

Università LUISS Guido Carli, A.A. 2010-2011 (II semestre)

Metodi Matematici per la Finanza

Prof. Davide Vergni

Autovalori e autovettori

Indice

1	Autovalori e autovettori	5
1.0.1	Obiettivi principali del capitolo	5
1.1	Sottospazi invarianti	5
1.1.1	Autovalori e autovettori di matrici	7
1.2	Considerazioni su autovalori e autovettori	8
1.2.1	Proprietà degli autovalori	8
1.2.2	Proprietà degli autovettori	10
1.3	Decomposizione spettrale	11
1.3.1	Autovalori reali e distinti	11
1.3.2	Autovalori reali e coincidenti	13
1.3.3	Autovalori complessi coniugati	14
1.3.4	Azione generale degli operatori	15

Capitolo 1

Autovalori e autovettori

1.0.1 Obiettivi principali del capitolo

Gli operatori lineari su spazi vettoriali, in particolare gli endomorfismi di spazi vettoriali, sono uno strumento di lavoro molto potente della modellistica matematica. In questo capitolo approfondiremo lo studio dell'azione degli operatori lineari attraverso l'introduzione di autovalori ed autovettori, grandezze indissolubilmente legate al concetto di sottospazio invariante di un operatore.

Per semplicità ci limiteremo a trattare endomorfismi di $\mathbb{R}^n$, ossia operatori lineari $\phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$. In questo caso, essendo naturalmente presente nello spazio la base canonica, c'è un'associazione immediata e diretta tra operatori lineari, ϕ , e matrici, $\hat{A} \in \mathcal{M}(n, n)$. Quindi in questo capitolo ogni matrice sarà vista come un endomorfismo, tenendo però bene a mente che i principali concetti introdotti saranno assolutamente generali e validi anche per endomorfismi di spazi vettoriali astratti.

1.1 Sottospazi invarianti

Man mano che viene sviluppata la teoria appare sempre più importante il ruolo svolto dai sottospazi vettoriali. Questo perché se si riesce a dimostrare che un certo insieme di interesse (le soluzioni di un'equazione, gli zeri di un operatore o il suo codominio, ecc. ecc.) è un sottospazio vettoriale, allora per la sua determinazione è sufficiente riuscire a trovarne una base attraverso la quale poter rappresentare tutti i suoi elementi. Ed è intuitivo come la ricerca di una base di un sottospazio vettoriale sia molto più semplice e sintetica della ricerca di tutti gli elementi di un insieme generico.

Theorem 1.1.1. *Sia $\phi \equiv \hat{A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un endomorfismo di $\mathbb{R}^n$ e sia $U \subset \mathbb{R}^n$ un sottospazio vettoriale di $\mathbb{R}^n$. Allora, se indichiamo con $U_\phi \subset \mathbb{R}^n$ l'insieme immagine di U attraverso l'operatore $\hat{A}$,*

$$U_\phi = \left\{ \mathbf{v} \in \mathbb{R}^n \text{ tale che } \exists \mathbf{u} \in U \text{ con } \mathbf{v} = \hat{A}\mathbf{u} \right\}$$

(a volte anche indicato, con un leggero abuso di notazione, $U_\phi = \phi(U)$) si ha che U_ϕ è un sottospazio vettoriale.

Dimostrazione. Prendiamo $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{U}$ e una loro qualsiasi combinazione lineare $\mathbf{u} = \alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2 \in \mathbb{U}$. Poniamo $\mathbf{v}_1 = \hat{A}\mathbf{u}_1$ e $\mathbf{v}_2 = \hat{A}\mathbf{u}_2$ e calcoliamo $\hat{A}\mathbf{u}$:

$$\hat{A}\mathbf{u} = \hat{A}(\alpha_1 \mathbf{u}_1 + \alpha_2 \mathbf{u}_2) = \alpha_1 \hat{A}\mathbf{u}_1 + \alpha_2 \hat{A}\mathbf{u}_2 = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2.$$

Ne consegue che ogni combinazione lineare di vettori di $\mathbb{U}_\phi$ (ultimo termine di destra) è ancora un vettore di $\mathbb{U}_\phi$, visto che il primo termine a sinistra è necessariamente un vettore di $\mathbb{U}_\phi$ $\square$

Lavorando con operatori lineari su spazi a molte dimensioni non è semplice intuire e a volte nemmeno calcolare l'azione degli operatori sui vettori dello spazio. Tuttavia esistono particolari sottospazi vettoriali rispetto ai quali l'azione dell'operatore risulta banale.

Definizione 1.1.2. Sia $\phi \equiv \hat{A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ un endomorfismo di $\mathbb{R}^n$ e sia $\mathbb{U} \subset \mathbb{R}^n$ un sottospazio vettoriale di $\mathbb{R}^n$. Diremo che $\mathbb{U}$ è un sottospazio invariante rispetto a ϕ se ϕ trasforma vettori di $\mathbb{U}$ in vettori di $\mathbb{U}$:

$$\forall \mathbf{u} \in \mathbb{U} \rightarrow \phi(\mathbf{u}) \in \mathbb{U} \quad \rightarrow \quad \mathbb{U}_\phi = \mathbb{U}$$

Anche in questo caso, con un leggero abuso di notazione, si può scrivere $\phi(\mathbb{U}) = \mathbb{U}$. Quindi, sostanzialmente, il sottospazio $\mathbb{U}$ viene lasciato invariato dall'azione di ϕ . Un tipo particolarmente importante di sottospazi invarianti sono quelli unidimensionali, per i quali è semplice scrivere una relazione che ne identifichi gli elementi. Infatti, se $\mathbb{U}$ è unidimensionale, ci sarà un unico vettore di base proporzionale a tutti i vettori di $\mathbb{U}$: dato $\mathbf{u} \neq \mathbf{0} \in \mathbb{U}$, allora

$$\forall \mathbf{v} \in \mathbb{U} \exists \alpha \in \mathbb{R} \mid \mathbf{v} = \alpha \mathbf{u}.$$

Ma se $\mathbb{U}_\phi = \mathbb{U}$ significa che per ogni $\mathbf{v} \in \mathbb{U}$ $\phi(\mathbf{v}) \in \mathbb{U}$, ossia

$$\phi(\mathbf{v}) = \lambda \mathbf{v}. \tag{1.1}$$

La costante di proporzionalità, λ , tra i vettori $\mathbf{v} \in \mathbb{U}$ e le loro immagini attraverso ϕ non dipende dal vettore $\mathbf{v}$ scelto:

$$\phi(\mathbf{v}) = \phi(\alpha \mathbf{u}) = \alpha \phi(\mathbf{u}),$$

ma visto che il membro di sinistra è pari a $\lambda \mathbf{v}$ (vedi eq. (1.1) e che $\lambda \mathbf{v} = \lambda \alpha \mathbf{u}$ si ha anche che

$$\phi(\mathbf{u}) = \lambda \mathbf{u}.$$

La quantità λ viene detta **autovalore** ed il vettore $\mathbf{u}$ viene detto **autovettore**. Visto che ϕ è un operatore lineare, qualsiasi $\mathbf{u} \in \mathbb{U}$ può essere scelto come autovettore, basta che sia $\mathbf{u} \neq \mathbf{0}$, e come già mostrato, il valore di λ è indipendente dalla scelta dell'autovettore.

1.1.1 Autovalori e autovettori di matrici

Calcolare autovalori ed autovettori di endomorfismi generici su spazi vettoriali astratti può rivelarsi anche complicato. Ma, nelle ipotesi di apertura di questo capitolo, ossia limitando lo studio agli endomorfismi di $\mathbb{R}^n$, ogni operatore è definito da una matrice $\hat{A} \in \mathcal{M}(n, n)$, e la condizione di esistenza di un sottospazio invariante unidimensionale diventa

$$\hat{A}\mathbf{u} = \lambda\mathbf{u}, \quad (1.2)$$

che è l'equazione alla quale ci si riferisce quando si parla di autovalori ed autovettori di matrici:

Definizione 1.1.3. *Sia $\hat{A} \in \mathcal{M}(n, n)$. Allora se vale l'equazione (1.2) diremo che λ è un autovalore di $\hat{A}$ e $\mathbf{u}$ un autovettore associato a tale autovalore.*

Così come si determinano nucleo e immagine di un operatore matriciale, così esistono dei metodi che permettono di calcolare autovalori ed autovettori di operatori matriciali. L'equazione $\hat{A}\mathbf{u} = \lambda\mathbf{u}$ può essere scritta come $\hat{A}\mathbf{u} - \lambda\mathbf{u} = \mathbf{0}$. Per poter mettere in evidenza $\mathbf{u}$, sfruttiamo l'uguaglianza $\lambda\mathbf{u} = \lambda\hat{J}\mathbf{u}$ e quindi

$$\hat{A}\mathbf{u} = \lambda\mathbf{u} \rightarrow \hat{A}\mathbf{u} - \lambda\hat{J}\mathbf{u} = \mathbf{0} \rightarrow (\hat{A} - \lambda\hat{J})\mathbf{u} = \mathbf{0},$$

che è un'equazione omogenea per la matrice $(\hat{A} - \lambda\hat{J})$. Tale equazione omogenea avrà sempre la soluzione banale, $\mathbf{u} = \mathbf{0}$, alla quale non siamo ovviamente interessati (il sottospazio vettoriale $\mathbf{0}$ è in effetti un sottospazio invariante, però non ci fornisce nessun tipo di informazione sull'operatore), mentre avrà soluzioni non banali se

$$\det(\hat{A} - \lambda\hat{J}) = 0. \quad (1.3)$$

Tale equazione, che bisogna effettivamente risolvere per calcolare gli autovalori, viene detta **equazione caratteristica**. Successivamente, una volta calcolati gli autovalori, questi vengono inseriti nel sistema omogeneo

$$(\hat{A} - \lambda\hat{J})\mathbf{u} = \mathbf{0}, \quad (1.4)$$

le cui soluzioni sono gli autovettori associati all'autovalore λ . Questo sistema avrà sicuramente soluzioni non banali perché abbiamo sostituito proprio quei valori di λ per cui $\det(\hat{A} - \lambda\hat{J}) = 0$.

Esempio 1.1.4. *Calcolare autovalori e autovettori dell'operatore $\hat{A} = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}$.*

Iniziamo con gli autovalori:

$$\det(\hat{A} - \lambda\hat{J}) = \det\left(\begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} -\lambda & 1 \\ 4 & -\lambda \end{pmatrix} = 0.$$

Effettuando il calcolo:

$$\det\begin{pmatrix} -\lambda & 1 \\ 4 & -\lambda \end{pmatrix} = \lambda^2 - 4 = 0 \rightarrow \lambda = \pm 2.$$

Quindi la matrice $\hat{A}$ possiede due autovalori diversi, $\lambda_1 = -2$ e $\lambda_2 = 2$.

Calcoliamo gli autovettori associati ad ognuno di essi. Per quanto riguarda $\mathbf{u}_1$

$$(\hat{A} - \lambda_1 \hat{J})\mathbf{u} = (\hat{A} + 2\hat{J})\mathbf{u} = \left(\begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix} + 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Una prima verifica della correttezza del calcolo degli autovalori l'abbiamo osservando la matrice $(\hat{A} - \lambda_1 \hat{J})$: essa deve avere determinante 0. E si vede a occhio che in tale matrice le due colonne sono proporzionali. Per applicare Rouché-Capelli e risolvere il sistema bisogna trovare una sottomatrice che determini il rango. Scegliamo l'elemento in 1 come sottomatrice di ordine 1 che determina il rango ($\text{Rg}(\hat{A} - \lambda_1 \hat{J}) = 1$) e otteniamo il sistema ridotto

$$2x + y = 0 \rightarrow y = -2x \rightarrow \mathbf{u}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Poi, analogamente per quanto riguarda $\mathbf{u}_2$

$$(\hat{A} - \lambda_2 \hat{J})\mathbf{u} = (\hat{A} - 2\hat{J})\mathbf{u} = \left(\begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Anche in questo caso è semplice osservare che la matrice $(\hat{A} - \lambda_2 \hat{J})$ ha determinante 0, e ancora una volta scegliamo l'elemento in 1 come sottomatrice di ordine 1 che determina il rango ($\text{Rg}(\hat{A} - \lambda_2 \hat{J}) = 1$) e otteniamo il sistema ridotto

$$-2x + y = 0 \rightarrow y = 2x \rightarrow \mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Vale la pena di notare che esiste un'altra verifica importante rispetto al calcolo di autovettori, ossia la prova $\hat{A}\mathbf{v} = \lambda\mathbf{v}$ che viene lasciata per esercizio.

1.2 Considerazioni su autovalori e autovettori

Abbiamo appena visto come la definizione di autovalore e di autovettore la otteniamo a partire dal concetto di sottospazio invariante, mentre le equazioni attraverso le quali questi vengono calcolati sono un'equazione algebrica per quanto riguarda gli autovalori e un sistema omogeneo per quanto riguarda gli autovettori. Nelle prossime sottosezioni, discuteremo il tipo di soluzioni che ci possiamo aspettare dal calcolo di autovalori ed autovettori.

1.2.1 Proprietà degli autovalori

Definizione 1.2.1. Chiamiamo **polinomio caratteristico** la quantità

$$P_n(\lambda) = \det(\hat{A} - \lambda \hat{J}).$$

Il grado del polinomio è pari all'ordine della matrice $\hat{A}$ (infatti nel calcolo del determinante, comparirà al massimo un prodotto di λ per se stesso n volte, che risulta dal prodotto degli elementi sulla diagonale) e quindi il calcolo degli autovalori viene ricondotto alla ricerca degli zeri del polinomio caratteristico, per il quale vale il teorema fondamentale dell'algebra (vedi la dispensa sui numeri complessi). Il succitato teorema ci dice che ogni equazione algebrica di ordine n ammette n soluzioni reali o complesse, semplici o ripetute. Per cui, nella ricerca di autovalori di un operatore ci possiamo aspettare:

- autovalori reali e distinti;
- autovalori reali e coincidenti;
- autovalori complessi coniugati.

Ricordiamo qui (per i dettagli rimandiamo ancora una volta alla dispensa sui numeri complessi) che definiamo **molteplicità algebrica** di una radice, il numero di volte che tale radice compare nell'equazione algebrica. Per similitudine, definiamo **molteplicità algebrica di un autovalore** $= ma(\lambda)$ come il numero di volte che tale autovalore si presenta come soluzione dell'equazione caratteristica.

Definizione 1.2.2. Chiamiamo **spettro** di $\hat{A}$ e lo indichiamo con $\sigma(\hat{A})$ l'insieme degli autovalori distinti di $\hat{A}$.

Esempio 1.2.3. Supponiamo che il polinomio caratteristico abbia la forma:

$$P_n(\lambda) = \lambda^5 - \lambda^4 - \lambda + 1.$$

Allora mettendo in evidenza singoli termini (oppure usando la tecnica di Ruffini) otteniamo

$$\begin{aligned} P_n(\lambda) &= \lambda^5 - \lambda^4 - \lambda + 1 = \lambda^4(\lambda - 1) - \lambda + 1 = (\lambda^4 - 1)(\lambda - 1) = \\ &= (\lambda^2 - 1)(\lambda^2 + 1)(\lambda - 1) = (\lambda - 1)^2(\lambda + 1)(\lambda^2 + 1). \end{aligned}$$

Tale fattorizzazione ci porta a determinare i seguenti autovalori

- un autovalore singolo $\lambda = -1$ (con $ma(-1) = 1$);
- due autovalori coincidenti $\lambda = 1$ (con $ma(1) = 2$);
- due autovalori complessi coniugati $\lambda = \pm i$ (con $ma(i) = ma(-i) = 1$).

Quindi, lo spettro di $\hat{A}$ sarà pari a

$$\sigma(\hat{A}) = (-1, 1, i, -i).$$

1.2.2 Proprietà degli autovettori

Una volta identificati gli autovalori, per trovare gli autovettori è necessario risolvere un problema omogeneo del tipo $(\hat{A} - \lambda \hat{J})\mathbf{u} = \mathbf{0}$, (vedi equazione (1.4)) per ogni autovalore che risolve l'equazione caratteristica. Ora, per il teorema di Rouchè-Capelli (come riportato nella dispensa sugli operatori lineari), tale sistema omogeneo ammette $\infty^{n-Rg(\hat{A}-\lambda\hat{J})}$ soluzioni che determinano un sottospazio vettoriale. Qualsiasi vettore di tale sottospazio vettoriale (tranne il vettore nullo) può essere considerato un autovettore associato all'autovalore λ . Essendo $\det(\hat{A} - \lambda \hat{J}) = 0$, relazione vera se λ è un autovalore, siamo sicuri di avere almeno ∞^1 soluzioni per il sistema omogeneo. Ma quante ce ne sono in generale?

Definizione 1.2.4. *Sia data $\hat{A} \in \mathcal{M}(n, n)$ e λ tale che $\det(\hat{A} - \lambda \hat{J}) = 0$. Allora la dimensione del sottospazio vettoriale delle soluzioni dall'equazione omogenea*

$$(\hat{A} - \lambda \hat{J})\mathbf{u} = \mathbf{0}$$

viene detta molteplicità geometrica di λ :

$$mg(\lambda) = n - Rg(\hat{A} - \lambda \hat{J}). \quad (1.5)$$

Per cui si definisce molteplicità geometrica la dimensione del sottospazio vettoriale invariante generato da tutti gli autovettori (indipendenti) associati all'autovalore λ . Ci limitiamo a riportare senza dimostrazione un risultato fondamentale nella teoria dei sottospazi invarianti:

Theorem 1.2.5. *Data una matrice $\hat{A} \in \mathcal{M}(n, n)$ per ogni autovalore di $\hat{A}$ vale la seguente disuguaglianza:*

$$ma(\lambda) \geq mg(\lambda) \geq 1$$

Per cui, associati alle varie tipologie di autovalori che si possono presentare avremo:

- un autovettore associato ad ogni autovalore reale semplice (ossia non ripetuto);
- uno o più autovettori associati ad ogni autovalore reale e ripetuto;
- una coppia di autovettori complessi coniugati associati ad ogni coppia di autovalori complessi coniugati semplici (ossia non ripetuti. Il caso di autovalori complessi ripetuti non verrà qui preso in esame).

Per concludere questo paragrafo presenteremo un risultato importante che ci garantisce che autovettori associati ad autovalori distinti sono indipendenti:

Theorem 1.2.6. *Sia $\hat{A} \in \mathcal{M}(n, n)$ e λ e μ due autovalori distinti di $\hat{A}$. Allora se chiamiamo $\mathbf{u}_\lambda$ e $\mathbf{u}_\mu$ due autovettori associati a λ e μ , rispettivamente, allora $\mathbf{u}_\lambda$ e $\mathbf{u}_\mu$ sono linearmente indipendenti.*

Dimostrazione. Supponiamo che i due autovettori siano linearmente dipendenti, ossia $\mathbf{u}_\lambda = \alpha \mathbf{u}_\mu$ per qualche α . Si ha che

$$\hat{A}\mathbf{u}_\lambda = \lambda \mathbf{u}_\lambda$$

per costruzione, ma dalla dipendenza lineare tra $\mathbf{u}_\lambda$ e $\mathbf{u}_\mu$ si ha anche

$$\hat{A}\mathbf{u}_\lambda = \hat{A}\alpha \mathbf{u}_\mu = \alpha \hat{A}\mathbf{u}_\mu = \alpha \mu \mathbf{u}_\mu.$$

Ma visto che $\mathbf{u}_\lambda = \alpha \mathbf{u}_\mu$ l'ultima uguaglianza ci dice che $\hat{A}\mathbf{u}_\lambda = \mu \mathbf{u}_\lambda$ e quindi $\lambda = \mu$ contro le ipotesi del teorema. $\square$

È possibile generalizzare questo teorema a k autovalori distinti di $\hat{A}$, ma ne omettiamo la dimostrazione in quanto non troppo elegante considerando anche che non aggiunge molto alla comprensione generale dell'argomento:

Theorem 1.2.7. *Sia $\hat{A} \in \mathcal{M}(n, n)$ e λ_i con $i = 1 \cdots k$, k autovalori distinti di $\hat{A}$. Allora gli autovettori $\mathbf{u}_i$, con $i = 1 \cdots k$, associati ai vari λ_i , sono linearmente indipendenti.*

1.3 Decomposizione spettrale

A questo punto possiamo tentare di fare qualche considerazione generale. I sottospazi invarianti di un operatore sono quei sottospazi nei quali l'azione dell'operatore è particolarmente semplice. La base dei sottospazi invarianti è formata da autovettori linearmente indipendenti. Per cui, se potessimo avere una base completa formata solo di autovettori, la rappresentazione dell'operatore in tale base sarebbe estremamente semplice. Purtroppo, come vedremo, non sempre esiste una base completa di autovettori. Ma, in generale, la rappresentazione di un operatore rispetto ai vettori che generano i suoi sottospazi invarianti è sempre la rappresentazione più semplice possibile. Questo procedimento, ossia la determinazione dei sottospazi invarianti di un operatore e la rappresentazione di tale operatore rispetto ai vettori che generano i sottospazi invarianti va sotto il nome di **decomposizione spettrale**.

1.3.1 Autovalori reali e distinti

Supponiamo che una matrice $\hat{A} \in \mathcal{M}(n, n)$ possieda n autovalori reali e distinti, $\lambda_i \in \mathbb{R}$. In questo caso, visto che ad ogni autovalore è associato un autovettore indipendente, esiste una base completa di autovettori. Esistono, cioè, n autovettori $\mathbf{u}_i$ linearmente indipendenti e tali che $\hat{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i$. Per calcolare la rappresentazione di $\hat{A}$ nella base degli autovettori, $\mathbb{D} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$, si deve calcolare l'azione di $\hat{A}$ su ogni vettore $\mathbf{u}_i$. Ma, proprio per le proprietà degli autovettori, l'azione di $\hat{A}$ su di essi è banale:

$$\begin{cases} \hat{A}\mathbf{u}_1 &= \lambda_1 \mathbf{u}_1 &= \lambda_1 \mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + 0\mathbf{u}_n, \\ \hat{A}\mathbf{u}_2 &= \lambda_2 \mathbf{u}_2 &= 0\mathbf{u}_1 + \lambda_2 \mathbf{u}_2 + \cdots + 0\mathbf{u}_n, \\ &\vdots &\vdots \\ \hat{A}\mathbf{u}_n &= \lambda_n \mathbf{u}_n &= 0\mathbf{u}_1 + 0\mathbf{u}_2 + \cdots + \lambda_n \mathbf{u}_n. \end{cases}$$

E quindi, per il teorema di rappresentazione, $\hat{A}$ nella base degli autovettori, avrà una forma diagonale, i.e.,

$$\hat{A}_{\mathbb{D}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \quad (1.6)$$

Per arrivare a questo stesso risultato avremmo potuto passare per la matrice del cambiamento di base dalla base canonica $\mathbb{E} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ alla base degli autovettori $\mathbb{D} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$, $\mathbb{D} = \mathbb{E}\hat{U}$ con:

$$\hat{U} = \left(\begin{pmatrix} \mathbf{u}_1 \end{pmatrix} \quad \begin{pmatrix} \mathbf{u}_2 \end{pmatrix} \quad \cdots \quad \begin{pmatrix} \mathbf{u}_n \end{pmatrix} \right)$$

e calcolare

$$\hat{A}_{\mathbb{D}} = \hat{U}^{-1} \hat{A} \hat{U},$$

ma è molto più potente e semplice usare il teorema di rappresentazione.

Vediamo, in generale, in quali casi una matrice è diagonalizzabile:

Theorem 1.3.1. *Data una matrice $\hat{A} \in \mathcal{M}(n, n)$, essa ammette una forma diagonale se tutti i suoi autovalori sono reali e se, supponendo k il numero di autovalori distinti, si ha*

$$\sum_{i=1}^k mg(\lambda_i) = n.$$

Ovvero, equivalentemente,

$$\forall i \quad ma(\lambda_i) = mg(\lambda_i).$$

Dimostrazione. La dimostrazione di questo teorema è banale perché dire che la somma delle molteplicità geometriche di tutti gli autovalori è pari ad n significa affermare che esistono n autovettori linearmente indipendenti con i quali costruire una base. Per quanto abbiamo visto prima, la rappresentazione dell'operatore rispetto a questa base sarà diagonale. $\square$

Un caso particolare di matrici che hanno sempre una rappresentazione diagonale è costituito dalle matrici simmetriche. Tale risultato fondamentale della teoria degli operatori lineari va sotto il nome di teorema spettrale:

Theorem 1.3.2. *Data una matrice $\hat{A} \in \mathcal{M}(n, n)$, se $\hat{A}$ è simmetrica, ossia $\hat{A} = \hat{A}^t$, essa avrà autovalori tutti reali, λ_i , per i quali varrà $ma(\lambda_i) = mg(\lambda_i)$, ossia ammetterà una base completa di autovettori.*

Nello sviluppo della teoria useremo solo incidentalmente matrici simmetriche, ma esistono molti contesti (per esempio nella teoria delle forme quadratiche, oppure nella teoria dell'ottimizzazione statica, ecc. ecc.) nei quali le matrici simmetriche hanno un ruolo molto importante, ed il teorema spettrale garantisce per queste matrici una rappresentazione in forma diagonale.

1.3.2 Autovalori reali e coincidenti

Non tutti gli operatori sono diagonalizzabili, anche nel caso in cui abbiamo solo autovalori reali. Per semplicità di esposizione, supponiamo di avere $\hat{A} \in \mathcal{M}(n, n)$ con $n = 2$ e che nello spettro di $\hat{A}$ ci sia un autovalore λ ripetuto, con molteplicità algebrica $ma(\lambda) = 2$ e molteplicità geometrica $mg(\lambda) = 1$. Questo fa sì che associato a λ ci sia solo un autovettore, non sufficiente a diagonalizzare la matrice $\hat{A}$, anche se il sottospazio invariante associato a λ , che chiameremo $\mathbb{U}$, ha dimensione 2. Possiamo comunque trovare una rappresentazione semplice per $\hat{A}$? La risposta è sì, e l'idea è di cercare un vettore $\mathbf{u}' \in \mathbb{U}$ facendo in modo che l'azione di $\hat{A}$ su $\mathbf{u}'$ sia particolarmente semplice e che non si esca da $\mathbb{U}$. Ovviamente la condizione $\hat{A}\mathbf{u}' \in \mathbb{U}$ è più debole della condizione $\hat{A}\mathbf{u} = \lambda\mathbf{u}$ ma garantirà comunque una rappresentazione abbastanza semplice per la matrice $\hat{A}$ relativamente al sottospazio $\mathbb{U}$. Dobbiamo quindi aggiungere all'autovettore un **autovettore generalizzato**.

Definizione 1.3.3. *Data una matrice $\hat{A} \in \mathcal{M}(n, n)$ e un autovalore λ di $\hat{A}$, definiamo autovettori generalizzati di ordine k , $\mathbf{u}^{(k)}$, i vettori che verificano la seguente equazione*

$$(\hat{A} - \lambda\hat{J})^{k+1}\mathbf{u}^{(k)} = \mathbf{0}.$$

Resta inteso, dalla definizione appena data, che gli autovettori generalizzati di ordine 0 sono gli autovettori semplici associati a λ (confronta l'ultima equazione con $k = 0$ con l'equazione (1.4) che definisce gli autovettori). Sfruttando l'esempio di apertura di questo paragrafo, calcoliamo $\mathbf{u}^{(1)}$:

$$(\hat{A} - \lambda\hat{J})^2\mathbf{u}^{(1)} = (\hat{A} - \lambda\hat{J})(\hat{A} - \lambda\hat{J})\mathbf{u}^{(1)} = \mathbf{0}. \quad (1.7)$$

È evidente che una soluzione semplice è data da $(\hat{A} - \lambda\hat{J})\mathbf{u} = \mathbf{0}$ visto che la successiva applicazione di $(\hat{A} - \lambda\hat{J})$ sul vettore nullo non potrà che dare il vettore nullo. Possiamo però trovare un'altra soluzione. Per verificare l'ultima uguaglianza a destra dell'equazione (1.7), è sufficiente che sia

$$(\hat{A} - \lambda\hat{J})\mathbf{u}^{(1)} = \mathbf{u},$$

visto che se $\mathbf{u}$ è un autovettore, allora $(\hat{A} - \lambda\hat{J})\mathbf{u} = \mathbf{0}$. Per cui si ha la seguente relazione

$$(\hat{A} - \lambda\hat{J})^2\mathbf{u}^{(1)} = \mathbf{0} \quad \rightarrow \quad \begin{cases} (\hat{A} - \lambda\hat{J})\mathbf{u} = \mathbf{0} \\ (\hat{A} - \lambda\hat{J})\mathbf{u}^{(1)} = \mathbf{u} \end{cases} \quad (1.8)$$

Quindi, avendo ricavato i due vettori $\mathbf{u}, \mathbf{u}^{(1)}$ li usiamo come base per generare il sottospazio $\mathbb{U}$, $\mathbb{U} = \text{Span}(\mathbf{u}, \mathbf{u}^{(1)})$. Dalle equazioni di destra della (1.8) si riesce a dimostrare che il sottospazio vettoriale $\mathbb{U}$ è un sottospazio invariante di dimensione 2. Infatti si ha che

$$\begin{cases} \hat{A}\mathbf{u} = \lambda\mathbf{u} \\ \hat{A}\mathbf{u}^{(1)} = \mathbf{u} + \lambda\mathbf{u}^{(1)} \end{cases} \quad (1.9)$$

e quindi, se prendiamo il generico vettore $\mathbf{v} \in \mathbb{U}$, con $\mathbf{v} = \alpha\mathbf{u} + \beta\mathbf{u}^{(1)}$, e calcoliamo $\hat{A}\mathbf{v}$ otteniamo

$$\hat{A}(\alpha\mathbf{u} + \beta\mathbf{u}^{(1)}) = \alpha\lambda\mathbf{u} + \beta(\mathbf{u} + \lambda\mathbf{u}^{(1)}) = (\alpha\lambda + \beta)\mathbf{u} + \beta\lambda\mathbf{u}^{(1)},$$

ossia ogni vettore di $\mathbb{U}$ viene trasformato da $\hat{A}$ in un altro vettore di $\mathbb{U}$.
Qual'è la rappresentazione di $\hat{A}$ rispetto alla base $\mathbb{J} = (\mathbf{u}, \mathbf{u}^{(1)})$ scelta per $\mathbb{U}$?

$$\begin{cases} \hat{A}\mathbf{u} = \lambda\mathbf{u} = (\lambda)\mathbf{u} + (0)\mathbf{u}^{(1)} = \mathbb{J} \begin{pmatrix} \lambda \\ 0 \end{pmatrix}, \\ \hat{A}\mathbf{u}^{(1)} = \mathbf{u} + \lambda\mathbf{u}^{(1)} = (1)\mathbf{u} + (\lambda)\mathbf{u}^{(1)} = \mathbb{J} \begin{pmatrix} 1 \\ \lambda \end{pmatrix}, \end{cases}$$

E quindi, per il teorema di rappresentazione, $\hat{A}$ nella base di $\mathbb{U}$, avrà una forma del tipo

$$\hat{A}_{\mathbb{J}} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \quad (1.10)$$

Questa forma matriciale si chiama **blocchetto di Jordan** di ordine 2.

In generale, si dimostra che se $\hat{A} \in \mathcal{M}(n, n)$ e se λ è autovalore di $\hat{A}$ con $ma(\lambda) = k$ e $mg(\lambda) = 1$ allora esistono $k-1$ autovettori generalizzati di $\hat{A}$ associati all'autovalore λ . Tali autovettori si determinano con la regola della catena

$$(\hat{A} - \lambda\hat{J})^{k+1}\mathbf{u}^{(k)} = \mathbf{0} \quad \rightarrow \quad \begin{cases} (\hat{A} - \lambda\hat{J})\mathbf{u} = \mathbf{0} \\ (\hat{A} - \lambda\hat{J})\mathbf{u}^{(1)} = \mathbf{u} \\ (\hat{A} - \lambda\hat{J})\mathbf{u}^{(2)} = \mathbf{u}^{(1)} \\ \vdots \\ (\hat{A} - \lambda\hat{J})\mathbf{u}^{(k-1)} = \mathbf{u}^{(k-2)} \end{cases} \quad (1.11)$$

Se fissiamo una base composta da autovettore e autovettori generalizzati, $\mathbb{J} = (\mathbf{u}, \mathbf{u}^{(1)}, \mathbf{u}^{(2)}, \dots, \mathbf{u}^{(k-1)})$, allora rispetto a tale base, la rappresentazione di $\hat{A}$ avrà la forma

$$\hat{A}_{\mathbb{J}} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix} \quad (1.12)$$

1.3.3 Autovalori complessi coniugati

Nel caso in cui nello spettro di un operatore ci siano degli autovalori complessi coniugati non esiste nessuna rappresentazione dell'operatore rispetto ad una base reale che porti l'operatore ad avere forma diagonale. È fondamentale specificare "base reale" perché nel campo complesso, ossia usando vettori di base complessi, l'operatore sarebbe diagonale, ma noi siamo interessati a lavorare sempre e solo con basi reali.

Theorem 1.3.4. *Sia $\hat{A}$ una matrice reale con λ autovalore complesso e $\mathbf{u}$ autovettore complesso associato. Allora anche il complesso coniugato di λ , $\bar{\lambda}$, sarà autovalore di $\hat{A}$, con autovettore associato $\bar{\mathbf{u}}$.*

Dimostrazione. Per le ipotesi del teorema si ha che $\hat{A}\mathbf{u} = \lambda\mathbf{u}$. Ma, prendendo il complesso coniugato di questa espressione, e considerando che essendo $\hat{A}$ reale si ha che $\overline{\hat{A}} = \hat{A}$, allora

$$\overline{\hat{A}\mathbf{u} = \lambda\mathbf{u}} \rightarrow \overline{\hat{A}\mathbf{u}} = \overline{\lambda\mathbf{u}} \rightarrow \hat{A}\overline{\mathbf{u}} = \overline{\lambda}\overline{\mathbf{u}},$$

che implica, appunto, la tesi del teorema. $\square$

Quindi gli autovalori complessi compariranno sempre in coppia, così come gli autovettori associati, che formeranno un sottospazio complesso bidimensionale. Vediamo come sia possibile trovare un sottospazio vettoriale bidimensionale reale, $\mathbb{U}_{\mathbb{C}}$, a partire dal sottospazio complesso formato dagli autovettori complessi coniugati tale che l'operatore $\hat{A}$ sia invariante rispetto a tale sottospazio, e che la rappresentazione di $\hat{A}$ rispetto ai vettori che generano $\mathbb{U}_{\mathbb{C}}$ sia relativamente semplice. Visto che λ e $\mathbf{u}$ sono complessi, scomponiamoli nella loro parte reale e parte immaginaria:

$$\lambda = \lambda_a + i\lambda_b, \quad \mathbf{u} = \mathbf{u}_a + i\mathbf{u}_b.$$

Non perdiamo di generalità se scegliamo λ_b positiva. Vediamo come agisce $\hat{A}$ sulla parte reale e sulla parte immaginaria dell'autovettore $\mathbf{u}$:

$$\begin{aligned} \hat{A}\mathbf{u} = \lambda\mathbf{u} &\rightarrow \hat{A}(\mathbf{u}_a + i\mathbf{u}_b) = (\lambda_a + i\lambda_b)(\mathbf{u}_a + i\mathbf{u}_b) \rightarrow \\ &\rightarrow \hat{A}\mathbf{u}_a + i\hat{A}\mathbf{u}_b = \lambda_a\mathbf{u}_a - \lambda_b\mathbf{u}_b + i(\lambda_b\mathbf{u}_a + \lambda_a\mathbf{u}_b). \end{aligned}$$

Ora, uguagliando separatamente parte reale e parte immaginaria si ha

$$\begin{cases} \hat{A}\mathbf{u}_a &= \lambda_a\mathbf{u}_a - \lambda_b\mathbf{u}_b \\ \hat{A}\mathbf{u}_b &= \lambda_b\mathbf{u}_a + \lambda_a\mathbf{u}_b \end{cases} \quad (1.13)$$

da cui otteniamo la rappresentazione di $\hat{A}$ rispetto ai vettori di base $\mathbb{C} = (\mathbf{u}_a, \mathbf{u}_b)$:

$$\hat{A}_{\mathbb{C}} = \begin{pmatrix} \lambda_a & \lambda_b \\ -\lambda_b & \lambda_a \end{pmatrix} \quad (1.14)$$

Come nel caso di Jordan, anche qui non abbiamo una rappresentazione diagonale dell'operatore, ma è comunque una rappresentazione abbastanza semplice (e tutta la sua semplicità apparirà nel prossimo capitolo), con la quale sarà possibile effettuare tutta una serie di operazioni che ci saranno utili nel seguito.

1.3.4 Azione generale degli operatori

Un'altro modo per apprezzare la potenza dei sottospazi invarianti è quello di vedere l'azione degli operatori sui vettori dello spazio rappresentati rispetto alla base dei sottospazi invarianti. Supponiamo inizialmente che una matrice $\hat{A}$ possieda una base completa di autovettori. Esistono quindi n vettori $\mathbf{u}_i$ linearmente indipendenti e tali che $\hat{A}\mathbf{u}_i = \lambda_i\mathbf{u}_i$. Se decidiamo di rappresentare i vettori di $\mathbb{R}^n$ nella base degli autovettori si ha

$$\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots \alpha_n\mathbf{u}_n$$

dove le α_i sono le coordinate del vettore $\mathbf{v}$ rispetto alla base degli autovettori $\mathbb{D} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n)$, e quindi $\mathbf{v} = \mathbb{D}\mathbf{v}_{\mathbb{D}}$, dove $\mathbf{v}_{\mathbb{D}}$ è il vettore colonna che contiene i coefficienti α_i del vettore $\mathbf{v}$ rispetto alla base $\mathbb{D}$:

$$\mathbf{v}_{\mathbb{D}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Per le proprietà degli autovettori un semplice calcolo diretto porta a:

$$\begin{aligned} \hat{A}\mathbf{v} &= \hat{A}(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \dots + \alpha_n\mathbf{u}_n) = \alpha_1\hat{A}\mathbf{u}_1 + \alpha_2\hat{A}\mathbf{u}_2 + \dots + \alpha_n\hat{A}\mathbf{u}_n = \\ &= \alpha_1\lambda_1\mathbf{u}_1 + \alpha_2\lambda_2\mathbf{u}_2 + \dots + \alpha_n\lambda_n\mathbf{u}_n. \end{aligned}$$

Quindi l'applicazione della matrice su un vettore rappresentato nella base degli autovettori si riduce al prodotto delle componenti del vettore per gli autovalori. Questo non è nient'altro che l'applicazione della rappresentazione di $\hat{A}$ rispetto alla base $\mathbb{D}$ degli autovettori (rappresentazione diagonale) ed i vettori espressi nella stessa base $\mathbb{D}$:

$$\hat{A}\mathbf{v} = \hat{A}_{\mathbb{D}}\mathbf{v}_{\mathbb{D}} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1\lambda_1\mathbf{u}_1 + \alpha_2\lambda_2\mathbf{u}_2 + \dots + \alpha_n\lambda_n\mathbf{u}_n.$$

In generale, l'azione di una matrice su un sottospazio invariante sarà:

- se il sottospazio è unidimensionale ed associato ad un autovalore reale semplice

$$\hat{A}\mathbf{u} = \lambda\mathbf{u}; \quad (1.15)$$

- se il sottospazio è bidimensionale ed associato ad un autovalore reale con molteplicità algebrica pari a 2

$$\hat{A}\mathbf{u} = \hat{A}(\alpha_1\mathbf{u} + \alpha_2\mathbf{u}^{(1)}) = (\lambda\alpha + \beta)\mathbf{u} + \lambda\beta\mathbf{u}^{(1)}; \quad (1.16)$$

- se il sottospazio è bidimensionale ed associato ad una coppia di autovalori complessi coniugati

$$\hat{A}\mathbf{u} = \hat{A}(\alpha_1\mathbf{u}_a + \alpha_2\mathbf{u}_b) = (\lambda_a\alpha + \lambda_b\beta)\mathbf{u}_a + (-\lambda_b\alpha + \lambda_a\beta)\mathbf{u}_b. \quad (1.17)$$

Quindi, nel caso di un generico operatore $\hat{A} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ esso può avere differenti tipologie di sottospazi invarianti, e quindi la nuova base rispetto alla quale calcolare la rappresentazione di $\hat{A}$ può essere composta da autovettori semplici, autovettori e relativi autovettori generalizzati e parte reale e parte immaginaria di autovettori complessi coniugati,

$$\mathbb{F} = (\mathbf{u}_1, \dots, \mathbf{u}_l, \mathbf{v}_1, \mathbf{v}_1^{(1)}, \dots, \mathbf{v}_m, \mathbf{v}_m^{(1)}, \mathbf{w}_{1a}, \mathbf{w}_{1b}, \dots, \mathbf{w}_{pa}, \mathbf{w}_{pb})$$

dove gli $\mathbf{u}_i$ per $i = 1 \cdots l$ sono l autovettori associati agli autovalori semplici λ_i , gli $\mathbf{v}_j$ e $\mathbf{v}_j^{(1)}$ per $j = 1 \cdots m$ sono autovettori e relativo primo autovettore generalizzato associato agli autovalori doppi λ_j (dove per semplicità si è supposto che $ma(\lambda_j) = 2$ e $mg(\lambda_j) = 1$) e infine gli $\mathbf{w}_{a_k}$ e $\mathbf{w}_{b_k}$ per $k = 1 \cdots p$ sono parte reale e parte immaginaria degli autovettori complessi coniugati $\mathbf{w} = \mathbf{w}_{k_a} \pm i\mathbf{w}_{k_b}$ associati agli autovalori complessi coniugati $\lambda_k = \lambda_{k_a} \pm i\lambda_{k_b}$. Facendo un semplice conto dimensionale dovrà essere $n = l + 2m + 2p$ e la rappresentazione dell'operatore $\hat{A}$ nella base $\mathbb{F}$ sarà:

$$\hat{A}_{\mathbb{F}} = \begin{pmatrix} \begin{pmatrix} \ddots & 0 & 0 \\ 0 & \lambda_i & 0 \\ 0 & 0 & \ddots \end{pmatrix} & 0 & 0 \\ 0 & \begin{pmatrix} \ddots & 0 & 0 & \cdots \\ 0 & \lambda_j & 1 & 0 \\ 0 & 0 & \lambda_j & 0 \\ \cdots & 0 & 0 & \ddots \end{pmatrix} & 0 \\ 0 & 0 & \begin{pmatrix} \ddots & 0 & 0 & \cdots \\ 0 & \lambda_{a_k} & \lambda_{b_k} & 0 \\ 0 & -\lambda_{b_k} & \lambda_{a_k} & 0 \\ \cdots & 0 & 0 & \ddots \end{pmatrix} \end{pmatrix}$$