltiplicatore associato al vincolo.

La condizione di stazionarietà rispetto a $\mathbf{p}$ fornisce

$$\left(J^{(\ell)}\right)^T \left(\mathbf{r}^{(\ell)} + J^{(\ell)}\mathbf{p}\right) + \lambda_\ell \mathbf{p} = \mathbf{0}, \tag{4.84}$$

ossia

$$\left[\left(J^{(\ell)}\right)^T J^{(\ell)} + \lambda_\ell I\right] \mathbf{p} = -\left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)}. \tag{4.85}$$

Una soluzione $\mathbf{p}_{\mathrm{LM}}^{(\ell)}$ del problema vincolato deve inoltre soddisfare

$$\lambda_\ell \geq 0, \qquad \|\mathbf{p}_{\mathrm{LM}}^{(\ell)}\|_2 \leq \Delta_\ell, \tag{4.86}$$

e la condizione di complementarietà

$$\lambda_\ell \left(\Delta_\ell - \|\mathbf{p}_{\mathrm{LM}}^{(\ell)}\|_2\right) = 0. \tag{4.87}$$

Poiché il modello m_ℓ è una funzione quadratica convessa e la regione di confidenza è convessa, queste condizioni, insieme a (4.85), sono necessarie e sufficienti per l'ottimalità globale del passo nel problema vincolato.

Consideriamo la soluzione di norma euclidea minima delle equazioni di Gauss-Newton, indicata con $\mathbf{p}_{\mathrm{GN}}^{(\ell)}$; questa scelta permette di trattare anche il caso in cui tali equazioni ammettano più soluzioni. Se

$$\|\mathbf{p}_{\mathrm{GN}}^{(\ell)}\|_2 \leq \Delta_\ell, \tag{4.88}$$

questa soluzione è ammissibile e risolve anche il problema con regione di confidenza. In tal caso si può scegliere $\lambda_\ell = 0$, e (4.85) coincide con le equazioni di Gauss-Newton.

Se invece

$$\|\mathbf{p}_{\mathrm{GN}}^{(\ell)}\|_2 > \Delta_\ell, \tag{4.89}$$

nessuna soluzione del problema non vincolato appartiene alla regione di confidenza. Il vincolo è allora attivo e si ha

$$\lambda_\ell > 0, \qquad \|\mathbf{p}_{\mathrm{LM}}^{(\ell)}\|_2 = \Delta_\ell. \tag{4.90}$$

Il valore di λ_ℓ è determinato in relazione al raggio Δ_ℓ, in modo che il passo rispetti il vincolo imposto sulla sua norma.

Per $\lambda_\ell > 0$ la matrice

$$\left(J^{(\ell)}\right)^T J^{(\ell)} + \lambda_\ell I \tag{4.91}$$

è definita positiva anche quando la Jacobiana non ha rango pieno; il sistema (4.85) ammette pertanto un'unica soluzione.

Quando la Jacobiana ha rango pieno, per valori piccoli di λ_ℓ il passo è vicino a quello di Gauss-Newton. Se la Jacobiana perde rango, il limite per $\lambda_\ell \to 0^+$ seleziona invece la soluzione di Gauss-Newton di norma minima. Per valori elevati di λ_ℓ, si ha l'approssimazione

$$\mathbf{p}_{\mathrm{LM}}^{(\ell)} \simeq -\frac{1}{\lambda_\ell}\left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)} = -\frac{1}{\lambda_\ell}\nabla\widetilde{\Phi}(\boldsymbol{\theta}^{(\ell)}), \tag{4.92}$$

che corrisponde a un passo di norma ridotta, orientato approssimativamente nella direzione opposta al gradiente.

4.5.3 Convergenza del metodo di Levenberg-Marquardt

Nocedal e Wright [9], nel Teorema 10.3, forniscono un risultato di convergenza globale per la formulazione del metodo basata su una regione di confidenza. La strategia di aggiornamento del raggio è quella dell'Algoritmo 4.1, descritta nello schema iterativo riportato nel seguito.

Consideriamo l'insieme di livello

$$\mathcal{L} = \left\{\boldsymbol{\theta} \in \mathbb{R}^5 : \widetilde{\Phi}(\boldsymbol{\theta}) \le \widetilde{\Phi}(\boldsymbol{\theta}^{(0)})\right\}. \tag{4.93}$$

Supponiamo che $\mathcal{L}$ sia limitato e che ciascuna funzione residuo $\widetilde{r}_i$, $i = 1, \ldots, N$, sia due volte continuamente differenziabile in un intorno $\mathcal{N}$ di $\mathcal{L}$.

Il problema vincolato non deve essere necessariamente risolto in modo esatto. Si richiede tuttavia che il passo ammissibile $\mathbf{p}^{(\ell)}$, con $\|\mathbf{p}^{(\ell)}\|_2 \le \Delta_\ell$, produca una diminuzione sufficiente del modello locale. Precisamente, deve esistere una costante $c > 0$, indipendente da ℓ, tale che

$$\begin{aligned} &m_\ell(\mathbf{0}) - m_\ell(\mathbf{p}^{(\ell)}) \\ &\quad \ge c \left\|\left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)}\right\|_2 \min\left\{\Delta_\ell, \frac{\left\|\left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)}\right\|_2}{\left\|\left(J^{(\ell)}\right)^T J^{(\ell)}\right\|_2}\right\}. \end{aligned} \tag{4.94}$$

La condizione viene considerata nelle iterazioni non stazionarie; in tali iterazioni il denominatore della frazione è strettamente positivo. Una soluzione esatta del problema con regione di confidenza soddisfa la richiesta di diminuzione sufficiente, così come un passo approssimato che produca almeno la diminuzione ottenibile mediante il punto di Cauchy.

Si consideri inoltre un parametro di accettazione

$$0 < \eta < \frac{1}{4}. \tag{4.95}$$

Se le iterazioni sono generate applicando la strategia sulla regione di confidenza indicata sopra e i passi soddisfano (4.94), allora

$$\lim_{\ell \to \infty} \nabla \widetilde{\Phi}(\boldsymbol{\theta}^{(\ell)}) = \lim_{\ell \to \infty} \left(J^{(\ell)}\right)^T \mathbf{r}^{(\ell)} = \mathbf{0}. \tag{4.96}$$

Il risultato garantisce dunque che la norma del gradiente della funzione obiettivo tenda a zero lungo la successione delle iterazioni, sotto le ipotesi dichiarate. Non garantisce, da solo, che l'intera successione converga a un minimizzatore locale o globale.

4.5.4 Relazione tra convergenza numerica e identificabilità

I risultati di convergenza richiamati nelle sezioni precedenti garantiscono, sotto le rispettive ipotesi, che il gradiente della funzione obiettivo tenda a zero. Pertanto, per continuità del gradiente, ogni eventuale punto limite della successione numerica è un punto stazionario di $\widetilde{\Phi}$. Ciò non implica necessariamente l'annullamento dei residui, né l'individuazione del parametro fisico vero.

L'identificabilità strutturale studiata nel capitolo precedente riguarda l'intera relazione ingresso-uscita in condizioni ideali. Il problema numerico impiega invece un numero finito di osservazioni e richiede quindi un'analisi distinta.

Sia $\mathbf{k}^* \in (0, +\infty)^5$ il vettore dei parametri veri e poniamo

$$\boldsymbol{\theta}^* = E^{-1}(\mathbf{k}^*). \tag{4.97}$$

Supponiamo che i dati siano esatti e generati dal modello, ossia

$$\mathbf{y} = \widetilde{F}(\boldsymbol{\theta}^*). \tag{4.98}$$

Allora

$$\widetilde{\mathbf{r}}(\boldsymbol{\theta}) = \widetilde{F}(\boldsymbol{\theta}) - \widetilde{F}(\boldsymbol{\theta}^*), \tag{4.99}$$

e, di conseguenza,

$$\widetilde{\Phi}(\boldsymbol{\theta}) = 0 \quad \Longleftrightarrow \quad \widetilde{F}(\boldsymbol{\theta}) = \widetilde{F}(\boldsymbol{\theta}^*). \tag{4.100}$$

L'identificabilità strutturale della relazione continua non assicura, da sola, che la mappa discreta $\widetilde{F} : \mathbb{R}^5 \longrightarrow \mathbb{R}^N$ sia iniettiva per una qualunque scelta degli istanti

di osservazione. Possiamo tuttavia ottenere un risultato di unicità locale mediante il rango della Jacobiana.

Supponiamo infatti che

$$\operatorname{rank} J_{\boldsymbol{\theta}}(\boldsymbol{\theta}^*) = 5. \tag{4.101}$$

Poiché la Jacobiana ha cinque colonne, esistono cinque righe linearmente indipendenti. Possiamo dunque scegliere cinque indici distinti $i_1, \dots, i_5$ e definire la mappa $G : \mathbb{R}^5 \to \mathbb{R}^5$ mediante

$$G(\boldsymbol{\theta}) = \begin{pmatrix} y_c(t_{i_1}; E(\boldsymbol{\theta})) \\ y_c(t_{i_2}; E(\boldsymbol{\theta})) \\ y_c(t_{i_3}; E(\boldsymbol{\theta})) \\ y_c(t_{i_4}; E(\boldsymbol{\theta})) \\ y_c(t_{i_5}; E(\boldsymbol{\theta})) \end{pmatrix}. \tag{4.102}$$

La matrice $DG(\boldsymbol{\theta}^*)$ è una sottomatrice quadrata 5×5 non singolare di $J_{\boldsymbol{\theta}}(\boldsymbol{\theta}^*)$.

Per il teorema della funzione inversa esiste un intorno U di $\boldsymbol{\theta}^*$ nel quale G è iniettiva. Se $\boldsymbol{\theta} \in U$ soddisfa

$$\widetilde{F}(\boldsymbol{\theta}) = \widetilde{F}(\boldsymbol{\theta}^*), \tag{4.103}$$

allora anche $G(\boldsymbol{\theta}) = G(\boldsymbol{\theta}^*)$, e quindi $\boldsymbol{\theta} = \boldsymbol{\theta}^*$.

Pertanto, nell'intorno U,

$$\widetilde{\Phi}(\boldsymbol{\theta}) = 0 \quad \Longrightarrow \quad \boldsymbol{\theta} = \boldsymbol{\theta}^*. \tag{4.104}$$

Abbiamo così ottenuto l'unicità locale del vettore di parametri che annulla i residui del problema discreto.

Esaminiamo infine la natura del punto $\boldsymbol{\theta}^*$ rispetto alla funzione obiettivo. Poiché $\widetilde{\mathbf{r}}(\boldsymbol{\theta}^*) = \mathbf{0}$, si ha

$$\nabla \widetilde{\Phi}(\boldsymbol{\theta}^*) = \mathbf{0}. \tag{4.105}$$

Richiamando la decomposizione della Hessiana, si ottiene inoltre

$$\nabla^2 \widetilde{\Phi}(\boldsymbol{\theta}^*) = J_{\boldsymbol{\theta}}(\boldsymbol{\theta}^*)^T J_{\boldsymbol{\theta}}(\boldsymbol{\theta}^*). \tag{4.106}$$

Sotto l'ipotesi (4.101), per ogni $\mathbf{z} \in \mathbb{R}^5$ non nullo vale

$$\begin{aligned} \mathbf{z}^T \nabla^2 \widetilde{\Phi}(\boldsymbol{\theta}^*) \mathbf{z} &= \| J_{\boldsymbol{\theta}}(\boldsymbol{\theta}^*) \mathbf{z} \|_2^2 \\ &> 0. \end{aligned} \tag{4.107}$$

La Hessiana è quindi definita positiva e, per la condizione sufficiente del secondo ordine, $\boldsymbol{\theta}^*$ è un minimo locale stretto di $\widetilde{\Phi}$. Poiché i residui si annullano in questo

punto, esso è anche un minimo globale, non necessariamente l'unico sull'intero spazio $\mathbb{R}^5$.

Identificabilità strutturale, unicità locale per i dati discreti e convergenza del metodo iterativo costituiscono dunque proprietà distinte. Il rango pieno della Jacobiana nel parametro vero assicura l'unicità locale nell'esperimento discreto ideale, mentre la convergenza del gradiente non garantisce che il punto raggiunto coincida con quel parametro.

4.5.5 Schema iterativo

Consideriamo una strategia di Levenberg-Marquardt basata sulla regione di confidenza, coerente con l'Algoritmo 4.1 di Nocedal e Wright [9]. Fissiamo una stima iniziale $\boldsymbol{\theta}^{(0)} \in \mathbb{R}^5$, un raggio massimo $\overline{\Delta} > 0$, un raggio iniziale $\Delta_0 \in (0, \overline{\Delta})$ e un parametro di accettazione $\eta \in (0, 1/4)$.

A ogni iterazione, purché non sia già soddisfatto un criterio di arresto, si procede nel modo seguente:

1. Si calcolano il vettore dei residui e la Jacobiana nel punto corrente:

$$\mathbf{r}^{(\ell)} = \widetilde{\mathbf{r}}(\boldsymbol{\theta}^{(\ell)}), \qquad J^{(\ell)} = J_{\boldsymbol{\theta}}(\boldsymbol{\theta}^{(\ell)}).$$

2. Si determina un passo ammissibile $\mathbf{p}^{(\ell)} \in \mathbb{R}^5$ risolvendo, esattamente o approssimativamente, il problema

$$\min_{\|\mathbf{p}\|_2 \le \Delta_\ell} \frac{1}{2}\|\mathbf{r}^{(\ell)} + J^{(\ell)}\mathbf{p}\|_2^2.$$

   Nel caso approssimato si richiede la diminuzione sufficiente (4.94).

3. Se il gradiente non è nullo, si calcola il rapporto ρ_ℓ definito in (4.81); per i passi considerati la diminuzione prevista è strettamente positiva.

4. Si aggiorna il raggio secondo la regola

$$\Delta_{\ell+1} = \begin{cases} \dfrac{1}{4}\|\mathbf{p}^{(\ell)}\|_2, & \rho_\ell < \dfrac{1}{4}, \\[2ex] \min\{2\Delta_\ell, \overline{\Delta}\}, & \rho_\ell > \dfrac{3}{4}, \ \|\mathbf{p}^{(\ell)}\|_2 = \Delta_\ell \\[2ex] \Delta_\ell, & \text{altrimenti.} \end{cases},$$

5. Il passo viene accettato se $\rho_\ell > \eta$; in caso contrario viene rifiutato. In formule,

$$\boldsymbol{\theta}^{(\ell+1)} = \begin{cases} \boldsymbol{\theta}^{(\ell)} + \mathbf{p}^{(\ell)}, & \rho_\ell > \eta, \\ \boldsymbol{\theta}^{(\ell)}, & \rho_\ell \leq \eta. \end{cases}$$

Il procedimento viene ripetuto fino al soddisfacimento di un opportuno criterio di arresto, basato, per esempio, sulla norma del gradiente, sulla dimensione del passo o sulla variazione della funzione obiettivo.

A ogni iterazione, i coefficienti cinetici fisici si ricavano mediante

$$\mathbf{k}^{(\ell)} = E(\boldsymbol{\theta}^{(\ell)}). \tag{4.108}$$

Poiché le componenti di E sono esponenziali, si ha automaticamente $\mathbf{k}^{(\ell)} \in (0, +\infty)^5$ per ogni ℓ; non sono quindi necessarie procedure di proiezione per garantire la positività dei coefficienti.

Conclusioni

Nel nostro lavoro abbiamo introdotto un modello compartimentale per descrivere la cinetica della FDOPA e del suo metabolita, la 3-OMFD, con particolare riferimento alla regione cerebellare. Abbiamo descritto un sistema dinamico lineare che rappresenta gli scambi tra plasma e cervelletto e la conversione periferica della FDOPA nel suo metabolita, collegando le variabili del modello alle quantità osservabili mediante esame PET.

Una volta formulato il problema diretto, abbiamo affrontato il problema inverso, distinguendo lo studio teorico dell'identificabilità strutturale dalla formulazione numerica della stima dei parametri.

Abbiamo inizialmente considerato il modello generale con cinque coefficienti cinetici indipendenti, soggetti al vincolo di positività. Mediante lo studio della funzione di trasferimento abbiamo individuato le condizioni di irriducibilità della corrispondente rappresentazione razionale e introdotto la mappa che associa ai parametri i coefficienti dei polinomi della funzione di trasferimento.

Lo studio del rango della Jacobiana di tale mappa ha permesso di individuare punti regolari nei quali i parametri sono localmente univocamente determinati dalla funzione di trasferimento. Nei punti in cui il rango viene meno, il criterio differenziale non è invece sufficiente, da solo, a stabilire l'iniettività locale.

Per affrontare il problema globale abbiamo ricondotto il sistema algebrico associato ai coefficienti della funzione di trasferimento a un problema in una sola incognita, scelta pari a k_{pf}. Abbiamo così mostrato che, per un insieme fissato di coefficienti, il numero di vettori positivi compatibili è finito. Lo studio delle radici comuni di due polinomi e del loro risultante ha poi permesso di esaminare i casi di possibile perdita di unicità.

In particolare, abbiamo dimostrato che, al di fuori di un insieme eccezionale di misura di Lebesgue nulla nello spazio dei parametri, la funzione di trasferimento determina univocamente i cinque coefficienti cinetici. Il modello generale è pertanto genericamente globalmente identificabile. Abbiamo inoltre costruito un esempio esplicito di due vettori positivi distinti che producono la stessa funzione di trasferimento,

mostrando che l'iniettività non vale in tutto lo spazio dei parametri.

Successivamente abbiamo introdotto ulteriori relazioni tra i coefficienti cinetici, motivate da risultati presenti in letteratura [5, 7]. Tali ipotesi fisiologiche permettono di ottenere un modello ridotto a tre parametri, ma non sono necessarie per dimostrare l'identificabilità generica del modello generale. Nella formulazione numerica abbiamo invece considerato il modello a cinque parametri indipendenti, senza imporre queste ulteriori relazioni.

L'ultima parte del lavoro è stata dedicata alla formulazione numerica del problema inverso e allo studio di metodi per la stima dei coefficienti cinetici. Abbiamo ricondotto il problema a minimi quadrati non lineari, introducendo il vettore dei residui e la funzione obiettivo. Il vincolo di positività è stato incorporato mediante una riparametrizzazione logaritmica; abbiamo quindi ricavato la Jacobiana del problema attraverso le equazioni di sensibilità.

Infine, abbiamo presentato i metodi di Gauss-Newton e Levenberg-Marquardt, discutendone le principali proprietà e le ipotesi necessarie per i risultati di convergenza. È emersa in particolare la necessità di distinguere l'identificabilità strutturale dalla stabilità e dal comportamento numerico del problema discretizzato: l'unicità strutturale generica in condizioni ideali non garantisce, da sola, che la stima ottenuta da dati discreti e perturbati sia ben condizionata né che un algoritmo individui il parametro vero.

Bibliografia

[1] Dale L. Bailey et al., cur. *Nuclear Medicine Physics: A Handbook for Teachers and Students*. Vienna: International Atomic Energy Agency, 2014. ISBN: 978-92-0-143810-2.

[2] Claudio Cobelli, David Foster e Gianna Toffolo. *Tracer Kinetics in Biomedical Research: From Data to Model*. New York: Kluwer Academic/Plenum Publishers, 2000. ISBN: 0-306-46427-6. DOI: 10.1007/b112199.

[3] Mattia De Francisci et al. «EMATA: a toolbox for the automatic extraction and modeling of arterial inputs for tracer kinetic analysis in [18F]FDG brain studies». In: *EJNMMI Physics* 11 (2024), p. 105. DOI: 10.1186/s40658-024-00707-2.

[4] Paul A. Fuhrmann. *A Polynomial Approach to Linear Algebra*. 2ª ed. New York: Springer, 2012. DOI: 10.1007/978-1-4614-0338-8.

[5] Sung-Cheng Huang et al. «Kinetics and Modeling of L-6-[18F]Fluoro-DOPA in Human Positron Emission Tomographic Studies». In: *Journal of Cerebral Blood Flow and Metabolism* 11 (1991), pp. 898–913. DOI: 10.1038/jcbfm.1991.155.

[6] Steven G. Krantz e Harold R. Parks. *A Primer of Real Analytic Functions*. 2ª ed. Birkhäuser, 2002. DOI: 10.1007/978-0-8176-8134-0.

[7] Yoshitaka Kumakura et al. «PET studies of net blood–brain clearance of FDOPA to human brain: age-dependent decline of [18F]fluorodopamine storage capacity». In: *Journal of Cerebral Blood Flow & Metabolism* 25.7 (2005), pp. 807–819. DOI: 10.1038/sj.jcbfm.9600079.

[8] William P. Melega et al. «L-6-[18F]Fluoro-DOPA Metabolism in Monkeys and Humans: Biochemical Parameters for the Formulation of Tracer Kinetic Models with Positron Emission Tomography». In: *Journal of Cerebral Blood Flow and Metabolism* 11 (1991), pp. 890–897. DOI: 10.1038/jcbfm.1991.154.

[9] Jorge Nocedal e Stephen J. Wright. *Numerical Optimization*. 2ª ed. New York: Springer, 2006. DOI: 10.1007/978-0-387-40065-5.

[10] Christiaan Schiepers et al. «18F-FDOPA Kinetics in Brain Tumors». In: *Journal of Nuclear Medicine* 48 (2007), pp. 1651–1661. DOI: 10.2967/jnumed.106.039321.

[11] Henry N. Jr. Wagner. «A Brief History of Positron Emission Tomography (PET)». In: *Seminars in Nuclear Medicine* 28.3 (1998), pp. 213–220. DOI: 10.1016/S0001-2998(98)80027-5.

[12] Timothée Zaragori et al. «Dynamic 18F-FDopa PET Imaging for Newly Diagnosed Gliomas: Is a Semiquantitative Model Sufficient?» In: *Frontiers in Oncology* 11 (2021), p. 735257. DOI: 10.3389/fonc.2021.735257.