Algebra lineare

18 maggio 2023

Indice

Capitolo 1

Matrici

Contents

In R carichiamo il pacchetto `lbla` con utilities e funzionalità (tra l'altro con l'assegnazione `i <- 1i` che semplifica la scrittura con i complessi).

1.1 Matrici

Definizione 1.1.1 (Matrice $m \times n$)**.** È una tabella di m righe ed n colonne di numeri complessi[1] (o di simboli che rappresentano complessi), racchiusa tra parentesi. Si denota con:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mj} & \dots & a_{mn} \end{bmatrix} = [a_{ij}]_{j \le n}^{i \le m} \tag{1.1}$$

o più semplicemente con $[a_{ij}]$ se le dimensioni sono note.

Esempio 1.1.1. Le seguenti

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 \\ -\pi \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

costituiscono esempi rispettivamente di matrici 2×2 a coefficienti complessi, 2×1 a coefficienti reali e 3×4 a coefficienti naturali.

In R per creazione di matrici si usi `matrix` o `rmatrix` (per il riempimento per riga di default), per ottenerne le dimensioni `dim`, `nrow`, `ncol`:

```
(A <- rmatrix(c(1 + i, 2 - i, 3, -i), nrow = 2))

##      [,1] [,2]
## [1,] 1+1i 2-1i
## [2,] 3+0i 0-1i

(C <- rmatrix(1:12, nrow = 3, ncol = 4))

##      [,1] [,2] [,3] [,4]
## [1,]    1    2    3    4
## [2,]    5    6    7    8
## [3,]    9   10   11   12

dim(C); nrow(C); ncol(C)

## [1] 3 4
## [1] 3
## [1] 4
```

[1] Alcune trattazioni si riferiscono ad elementi di $\mathbb{K}$ come un modo per indicare equivalentemente $\mathbb{C}$ o $\mathbb{R}$. Ma essendo $\mathbb{R}$ un sottoinsieme di $\mathbb{C}$, ci si riferisce a quest'ultimo per non introdurre notazione inutile.

In maxima per la creazione delle medesime matrici la sintassi è:

```
##
## A:matrix([1+%i,2-%i],[3,-%i])
##                              [ %i + 1  2 - %i ]
##                              [                ]
##                              [   3      - %i  ]
## C:matrix([1,2,3,4],[5,6,7,8],[9,10,11,12])
##                               [ 1  2   3   4  ]
##                               [               ]
##                               [ 5  6   7   8  ]
##                               [               ]
##                               [ 9  10  11  12 ]
```

Definizione 1.1.2 (Elemento/coefficiente di posto (i,j))**.** È il coefficiente $a_{ij} \in \mathbb{C}$ della matrice, che si trova nella i-esima riga e nella j-esima colonna.

Osservazione 1. A volte si indica con $a_{i,j}$ ad esempio quando si indica l'ultimo elemento della penultima riga come $a_{m-1,n}$.

Osservazione 2. Una matrice $m \times n$ può essere pensata come un vettore con $m \times n$ componenti quindi come un elemento dall'insieme $\mathbb{C}^{m \times n}$. A tal proposito diamo le seguenti definizioni.

Definizione 1.1.3 (Insiemi di matrici a coefficienti complessi)**.** Indichiamo con $M_{n\times p}(\mathbb{C})$ l'insieme delle matrici $n \times p$ a coefficienti complessi

Osservazione 3. Vediamo sin da subito alcuni tipi speciali di matrici

Definizione 1.1.4 (Matrice nulla)**.** Matrice in cui tutti i coefficienti sono uguali a 0; si denota con $\mathbb{O}_{mn}$, o semplicemente con $\mathbb{O}$ (se chiaro dal contesto la dimensione della matrice).

Osservazione 4. La matrice nulla ha proprietà analoghe a quelle di 0 tra i numeri: è l'elemento neutro per la somma e annulla tutti i prodotti tra matrici, come si avrà modo di constatare.

Definizione 1.1.5 (Matrice quadrata)**.** Matrice in cui numero righe è uguale al numero colonne ($m = n$). Solo per queste matrici si può parlare di ordine e diagonali.

Definizione 1.1.6 (Ordine di matrice)**.** In una matrice quadrata corrisponde al numero di righe (o ugualmente colonne).

Definizione 1.1.7 (Algebra delle matrici complesse $n\times n$)**.** Indicato con $M_n(\mathbb{C})$, consiste nell'insieme di tutte le matrici a coefficienti complessi di ordine n, dotato delle operazioni di prodotto per scalari, somma e prodotto (righe per colonne)

Osservazione 5. Analogamente si può parlare di $M_n(\mathbb{R})$ e $M_n(\mathbb{Q})$, ma teniamo $\mathbb{C}$ per generalità.

Definizione 1.1.8 (Diagonale principale)**.** In una matrice quadrata la *diagonale principale* (o semplicemente diagonale) è la diagonale NO-SE ossia quella costituita dagli elementi di posto (i,i), con $i = 1, \ldots, m$

Definizione 1.1.9 (Diagonale secondaria)**.** In una matrice quadrata la *diagonale secondaria* è la diagonale NE-SO, ossia costituita dagli elementi di posto $(i, m - i + 1)$.

Esempio 1.1.2. La matrice

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

è quadrata di ordine 3, con diagonale $(1, 6, 11)$ e diagonale secondaria $(3, 6, 9)$

Definizione 1.1.10 (Matrice diagonale)**.** Matrice quadrata con elementi al di fuori della diagonale nulli, ossia $d_{ij} = 0$ se $i \neq j$; si indica con $\mathbf{Diag}(\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n)$ la matrice diagonale in cui $d_{ii} = \lambda_i$

$$\mathbf{Diag}(\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_2 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_3 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_n \end{bmatrix} \tag{1.2}$$

Definizione 1.1.11 (Matrice psedo-diagonale)**.** Matrice (non quadrata) $\mathbf{A}$ di tipo $m \times n$ con $m \neq n$ i cui elementi al di fuori della diagonale sono nulli, ossia con $a_{ij} = 0$ se $i \neq j$.

Definizione 1.1.12 (Matrice scalare)**.** Matrice diagonale con coefficienti sulla diagonale uguali fra loro, quindi del tipo $\mathbf{Diag}(d, d, d, \ldots, d)$; sono le matrici che dal punto di vista algebrico più si avvicinano agli scalari.

Definizione 1.1.13 (Matrice identità/unità)**.** Matrice scalare con coefficienti della diagonale pari a 1. Si indica con $\mathbf{I}_n$ (n è il numero di righe/colonne) ed è definita da

$$\mathbf{I}_n = [\delta_{ij}] = [\mathbb{1}(i = j)], \qquad i, j \leq n \tag{1.3}$$

con $\mathbb{1}(i = j)$ funzione indicatrice.

Esempio 1.1.3.

$$\mathbf{I}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In R la funzione `diag` serve per creare matrici diagonali di varia tipologia o per modificare la diagonale di una matrice esistente.

```
## Creazione di matrice diagonale generica
## (se x è vettore specifica diagonale)
diag(x = 1:2)

##      [,1] [,2]
## [1,]    1    0
## [2,]    0    2

## Matrice scalare (con 2 sulla diagonale)
diag(x = 2, nrow = 3)

##      [,1] [,2] [,3]
## [1,]    2    0    0
## [2,]    0    2    0
## [3,]    0    0    2

## Matrice identità (se x è scalare specifica la dimensione)
diag(x = 2)

##      [,1] [,2]
## [1,]    1    0
## [2,]    0    1

## Uso di diag per ottenere/modifica la diagonale di
## una matrice (x, data in input)
(m <- matrix(rnorm(4), ncol = 2))

##            [,1]      [,2]
## [1,]  1.0430495 0.3322731
## [2,] -0.2817964 1.0200671

diag(m)

## [1] 1.043050 1.020067

diag(m) <- 0
m

##            [,1]      [,2]
## [1,]  0.0000000 0.3322731
## [2,] -0.2817964 0.0000000
```

In maxima si usa `diagmatrix`.

Definizione 1.1.14 (Matrice triangolare superiore)**.** Matrice quadrata con coefficienti al di sotto della diagonale nulli ossia,

$$a_{ij} = 0, \quad \forall i > j \tag{1.4}$$

Tali matrici vengono solitamente denotate con **T** o **U** (*upper*).

Esempio 1.1.4. Una generica matrice triangolare superiore 3×3 è del tipo

$$\begin{bmatrix} t_{11} & t_{12} & t_{13} \\ 0 & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{bmatrix}$$

Definizione 1.1.15 (Matrice uni-triangolare superiore)**.** Matrice triangolare superiore con coefficienti della diagonale principale tutti pari a 1.

Definizione 1.1.16 (Matrice triangolare inferiore)**.** Matrice con coefficienti al di sopra della diagonale nulli, ossia

$$a_{ij} = 0, \quad \forall i < j \tag{1.5}$$

denotata con **T** o **L** (*lower*).

Esempio 1.1.5. Una generica matrice triangolare inferiore 3×3 è del tipo

$$\begin{bmatrix} t_{11} & 0 & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & t_{33} \end{bmatrix}$$

Definizione 1.1.17 (Matrice uni-triangolare inferiore)**.** Matrice triangolare inferiore con coefficienti della diagonale principale tutti pari a 1.

Ai fini delle matrici triangolari, `upper.tri` e `lower.tri` restituiscono una matrice di logici della stessa dimensione della matrice fornitagli, con la possibilità di specificare se la diagonale deve essere inclusa. Questi indici possono essere utilizzati leggere/scrivere elementi della matrice e quindi creare matrici triangolari superiori o inferiori. Un esempio con `upper.tri` (`lower.tri` allo stesso modo)

```
## parto da matrice nulla
O <- matrix(0, nrow = 2, ncol = 2)

## matrice triangolare superiore
U <- O
upper.tri(U, diag = TRUE)

##       [,1] [,2]
## [1,]  TRUE TRUE
## [2,] FALSE TRUE

U[upper.tri(U, diag = TRUE)] <- 1
U

##      [,1] [,2]
## [1,]    1    1
## [2,]    0    1
```

Definizione 1.1.18 (Matrici uguali)**.** Si dice di due matrici che hanno lo stesso numero di righe e colonne e hanno uguali tutti gli elementi corrispondenti.

Definizione 1.1.19 (Matrici della base canonica)**.** Data una matrice $m \times n$ le matrici della base canonica sono le $m \cdot n$ matrici $\{\mathbf{E}_{ij}\}$, con $1 \leq i \leq m$, $i \leq j \leq n$ di tipo $m \times n$ che ha 1 come elemento nel posto (i, j) e tutti gli altri elementi nulli.

Esempio 1.1.6. Le matrici della base canonica 2×2 sono

$$\mathbf{E}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \mathbf{E}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \mathbf{E}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \mathbf{E}_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Definizione 1.1.20 (Matrice colonna (vettore))**.** Anche detto vettore, si tratta di una matrice $m \times 1$. I vettori si denotano con lettere minuscole in grassetto: $\mathbf{v}$, $\mathbf{u}$, $\mathbf{w}$.

Definizione 1.1.21 (Matrice riga (vettore riga))**.** Si tratta di una matrice $1 \times n$. I vettori riga vengono denotati come: $\mathbf{v}^T, \mathbf{u}^T, \mathbf{w}^T$.

Osservazione 6. La notazione con T come esponente sarà chiara quando si introdurrà l'operazione di trasposizione.

Definizione 1.1.22 (Coordinate di un vettore)**.** Sono chiamati alternativamente così i coefficienti che lo compongono

Esempio 1.1.7. $\mathbf{w}$ è vettore e $\mathbf{v}^T$ vettore riga

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \quad \mathbf{v}^T = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}$$

I vettori di R sono considerati vettori colonna; per i vettori riga si costruisca una matrice con una riga solamente (oppure si usi la trasposizione su un vettore).

Viceversa quelli di maxima, definiti ad esempio mediante `[1,2,3]`, sono considerati di default vettori riga e occorre applicare `transpose` per avere il vettore colonna

```
##
## [1,2,3]
##                                    [1, 2, 3]
## transpose([1,2,3])
##                                      [ 1 ]
##                                      [   ]
##                                      [ 2 ]
##                                      [   ]
##                                      [ 3 ]
```

È possibile accedere ai vettori che compongono una matrice utilizzando le funzioni `row` e `col`, specificando matrice e indice di riga/colonna desiderata, mentre per accedere ad elementi specifici della matrice si usa una sintassi simile ad R

```
##
## A:matrix([1,2,3],[4,5,6])
## row(A,2)
##                                   [ 4  5  6 ]
## col(A,1)
##                                      [ 1 ]
##                                      [   ]
##                                      [ 4 ]
## A[2,2]
##                                        5
```

Definizione 1.1.23 (Insieme dei vettori colonna)**.** Indichiamo con $\mathbb{C}^n$ l'insieme dei vettori colonna con n coordinate, dotato dell'operazione di prodotto per scalari e somma tra vettori.

Definizione 1.1.24 (Insieme dei vettori riga)**.** Speculare, identificato con $\mathbb{C}_n$,

Osservazione 7 (Matrice come vettore di vettori). Una matrice di tipo (m, n) può essere vista mettendo in evidenza i suoi vettori riga o colonna; nel primo caso la matrice può essere riscritta come vettore colonna di m componenti, con ciascuno di questi costituente un vettore riga di n componenti; nel secondo caso corrisponde ad un vettore riga di n componenti, ciascuno di questi un vettore colonna di m elementi

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_n \end{bmatrix}$$

Definizione 1.1.25 (Vettore nullo)**.** Vettore che ha tutte coordinate nulle, indicato con $\mathbf{0}$.

Definizione 1.1.26 (Vettori coordinati)**.** Vettori con tutte le coordinate nulle tranne una, la i-esima che è uguale a 1. Si indicano con $\mathbf{e}_i$ dove i è la coordinata unitaria.

Esempio 1.1.8.

$$\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{e}_3^T = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$$

Definizione 1.1.27 (Vettori della base canonica)**.** Data una dimensione n si dice base canonica l'insieme di n vettori coordinati $\mathbf{e}_i$ per $i = 1, \dots n$

Esempio 1.1.9. Per $n = 3$, la base canonica dei vettori è costituita dalla tripletta

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Esempio 1.1.10. Le colonne della matrice identità $\mathbf{I}_n$ sono i vettori della base canonica di $\mathbb{C}^n$

$$\mathbf{I}_n = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix}$$

Analogamente, le righe di $\mathbf{I}_n$ sono $\mathbf{e}_1^T, \dots, \mathbf{e}_n^T$.

Definizione 1.1.28 (Traccia)**.** Se $\mathbf{A} = [a_{ij}]$ è una matrice $n \times n$, la traccia di $\mathbf{A}$ è il numero:

$$\operatorname{Tr} \mathbf{A} = \sum_{i=1}^{n} a_{ii} \tag{1.6}$$

ossia la somma degli elementi sulla diagonale principale.

1.2 Algebra delle matrici

1.2.1 Prodotto per scalare

Definizione 1.2.1 (Prodotto di matrice per scalare)**.** Dato uno scalare α e una matrice $\mathbf{A}$ il prodotto $\alpha\mathbf{A}$, si effettua moltiplicando gli elementi della matrice per lo scalare:

$$\alpha\mathbf{A} = [\alpha a_{ij}] \tag{1.7}$$

Proposizione 1.2.1 (Proprietà del prodotto per scalare)**.**

$$\alpha\mathbf{A} = \mathbf{A}\alpha \tag{1.8}$$
$$1\mathbf{A} = \mathbf{A} \tag{1.9}$$
$$0\mathbf{A} = \mathbb{O} \tag{1.10}$$
$$(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A}) \tag{1.11}$$

Proposizione 1.2.2 (Traccia del prodotto per scalare)**.** *Se* $\mathbf{A}$ *è quadrata e* $\alpha \in \mathbb{C}$ *si ha*

$$\operatorname{Tr}(\alpha\mathbf{A}) = \alpha \operatorname{Tr}(\mathbf{A})$$

Dimostrazione. Banale □

Definizione 1.2.2 (Matrice opposta)**.** Matrice ottenuta da quella di partenza, moltiplicandola per -1: l'opposta di $\mathbf{A}$ si indica con $-\mathbf{A}$.

Dimostrazione. Derivano facilmente dalle proprietà del prodotto di due scalari □

Esempio 1.2.1. La matrice nulla e la matrice identità possono essere riscritte come matrici diagonali; più in generale, tutti i multipli scalari della matrice identità sono matrici diagonali:

$$\lambda \mathbf{I} = \mathbf{Diag}(\lambda, \ldots, \lambda)$$

Per il prodotto per scalari si usa `*`, per la matrice opposta `-`

```
i * A

##       [,1] [,2]
## [1,] -1+1i 1+2i
## [2,]  0+3i 1+0i

-C

##      [,1] [,2] [,3] [,4]
## [1,]   -1   -2   -3   -4
## [2,]   -5   -6   -7   -8
## [3,]   -9  -10  -11  -12
```

In maxima si usa il punto sia per la moltiplicazione scalare che, come si vedrà per quella vettoriale

```
##
## A:matrix([1,2,3],[4,5,6])
## %i . A
##                                 [  %i   2 %i  3 %i ]
##                                 [                  ]
##                                 [ 4 %i  5 %i  6 %i ]
## -A
##                                   [ - 1  - 2  - 3 ]
##                                   [               ]
##                                   [ - 4  - 5  - 6 ]
```

1.2.2 Somma di due matrici

Definizione 1.2.3 (Somma di due matrici)**.** Se $\mathbf{A}$ e $\mathbf{B}$ sono entrambi due matrici $m \times n$ la loro somma $\mathbf{A} + \mathbf{B}$ è la matrice $m \times n$ che ha come elemento di posto (i, j) la somma $a_{ij} + b_{ij}$

$$\mathbf{A} + \mathbf{B} = [a_{ij} + b_{ij}] \tag{1.12}$$

Proposizione 1.2.3 (Proprietà della somma di due matrici)**.**

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C} \tag{1.13}$$
$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \tag{1.14}$$
$$\mathbf{A} + \mathbb{O} = \mathbf{A} \tag{1.15}$$
$$\mathbf{A} + (-\mathbf{A}) = \mathbb{O} \tag{1.16}$$
$$\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B} \tag{1.17}$$
$$(\alpha + \beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A} \tag{1.18}$$

Proposizione 1.2.4 (Traccia della somma tra due matrici)**.** *Se* $\mathbf{A}, \mathbf{B}$ *sono quadrate di ordine* n

$$\mathrm{Tr}(\mathbf{A} + \mathbf{B}) = \mathrm{Tr}(\mathbf{A}) + \mathrm{Tr}(\mathbf{B})$$

Dimostrazione. Banale □

Definizione 1.2.4 (Differenza di due matrici)**.** La differenza di due matrici è calcolabile come la somma della prima matrice con la matrice opposta della seconda:

$$\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B}) \tag{1.19}$$

Osservazione 8. Qualsiasi matrice può essere espressa in maniera univoca come una combinazione lineare delle matrice della base canonica:

$$\mathbf{A} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} \mathbf{E}_{ij} \tag{1.20}$$

ove il coefficiente a_{ij} è l'elemento di posto (i, j) di $\mathbf{A}$.

Esempio 1.2.2. Una matrice generica 2×2 si scrive come combinazione delle matrici della base canonica per lo spazio delle matrici di tipo $(2, 2)$:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + a_{12} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + a_{21} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + a_{22} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

In R per la somma di matrici si usa semplicemente `+`

```
A <- matrix(c(-1+i, 3i,  1+2i,  1), nrow = 2)
B <- matrix(c( 1+i, 1i, -1+2i, -1), nrow = 2)
A+B

##      [,1] [,2]
## [1,] 0+2i 0+4i
## [2,] 0+4i 0+0i
```

Allo stesso modo `A + B;` per sommare le due matrici in maxima.

1.2.3 Prodotto di due matrici

Osservazione 9. Vi sono diversi modi di definire il prodotto di due matrici. Vediamo il più usato, detto prodotto righe per colonne

Definizione 1.2.5 (Prodotto di vettore riga per vettore (colonna))**.** Dati due vettori, $\mathbf{v}^T$ riga e $\mathbf{u}$ colonna, con lo stesso numero n di elementi

$$\mathbf{v}^T = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \end{bmatrix} \quad \mathbf{u} = \begin{bmatrix} u_{11} \\ u_{21} \\ \dots \\ u_{n1} \end{bmatrix} \tag{1.21}$$

Si definisce il loro prodotto come:

$$\mathbf{v}^T\mathbf{u} = v_{11}u_{11} + v_{12}u_{21} + \dots + v_{1n}u_{n1} = \sum_{j=1}^{n} v_{1j}u_{j1} \tag{1.22}$$

Osservazione 10. Tale prodotto non restituisce un vettore bensì uno scalare (ossia non è operazione interna all'insieme dei vettori).

Proposizione 1.2.5 (Proprietà del prodotto di vettori)**.**

$$\mathbf{v}^T\mathbf{u} = \mathbf{u}^T\mathbf{v} \tag{1.23}$$

$$\mathbf{v}^T\mathbf{0} = 0 \tag{1.24}$$

Esempio 1.2.3. Il prodotto di un vettore per l'i-esimo coordinato restituisce il valore dell'i-esima coordinata del vettore di partenza.

Esempio 1.2.4. Il prodotto di un vettore per un altro con coordinate unitarie restituisce la somma degli elementi del primo vettore.

Definizione 1.2.6 (Matrici conformi per il prodotto)**.** Lo sono due matrici $\mathbf{A}, \mathbf{B}$ quando il numero di colonne di $\mathbf{A}$ coincide con il numero di righe di $\mathbf{B}$ (quindi $\mathbf{A}$ è matrice $m \times n$ e $\mathbf{B}$ $n \times p$).

Definizione 1.2.7 (Prodotto righe per colonne)**.** Date due matrici $\mathbf{A}, \mathbf{B}$ conformi per il prodotto (del tipo rispettivamente $m \times n$ e $n \times p$), poniamo in evidenza la matrice $\mathbf{A}$ come colonna di vettori riga, e la $\mathbf{B}$ come riga di vettori colonna:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \dots \\ \mathbf{a}_m^T \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_p \end{bmatrix} \tag{1.25}$$

La matrice prodotto righe per colonne $\mathbf{AB}$, di dimensioni $m \times p$ è definita nella posizione (i, k) per $i \leq m$ e $k \leq p$, come prodotto della i-esima riga di $\mathbf{A}$ per la k-esima colonna di $\mathbf{B}$ (dato che il numero di elementi è lo stesso e il prodotto tra vettori è possibile):

$$[\mathbf{AB}_{ik}] = \left[\mathbf{a}_i^T\mathbf{b}_k\right]_{k\leq p}^{i\leq m} = \left[\sum_{j=1}^{n} a_{ij}b_{jk}\right]_{k\leq p}^{i\leq m} = [a_{i1}b_{1k} + a_{i2}b_{2k} + \dots + a_{in}b_{nk}]_{k\leq p}^{i\leq m} \tag{1.26}$$

Esempio 1.2.5. Date due matrici **A** e **B**

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} i \\ -1 \end{bmatrix}, \quad \mathbf{AB} = \begin{bmatrix} (1+i)i + (2-i)(-1) \\ 3i + i \end{bmatrix} = \begin{bmatrix} -3+2i \\ 4i \end{bmatrix}$$

mentre il prodotto **BA** non è calcolabile perché le matrici non sono in tal caso conformabili.

Esempio 1.2.6. È immediato verificare che il prodotto di due matrici diagonali è ancora diagonale e più precisamente

$$\mathbf{Diag}(\lambda_1, \ldots, \lambda_n)\mathbf{Diag}(\mu_1, \ldots, \mu_n) = \mathbf{Diag}(\lambda_1\mu_1, \ldots, \lambda_n\mu_n)$$

Osservazione 11. Il modo particolare con cui si è definita la moltiplicazione fa sì che per la moltiplicazione tra matrici valgano il maggior numero di proprietà valide per la moltiplicazione tra scalari.

In R per il prodotto righe per colonne si usa l'operatore `%*%`:

```
A <- rmatrix(c(1 + i, 2 - i,
                      3,   - i), nrow = 2)
B <- rmatrix(c( i,
                 -1),  nrow = 2)
A %*% B

##        [,1]
## [1,] -3+2i
## [2,]  0+4i

B %*% A ## .. e ovviamente

## Error in B %*% A: gli argomenti non sono compatibili
```

Per quello componente per componente si usa `*` mentre per quello tensoriale (o di Kronecker) `%x%`.

In maxima il prodotto righe per colonne si ottiene mediante il punto, mentre quello element-wise mediante l'asterisco;

```
##
## A:matrix([1+%i,2-%i],[3,-i%])
##                               [ %i + 1  2 - %i ]
##                               [                ]
##                               [   3      - i%  ]
## B:matrix([%i],[-1])
##                                     [ %i  ]
##                                     [     ]
##                                     [ - 1 ]
## A . B
##                            [ %i (%i + 1) + %i - 2 ]
##                            [                      ]
##                            [      i% + 3 %i       ]
## expand(A . B)
##                                  [ 2 %i - 3  ]
##                                  [           ]
##                                  [ i% + 3 %i ]
## A*2
##                           [ 2 (%i + 1)  2 (2 - %i) ]
##                           [                        ]
##                           [     6         - 2 i%   ]
```

Proprietà che valgono

Proposizione 1.2.6 (Prodotto per matrice nulla)**.** *Se* **A** *è una matrice* $n \times p$*,*

$$\mathbb{O}_{mn}\mathbf{A} = \mathbb{O}_{mp} \tag{1.27}$$

$$\mathbf{A}\mathbb{O}_{pq} = \mathbb{O}_{nq} \tag{1.28}$$

Proposizione 1.2.7 (Prodotto per matrice identità)**.** *Se* **A** *è una matrice* $n \times p$*,*

$$\mathbf{I}_n\mathbf{A} = \mathbf{A} = \mathbf{A}\mathbf{I}_p \tag{1.29}$$

Osservazione 12. La matrice identità è l'elemento neutro rispetto al prodotto di matrici. Sia **A** con n colonne e **B** con n righe, si ha

$$\mathbf{A}\mathbf{I}_n = \mathbf{A}$$
$$\mathbf{I}_n\mathbf{B} = \mathbf{B}$$

Proposizione 1.2.8 (Associativa 1)**.** *Se* **A** *e* **B** *sono conformi per il prodotto e* α *è uno scalare*

$$\alpha(\mathbf{AB}) = (\alpha\mathbf{A})\mathbf{B} = \mathbf{A}(\alpha\mathbf{B}) \tag{1.30}$$

Dimostrazione. Infatti, guardando all'elemento di posto i, k

$$\alpha\left[\sum_{j=1}^{n} a_{ij}b_{jk}\right] = \left[\sum_{j=1}^{n} \alpha a_{ij}b_{jk}\right] = \left[\sum_{j=1}^{n}(\alpha a_{ij})b_{jk}\right] = \left[\sum_{j=1}^{n} a_{ij}(\alpha b_{jk})\right]$$

□

Proposizione 1.2.9 (Associativa 2)**.** *Se* $\mathbf{A}$, $\mathbf{B}$ *e* $\mathbf{C}$ *sono matrici rispettivamente* $m \times n$, $n \times p$, $p \times q$

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C} \tag{1.31}$$

Dimostrazione. Sia $[a_{ij}] = \mathbf{A}$, $[b_{ij}] = \mathbf{B}$ e $[c_{ij}] = \mathbf{C}$. Sia inoltre $[d_{ij}] = \mathbf{A}(\mathbf{BC})$ e $[e_{ij}] = (\mathbf{AB})\mathbf{C}$. Per la proprietà distributiva risulta:

$$d_{ij} = \sum_{h=1}^{n} a_{ih} \sum_{k=1}^{p} b_{hk} c_{kj} = \sum_{k=1}^{p} c_{kj} \sum_{h=1}^{n} a_{ih} b_{hk} = \sum_{h=1}^{n} a_{ih} b_{hk} \sum_{k=1}^{p} c_{kj} = e_{ij}$$

□

Osservazione 13. Come nel caso del prodotto di scalari, la proprietà associativa ci consente di scrivere $\mathbf{ABC}$ al posto di $(\mathbf{AB})\mathbf{C}$ e $\mathbf{A}(\mathbf{BC})$ poiché il risultato non dipende da quale prodotto eseguiamo per primo.

Proposizione 1.2.10 (Distributiva a destra)**.** *Se* $\mathbf{A}$ *è una matrice* $m \times n$, $\mathbf{B}$ *e* $\mathbf{C}$ *sono* $n \times p$*:*

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \tag{1.32}$$

Dimostrazione. Sia $[d_{ij}] = \mathbf{A}(\mathbf{B} + \mathbf{C})$ ed $[e_{ij}] = \mathbf{AB} + \mathbf{AC}$. Risulta

$$d_{ij} = \sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj}) = \sum_{k=1}^{n} a_{ik} b_{kj} + \sum_{k=1}^{n} a_{ik} c_{kj} = e_{ij}$$

□

Proposizione 1.2.11 (Distributiva a sinistra)**.** *Se* $\mathbf{A}$ *e* $\mathbf{B}$ *sono* $m \times n$, *e* $\mathbf{C}$ *è* $n \times p$*:*

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} \tag{1.33}$$

Dimostrazione. Simile alla dimostrazione precedente □

Osservazione 14. Si noti che le proprietà distributive sono due e una non segue dall'altra perché il prodotto di matrici non è commutativo.

Esempio 1.2.7. Il lettore provi come esercizio che il prodotto di matrici scalari (rispettivamente: diagonali, triangolari superiori, triangolari inferiori) dello stesso ordine è ancora una matrice scalare (rispettivamente: diagonale, triangolare superiore, triangolare inferiore) con coefficienti diagonali i prodotto dei corrispondenti coefficienti diagonali dei due fattori.

Proposizione 1.2.12 (Traccia del prodotto di due matrici)**.** *Siano* $\mathbf{A}$ *e* $\mathbf{B}$ *conformabili per il prodotto in entrambi i sensi. Si ha*

$$\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$$

Dimostrazione. Effettuiamo la dimostrazione nel caso in cui entrambe siano quadrate di ordine n[2]. Si ha che il posto i, i sulla diagonale di $\mathbf{AB}$ è

$$\mathbf{AB}_{i,i} = \sum_{k=1}^{n} a_{ik} b_{ki}$$

[2]Il caso generale con $\mathbf{A}$ di tipo $m \times n$ e $\mathbf{B}$ non è così banale da essere affrontato con gli strumenti ad ora a disposizione

Allora sommando per i

$$\mathrm{Tr}(\mathbf{AB}) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}b_{kj}$$

Specularmente il posto i, i di $\mathbf{BA}$ è

$$\mathbf{BA}_{i,i} = \sum_{k=1}^{n} b_{ik}a_{ki}$$

e analogamente

$$\mathrm{Tr}(\mathbf{BA}) = \sum_{j=1}^{n}\sum_{k=1}^{n} b_{jk}a_{kj}$$

Ma

$$\mathrm{Tr}(\mathbf{BA}) = \sum_{j=1}^{n}\sum_{k=1}^{n} b_{jk}a_{kj} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{kj}b_{jk} = \sum_{k=1}^{n}\sum_{j=1}^{n} a_{kj}b_{jk} = \mathrm{Tr}(\mathbf{AB})$$

□

Proprietà che non valgono

Proposizione 1.2.13 (Annullamento del prodotto)**.** *Non vale per le matrici la legge di annullamento del prodotto: il prodotto di due matrici può essere la matrice nulla senza che nessuno dei fattori sia matrice nulla.*

Osservazione 15. In generale, se si può eseguire $\mathbf{AB}$ non è detto che si possa eseguire $\mathbf{BA}$ (solo se sono matrici quadrate si può); il risultato poi non è detto che sia il medesimo.

Proposizione 1.2.14 (Commutativa)**.** *In generale la moltiplicazione tra matrici non gode della proprietà commutativa, ovvero:*

$$\mathbf{AB} \neq \mathbf{BA} \tag{1.34}$$

Dimostrazione. Sia $\mathbf{A}$ di tipo $m \times n$ e $\mathbf{B}$ di tipo $n \times m$, con $m \neq n$. Le due matrici sono compatibili per sia per il prodotto $\mathbf{AB}$ che per $\mathbf{BA}$: ma $\mathbf{AB}$ è una matrice $m \times m$, mentre $\mathbf{BA}$ $n \times n$, ed essendo $m \neq n$, sicuramente $\mathbf{AB} \neq \mathbf{BA}$. Sia $\mathbf{A}$ di tipo $m \times n$ e $\mathbf{B}$ di tipo $n \times p$: l'uguaglianza non vale dato che il primo prodotto è calcolabile, ma non il secondo. □

Osservazione 16. Le matrici $\mathbf{A}, \mathbf{B}$ per le quali può eventualmente valere la proprietà sono quelle quadrate dello stesso ordine.

Definizione 1.2.8 (Matrici commutabili)**.** Si chiamano così le matrici quadrate $\mathbf{A}$ e $\mathbf{B}$ per le quali è $\mathbf{AB} = \mathbf{BA}$.

Osservazione 17. Si dice anche che $\mathbf{A}$ e $\mathbf{B}$ commutano.

Definizione 1.2.9 (Matrici non commutabili)**.** Si chiamano così le matrici quadrate $\mathbf{A}$ e $\mathbf{B}$ per le quali è $\mathbf{AB} \neq \mathbf{BA}$.

Osservazione 18. Proprio perché il prodotto può dipendere dall'ordine dei fattori, in algebra lineare è necessario distinguere il prodotto a destra dal prodotto a sinistra.

Definizione 1.2.10 (Pre-moltiplicazione)**.** Nel prodotto $\mathbf{BA}$, la matrice $\mathbf{A}$ si dice pre-moltiplicata, ovvero è moltiplicata a sinistra per la matrice $\mathbf{B}$.

Definizione 1.2.11 (Post-moltiplicazione)**.** In $\mathbf{AB}$, $\mathbf{A}$ è post-moltiplicata (moltiplicata a destra) per $\mathbf{B}$.

Esempio 1.2.8 (Pre/post moltiplicazione per matrice diagonale)**.** La pre-moltiplicazione di una matrice $\mathbf{A}$ per una diagonale $\mathbf{D}$ ha l'effetto di moltiplicare ogni riga di $\mathbf{A}$ per il corrispondente elemento diagonale di $\mathbf{D}$; la post-moltiplicazione per $\mathbf{D}$ ha l'effetto di moltiplicare ogni colonna di $\mathbf{A}$ per il corrispondente elemento diagonale di $\mathbf{D}$.

1.2.4 Prodotto di vettore colonna per vettore riga

Sulla base della definizione di prodotto di matrici e considerando che un vettore non è altro che una matrice di tipo $1 \times n$ (vettore riga) o $p \times 1$ (vettore colonna), il prodotto di un vettore colonna per uno riga, a differenza di quello riga per colonna, non produce uno scalare ma una matrice.

In R

```
1:3 %*% rmatrix(1:3, 1)
```

```
##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    2    4    6
## [3,]    3    6    9
```

In maxima i vettori sono di default considerati riga quindi per ottenere lo stesso risultato

```
##
## transpose([1,2,3]) . [1,2,3]
##                                        [ 1  2  3 ]
##                                        [         ]
##                                        [ 2  4  6 ]
##                                        [         ]
##                                        [ 3  6  9 ]
```

1.2.5 Potenze di una matrice quadrata

Dato che possiamo moltiplicare una matrice quadrata con se stessa giungiamo a definire la potenza come segue

Definizione 1.2.12 (Potenza n-esima di matrice)**.** Definita come:

$$\mathbf{A}^n = \underbrace{\mathbf{A}\mathbf{A}\dots\mathbf{A}}_{n \text{ volte}} \tag{1.35}$$

e ponendo, come per i numeri:

$$\mathbf{A}^0 = \mathbf{I} \tag{1.36}$$

$$\mathbf{A}^1 = \mathbf{A} \tag{1.37}$$

Calcoliamo

```
A <- rmatrix(1:4, 2)
matrix_pow(A, 0)
```

```
##      [,1] [,2]
## [1,]    1    0
## [2,]    0    1
```

```
matrix_pow(A, 1)
```

```
##      [,1] [,2]
## [1,]    1    2
## [2,]    3    4
```

```
cbind(matrix_pow(A, 2), matrix_pow(A, 3), matrix_pow(A, 4))
```

```
##      [,1] [,2] [,3] [,4]   [,5]   [,6]
## [1,]    7   10  199  290 165751 241570
## [2,]   15   22  435  634 362355 528106
```

In maxima la potenza di matrice

```
##
## A:matrix([1,2],[3,4])
## A^^0
##                                        [ 1  0 ]
##                                        [      ]
##                                        [ 0  1 ]
## A^^2
##                                       [ 7   10 ]
##                                       [        ]
##                                       [ 15  22 ]
```

1.3 Trasposizione e matrici connesse

1.3.1 Operatori di trasposizione

Definizione 1.3.1 (Trasposizione)**.** Data una matrice $\mathbf{A}$, di tipo $m \times n$ si dice trasposta di $\mathbf{A}$, indicata con $\mathbf{A}^T$, la matrice $(n \times m)$ che si ottiene scambiandone le righe con le colonne. Il coefficiente a_{ij} di $\mathbf{A}$ diviene a_{ji} in $\mathbf{A}^T$.

Esempio 1.3.1.

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \quad \mathbf{C}^T = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{bmatrix} \tag{1.38}$$

Osservazione 19. La trasposta di un vettore colonna è un vettore riga; per questo per indicare il vettore riga si appone la T.

Proposizione 1.3.1 (Proprietà della trasposizione)**.**

$$(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T \tag{1.39}$$

$$(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T \tag{1.40}$$

$$(\mathbf{A}^T)^T = \mathbf{A} \tag{1.41}$$

$$(\mathbf{AC})^T = \mathbf{C}^T\mathbf{A}^T \tag{1.42}$$

Si noti l'inversione dei fattori in quest'ultima.

Dimostrazione. Dimostriamo l'ultima, le altre sono immediate. Sia $\mathbf{A} = [a_{ij}]_{j \le n}^{i \le m}$ e $\mathbf{C} = [c_{jh}]_{j \le p}^{j \le n}$. L'elemento di posto (i,j) di $(\mathbf{AC})^T$ coincide con (j,i) di $\mathbf{AC}$, che è $\sum_{1 \le h \le n} a_{jh}c_{hi}$.
D'altra parte l'elemento di posto (i,j) di $\mathbf{C}^T\mathbf{A}^T$ è dato dal prodotto della i-esima riga di $\mathbf{C}^T$ (ossia la i-esima colonna di $\mathbf{C}$) per la j-esima colonna di $\mathbf{A}^T$ (j-esima riga di $\mathbf{A}$) quindi è $\sum_{1 \le h \le n} c_{hi}a_{jh}$. Per cui le due matrici sono uguali. □

Definizione 1.3.2 (Matrice coniugata)**.** La matrice coniugata di una matrice $\mathbf{A}$, indicata con $\overline{\mathbf{A}}$, è la matrice ottenuta coniugando tutti i coefficienti di $\mathbf{A}$.

$$\overline{\mathbf{A}} = [\overline{a}_{ij}] \tag{1.43}$$

Osservazione 20. Se $\mathbf{A}$ è reale, $\overline{\mathbf{A}} = \mathbf{A}$

Osservazione 21. La coniugata di una coniugata è uguale alla matrice di partenza: $\overline{\overline{\mathbf{A}}} = \mathbf{A}$

Definizione 1.3.3 (H-trasposizione)**.** La matrice H-trasposta di una matrice $\mathbf{A}$ è la coniugata della trasposta (o equivalentemente la trasposta della coniugata); si indica con $\mathbf{A}^H$

$$\mathbf{A}^H = \overline{\mathbf{A}}^T = \overline{\mathbf{A}^T} \tag{1.44}$$

Osservazione 22. La lettera H deriva dal matematico francese Hermite.

Esempio 1.3.2.

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix} \quad \mathbf{A}^H = \begin{bmatrix} 1-i & 3 \\ 2+i & i \end{bmatrix} \tag{1.45}$$

Proposizione 1.3.2 (Proprietà dell'H-trasposizione)**.**

$$(\alpha \mathbf{A})^H = \overline{\alpha} \mathbf{A}^H \tag{1.46}$$

$$(\mathbf{A} + \mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H \tag{1.47}$$

$$(\mathbf{A}^H)^H = \mathbf{A} \tag{1.48}$$

$$(\mathbf{AC})^H = \mathbf{C}^H \mathbf{A}^H \tag{1.49}$$

Si noti l'inversione dei fattori in quest'ultima.

Dimostrazione. Per la dimostrazione dell'ultima

$$(\mathbf{AC})^H = \overline{(\mathbf{AC})^T} \overset{(1)}{=} \overline{\mathbf{C}^T \mathbf{A}^T} \overset{(2)}{=} \overline{\mathbf{C}^T}\,\overline{\mathbf{A}^T} = \mathbf{C}^H \mathbf{A}^H$$

dove in (1) si è usata la 1.42 e in (2) il fatto che $\overline{\mathbf{XY}} = \overline{\mathbf{X}}\,\overline{\mathbf{Y}}$, dato che il coniugato di somma e prodotto di due complessi è rispettivamente la somma e il prodotto dei loro coniugati. □

La trasposizione si ottiene con `t`, il coniugio con `Conj`, l'H-trasposizione con `ht` (che è una combinazione dei primi due)

```
## Trasposta
t(rmatrix(1:12, 3))

##      [,1] [,2] [,3]
## [1,]    1    5    9
## [2,]    2    6   10
## [3,]    3    7   11
## [4,]    4    8   12

## H-trasposta
ht(rmatrix(c(1+i, 2-i, 3, -i), nrow = 2))

##      [,1] [,2]
## [1,] 1-1i 3+0i
## [2,] 2+1i 0+1i
```

In maxima per l'h-trasposta si usano congiuntamente `transpose` e `conjugate`:

```
##
## A:matrix([1+%i,2-%i],[3,-%i])
## conjugate(transpose(A))
##                                [ 1 - %i   3  ]
##                                [             ]
##                                [ %i + 2   %i ]
```

1.3.2 Matrici connesse alla trasposizione

Osservazione 23. Tra tutte le matrici quadrate rivestono particolare interesse le seguenti.

Definizione 1.3.4 (Matrice simmetrica)**.** Una matrice quadrata $\mathbf{A}$ di ordine n si dice *simmetrica* se coincide con la propria trasposta:

$$\mathbf{A} = \mathbf{A}^T \tag{1.50}$$

In altre parole è simmetrica una matrice quadrata in cui $a_{ij} = a_{ji}$, $\forall i, j \leq n$.

Osservazione 24. La simmetria è rispetto alla diagonale principale: scambiando gli indici di riga e di colonna gli elementi non cambiano

Esempio 1.3.3.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix} = \mathbf{A}^T \tag{1.51}$$

Definizione 1.3.5 (Matrice Hermitiana)**.** Matrice che coincide con la propria H-trasposta

Esempio 1.3.4.

$$\mathbf{A} = \begin{bmatrix} 2 & 1-i & i \\ 1+i & 0 & 3+2i \\ -i & 3-2i & 3 \end{bmatrix} = \mathbf{A}^H \tag{1.52}$$

Osservazione 25. Per le matrici reali le due nozioni di matrice simmetrica e matrice hermitiana vengono a coincidere dato che la coniugazione non modifica una matrice reale. Invece una matrice complessa non reale può risultare simmetrica senza essere hermitiana e viceversa.

Osservazione 26. I coefficienti diagonali di una matrice simmetrica possono essere del tutto arbitrari; quelli di una matrice hermitiana debbono essere reali poiché se no cambierebbero in seguito a coniugio (e debbono invece rimanere uguali).

Definizione 1.3.6 (Matrice anti-simmetrica)**.** Una matrice quadrata $\mathbf{A}$ si dice *anti-simmetrica* se coincide con l'opposta della trasposta:

$$\mathbf{A} = -\mathbf{A}^T \tag{1.53}$$

Analogamente una matrice è anti-simmetrica se quadrata e $a_{ij} = -a_{ji}$, $\forall i, j = 1, 2, \ldots, n$.

Esempio 1.3.5.

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 3 \\ -2 & 0 & 4 \\ -3 & -4 & 0 \end{bmatrix} = -\mathbf{A}^T \tag{1.54}$$

Osservazione 27. Segue dalle definizioni che gli elementi della diagonale di una matrice anti-simmetrica sono nulli

Definizione 1.3.7 (Matrice anti-hermitiana)**.** Matrice che coincide con l'opposta della propria H-trasposta.

Esempio 1.3.6.

$$\mathbf{A} = \begin{bmatrix} i & 1-i & i \\ -1-i & 0 & 3+2i \\ i & -3+2i & -2i \end{bmatrix} = -\mathbf{A}^H \tag{1.55}$$

Osservazione 28. Segue dalle definizioni che gli elementi della diagonale di una matrice anti-hermitiana debbono immaginari ossia del tipo ir con $r \in \mathbb{R}$

Si può testare la simmetricità delle matrici mediante le funzioni `is.symmetric` `is.hermitian` `is.antisymmetric` `is.antihermitian`

```
is.symmetric(rmatrix(c(1,2,3,
                       2,4,5,
                       3,5,6), nrow = 3))

## [1] TRUE
```

Osservazione 29. Vediamo ora alcune proprietà di questi tipi di matrici, la cui verifica è lasciata come esercizio.

Proposizione 1.3.3. *La somma di due matrici simmetriche, hermitiane, anti-simmetriche o anti-hermitiane rimane dello stesso tipo.*

Proposizione 1.3.4. *Se $\mathbf{A}$ è simmetrica (risp. hermitiana) e α è uno scalare (risp. un numero reale), allora $\alpha\mathbf{A}$ è simmetrica (risp. hermitiana)*

Proposizione 1.3.5. *Se $\mathbf{A}$ è anti-simmetrica (risp. anti-hermitiana) e α è uno scalare (risp. un numero reale), allora $\alpha\mathbf{A}$ è anti-simmetrica (risp. anti-hermitiana)*

Proposizione 1.3.6. *$\mathbf{A}$ è hermitiana se e solo se $i\mathbf{A}$ è anti-hermitiana*

Proposizione 1.3.7. *Il prodotto di due simmetriche (risp. hermitiane) è simmetrica (hermitiana) se e solo se le due matrici commutano tra loro;*

Proposizione 1.3.8. *Se $\mathbf{A}$ è una matrice complessa $m \times n$, $\mathbf{A}\mathbf{A}^T$ e $\mathbf{A}^T\mathbf{A}$ sono simmetriche, $\mathbf{A}\mathbf{A}^H$ e $\mathbf{A}^H\mathbf{A}$ hermitiane.*

Dimostrazione. Infatti, a titolo di esempio:

$$(\mathbf{A}\mathbf{A}^T)^T = (\mathbf{A}^T)^T\mathbf{A}^T = \mathbf{A}\mathbf{A}^T$$
$$(\mathbf{A}^H\mathbf{A})^H = \mathbf{A}^H(\mathbf{A}^H)^H = \mathbf{A}^H\mathbf{A}$$

□

1.3.3 Decomposizione di una matrice complessa

Osservazione 30. Come ogni numero complesso $z = a + ib$ può esser scritto come la somma di una parte reale e di una immaginaria, ogni matrice quadrata complessa $\mathbf{A}$ si può decomporre univocamente come $\mathbf{A} = \mathbf{B} + \mathbf{C}$ con $\mathbf{B}$ matrice hermitiana (detta *parte hermitiana*) e $\mathbf{C}$ (parte) anti-hermitiana.

Proposizione 1.3.9 (Decomposizione di matrice complessa). *Ogni matrice complessa quadrata* $\mathbf{A}$ *si scrive in uno e un solo modo nella forma* $\mathbf{A} = \mathbf{B} + \mathbf{C}$ *con* $\mathbf{B}$ *matrice hermitiana e* $\mathbf{C}$ *anti-hermitiana. E ciò avviene come segue:*

$$\mathbf{A} = \underbrace{\left[\frac{\mathbf{A} + \mathbf{A}^H}{2}\right]}_{\mathbf{B}} + \underbrace{\left[\frac{\mathbf{A} - \mathbf{A}^H}{2}\right]}_{\mathbf{C}} \tag{1.56}$$

Dimostrazione. Una verifica diretta mostra che $\mathbf{B}$ è hermitiana

$$\mathbf{B}^H = \left[\frac{\mathbf{A} + \mathbf{A}^H}{2}\right]^H = \overline{\left(\frac{1}{2}\right)}(\mathbf{A} + \mathbf{A}^H)^H = \frac{1}{2}(\mathbf{A}^H + \mathbf{A}) = \mathbf{B}$$

Similmente $\mathbf{C}$ è anti-hermitiana:

$$\mathbf{C}^H = \left[\frac{\mathbf{A} - \mathbf{A}^H}{2}\right]^H = \overline{\left(\frac{1}{2}\right)}(\mathbf{A} - \mathbf{A}^H)^H = \frac{1}{2}(\mathbf{A}^H - \mathbf{A}) = -\frac{\mathbf{A}^H - \mathbf{A}}{2} = -\mathbf{C}$$

per cui se $\mathbf{C}^H = -\mathbf{C}$, allora $\mathbf{C} = -\mathbf{C}^H$ come si voleva. Inoltre $\mathbf{B} + \mathbf{C} = \mathbf{A}$

$$\mathbf{B} + \mathbf{C} = \frac{\mathbf{A}}{2} + \frac{\mathbf{A}^H}{2} + \frac{\mathbf{A}}{2} - \frac{\mathbf{A}^H}{2} = 2 \cdot \frac{\mathbf{A}}{2} = \mathbf{A}$$

Quanto all'unicità ponendo per ipotesi che $\mathbf{A}$ si possa scrivere anche in un altro modo, ossia $\mathbf{A} = \mathbf{B}' + \mathbf{C}'$, con $\mathbf{B}'$ hermitiana e $\mathbf{C}'$ anti-hermitiana (e ovviamente $\mathbf{B}' \neq \mathbf{B}$, $\mathbf{C}' \neq \mathbf{C}$) risulta che, sfruttando la diversa uguaglianza che definisce $\mathbf{A}$

$$\mathbf{B} + \mathbf{C} = \mathbf{B}' + \mathbf{C}' \iff \mathbf{B} - \mathbf{B}' = \mathbf{C}' - \mathbf{C}$$

Ma $\mathbf{B} - \mathbf{B}'$ risulta hermitiana e $\mathbf{C}' - \mathbf{C}$ anti-hermitiana; l'unica matrice contemporaneamente hermitiana e anti-hermitiana (data l'uguaglianza dei due termini $\mathbf{B} - \mathbf{B}'$ e $\mathbf{C}' - \mathbf{C}$) è la matrice nulla, perciò $\mathbf{B} - \mathbf{B}' = \mathbb{O}$ e $\mathbf{C} - \mathbf{C}' = \mathbb{O}$, da cui si conclude (contrariamente alle ipotesi dalle quali eravamo partiti) che $\mathbf{B} = \mathbf{B}'$ e $\mathbf{C} = \mathbf{C}'$ □

Osservazione 31. Se si ha a che fare con una matrice reale, la parte hermitiana è reale e simmetrica, quella anti-hermitiana reale anti-simmetrica

Esempio 1.3.7.

$$\mathbf{A} = \begin{bmatrix} 1+i & 2-i \\ 3 & -i \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 2 & 5-i \\ 5+i & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2i & -1-i \\ 1-i & -2i \end{bmatrix} \tag{1.57}$$

Definizione 1.3.8 (Matrice normale). Una matrice quadrata $\mathbf{A}$ si dice normale se commuta con la sua H-trasposta, ossia

$$\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A} \tag{1.58}$$

Proposizione 1.3.10. *Sia* $\mathbf{A} = \mathbf{B} + \mathbf{C}$ *la decomposizione della matrice quadrata* $\mathbf{A}$. *Allora parte hermitiana e anti-hermitiana commutano, ossia*

$$\mathbf{B}\mathbf{C} = \mathbf{C}\mathbf{B} \tag{1.59}$$

se e solo se la matrice $\mathbf{A}$ *è normale.*

Dimostrazione. Poiché $\mathbf{A}^H = \mathbf{B}^H + \mathbf{C}^H$ e

$$\mathbf{B} - \mathbf{C} = \frac{\mathbf{A} + \mathbf{A}^H}{2} - \frac{\mathbf{A} - \mathbf{A}^H}{2} = \frac{\mathbf{A}}{2} + \frac{\mathbf{A}^H}{2} - \frac{\mathbf{A}}{2} + \frac{\mathbf{A}^H}{2} = \mathbf{A}^H$$

ossia anche che $\mathbf{A}^H = \mathbf{B} - \mathbf{C}$, allora si ha che

$$\mathbf{A}\mathbf{A}^H = (\mathbf{B} + \mathbf{C})(\mathbf{B} - \mathbf{C}) = \mathbf{B}^2 - \mathbf{B}\mathbf{C} + \mathbf{C}\mathbf{B} - \mathbf{C}^2$$

mentre

$$\mathbf{A}^H\mathbf{A} = (\mathbf{B} - \mathbf{C})(\mathbf{B} + \mathbf{C}) = \mathbf{B}^2 + \mathbf{B}\mathbf{C} - \mathbf{C}\mathbf{B} - \mathbf{C}^2$$

Perciò $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ se e solo se

$$-\mathbf{B}\mathbf{C} + \mathbf{C}\mathbf{B} = \mathbf{B}\mathbf{C} - \mathbf{C}\mathbf{B} \iff 2\mathbf{B}\mathbf{C} = 2\mathbf{C}\mathbf{B} \iff \mathbf{B}\mathbf{C} = \mathbf{C}\mathbf{B}$$

□

1.4 Sottomatrici e decomposizione a blocchi

1.4.1 Sottomatrici

Definizione 1.4.1 (Sottomatrice)**.** Data una matrice $\mathbf{A}$ di tipo $m \times n$, si dice sottomatrice quella formata dagli elementi che appartengono all'intersezione di $1 \le p \le m$ righe e $1 \le q \le n$ colonne, scelte in qualsiasi modo.

Definizione 1.4.2 (Sottomatrici proprie)**.** $\mathbf{A}$ è sottomatrice di se stessa; le altre sottomatrici di $\mathbf{A}$ si dicono sottomatrici proprie.

Definizione 1.4.3 (Sottomatrici principali)**.** Tra le sottomatrici di una matrice quadrata, importanti sono le *sottomatrici principali*, ottenute scegliendo righe e colonne con gli stessi indici.

Definizione 1.4.4 (Sottomatrice principale k-esima)**.** Sottomatrice delle prime k righe colonne (con $1 \le k \le n$).

Esempio 1.4.1. Alcune sottomatrici principali di

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$$

sono

$$\begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}; \begin{bmatrix} 6 & 7 \\ 10 & 11 \end{bmatrix}; \begin{bmatrix} 6 & 7 & 8 \\ 10 & 11 & 12 \\ 14 & 15 & 16 \end{bmatrix}$$

di cui la prima è la sottomatrice principale seconda.

1.4.2 Decomposizione a blocchi

1.4.2.1 Definizioni ed esempi

Definizione 1.4.5 (Decomposizione a blocchi)**.** Altra maniera di considerare porzioni di matrici è quella di decomporre a blocchi una matrice: ciò si fa tracciando righe orizzontali e/o verticali che tagliano la matrice in sottomatrici chiamate blocchi.

Esempio 1.4.2. La matrice $\mathbf{A}$ di prima può esser decomposta a blocchi ambivalentemente in:

$$\left[\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ \hline 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{array}\right] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

$$\left[\begin{array}{c|cc|c} 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ \hline 13 & 14 & 15 & 16 \end{array}\right] = \begin{bmatrix} \mathbf{A}'_{11} & \mathbf{A}'_{12} & \mathbf{A}'_{13} \\ \mathbf{A}'_{21} & \mathbf{A}'_{22} & \mathbf{A}'_{23} \\ \mathbf{A}'_{31} & \mathbf{A}'_{32} & \mathbf{A}'_{33} \end{bmatrix}$$

con dimensione (a blocchi) rispettivamente 2×2 e 3×3.

Osservazione 32. A volte non si decompone una matrice a blocchi, ma a partire da due matrici se ne definisce una terza (con la stessa notazione a blocchi).

Esempio 1.4.3.

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 6 & 7 \\ 9 & 10 & 11 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 4 \\ 8 \\ 12 \end{bmatrix}, \quad [\mathbf{Ab}] = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

1.4.2.2 Operazioni su matrici decomposte a blocchi

Osservazione 33. L'utilità delle decomposizioni a blocchi sta nel fatto che sulle matrici si possono eseguire le operazioni di somma e prodotto (righe per colonne) per blocchi, ossia considerando i blocchi come fossero coefficienti, a patto che:

- le dimensioni a blocchi delle due matrici siano le stesse (per la somma) oppure le matrici a blocchi siano conformi (per il prodotto);
- le coppie di singoli blocchi (i quali sono matrici a tutti gli effetti) abbiano le stesse dimensioni (nel caso della somma), quelli che si moltiplicano siano conformi rispetto al prodotto

Soddisfatti tali requisiti si può operare a blocchi, ottenendo lo stesso risultato. La giustificazione di questo fatto è del tutto evidente per la somma, mentre per quanto riguarda il prodotto, dipende essenzialmente dalla proprietà associativa della somma

Esempio 1.4.4. Date le due decomposizioni a blocchi delle matrici $\mathbf{A}, \mathbf{B}$

$$\mathbf{A} = \left[\begin{array}{c|cc|c} 1 & 2 & 3 & 1 \\ \hline 0 & -1 & 0 & -2 \\ 3 & 0 & 1 & 2 \end{array}\right] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \end{bmatrix}, \quad \mathbf{B} = \left[\begin{array}{c|c} 2 & 0 \\ \hline -1 & 2 \\ 0 & -1 \\ \hline 3 & -1 \end{array}\right] = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \\ \mathbf{B}_{31} & \mathbf{B}_{32} \end{bmatrix}$$

Notiamo che $\mathbf{A}$ è di tipo 2×3 mentre $\mathbf{B}$ di tipo 3×2; inoltre $\mathbf{A}_{11}$ è conformabile con $\mathbf{B}_{11}$ (entrambe 1×1), $\mathbf{A}_{12}$ con $\mathbf{B}_{21}$, e così via le rimanenti. Possiamo pertanto procedere con il prodotto a blocchi delle due matrici:

$$\begin{aligned}\mathbf{AB} &= \begin{bmatrix}\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{21}+\mathbf{A}_{13}\mathbf{B}_{31} & \mathbf{A}_{11}\mathbf{B}_{12}+\mathbf{A}_{12}\mathbf{B}_{22}+\mathbf{A}_{13}\mathbf{B}_{32}\\ \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{21}+\mathbf{A}_{23}\mathbf{B}_{31} & \mathbf{A}_{21}\mathbf{B}_{12}+\mathbf{A}_{22}\mathbf{B}_{22}+\mathbf{A}_{23}\mathbf{B}_{32}\end{bmatrix}\\ &= \left[\begin{array}{l|r} 3 & 0\\ \hline -5 & 0\\ 12 & -3\end{array}\right]\end{aligned}$$

1.4.2.3 Notazione generale

La notazione usuale per una matrice decomposta a blocchi, di dimensioni a blocchi $m \times n$ è la seguente:

$$\mathbf{A} = \begin{bmatrix}\mathbf{A}_{11} & \dots & \mathbf{A}_{1j} & \dots & \mathbf{A}_{1n}\\ \dots & \dots & \dots & \dots & \dots\\ \mathbf{A}_{i1} & \dots & \mathbf{A}_{ij} & \dots & \mathbf{A}_{in}\\ \dots & \dots & \dots & \dots & \dots\\ \mathbf{A}_{m1} & \dots & \mathbf{A}_{mj} & \dots & \mathbf{A}_{mn}\end{bmatrix} \tag{1.60}$$

Se ciascun blocco $\mathbf{A}_{ij}$ ha dimensioni $m_i \times n_j$, le dimensioni della matrice $\mathbf{A}$ sono $m \times n$ con $m = \sum_i m_i$ e $n = \sum_j n_j$.

1.4.2.4 Altre definizioni e operazioni

Osservazione 34. Molte definizioni date per le matrici si estendono alle matrici a blocchi

Definizione 1.4.6 (Matrice triangolare superiore a blocchi)**.** Si chiama così la matrice a blocchi $m \times n$ $\mathbf{A}_{ij}$ se $\mathbf{A}_{ij} = \mathbb{O}$ per $i > j$

Definizione 1.4.7 (Matrice triangolare inferiore a blocchi)**.** Definizione speculare alla precedente

Definizione 1.4.8 (Matrice diagonale a blocchi)**.** Matrice a blocchi in cui i blocchi non residenti sulla diagonale principale corrispondono a $\mathbb{O}$. Indicata con $\mathbf{Diag}(\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_n)$, o meno frequentemente $\mathbf{A}_1 \oplus \mathbf{A}_2 \oplus \dots \oplus \mathbf{A}_n$

Definizione 1.4.9 (Trasposizione/H-trasposizione)**.** Le trasposizioni si effettuano scambiando tra loro righe e colonne di blocchi, con l'avvertenza di trasporre (e se necessario coniugare) ciascun blocco

Esempio 1.4.5. Data

$$\mathbf{A} = \begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13}\\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23}\end{bmatrix}$$

Si ha

$$\mathbf{A}^T = \begin{bmatrix}\mathbf{A}_{11}^T & \mathbf{A}_{21}^T\\ \mathbf{A}_{12}^T & \mathbf{A}_{22}^T\\ \mathbf{A}_{13}^T & \mathbf{A}_{23}^T\end{bmatrix}, \quad \mathbf{A}^H = \begin{bmatrix}\mathbf{A}_{11}^H & \mathbf{A}_{21}^H\\ \mathbf{A}_{12}^H & \mathbf{A}_{22}^H\\ \mathbf{A}_{13}^H & \mathbf{A}_{23}^H\end{bmatrix}$$

1.4.2.5 Matrice in forma bordata

Si tratta di una decomposizione molto usata nelle dimostrazioni, consiste nel mettere in evidenza la prima riga e la prima colonna (oppure l'ultima riga e l'ultima colonna).

Definizione 1.4.10 (Matrice in forma bordata)**.** Una matrice $\mathbf{A}$, di tipo $m \times n$ definita come

$$\mathbf{A} = \begin{bmatrix} a & \mathbf{u}^T \\ \mathbf{v} & \mathbf{B} \end{bmatrix} \tag{1.61}$$

con a coefficiente di posto $(1,1)$, $\mathbf{u}^T$ è la prima riga privata del primo coefficiente, $\mathbf{v}$ è la prima colonna privata del primo coefficiente e $\mathbf{B}$ si ottiene da $\mathbf{A}$ eliminando prima riga e prima colonna.

1.4.2.6 Letture alternative del prodotto tra matrici

Osservazione 35. La decomposizione a blocchi ci permette di vedere il prodotto di matrici come combinazione lineare di differenti elementi delle matrici di partenza.

Osservazione 36. Consideriamo due matrici $\mathbf{A}, \mathbf{B}$ conformi per il prodotto di dimensioni $m \times n$ e $n \times p$. Mettiamo in evidenza in entrambe righe e colonne

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \mathbf{s}_1^T \\ \mathbf{s}_2^T \\ \dots \\ \mathbf{s}_n^T \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_p \end{bmatrix}$$

Possiamo eseguire la moltiplicazione a blocchi in modi diversi:

1. sfruttando le righe di $\mathbf{A}$ e le colonne di $\mathbf{B}$, ottenendo di fatto la definizione del prodotto righe per colonne:

   $$\mathbf{AB} = [\mathbf{r}_i^T \mathbf{b}_j]_{j \le p}^{i \le m}$$

   ossia ciascuno di questi prodotti produce uno scalare che va a riempire la casella corrispondente nella matrice risultante;

2. usando le colonne di $\mathbf{A}$ e per le righe di $\mathbf{B}$;

   - se sia $\mathbf{A}$ che $\mathbf{B}$ sono matrici, un prodotto colonna per riga produce una matrice e $\mathbf{AB}$ è la somma delle n matrici prodotte in questo modo

     $$\mathbf{AB} = \sum_{i \le n} \mathbf{a}_i \mathbf{s}_i^T$$

   - nel caso che $\mathbf{A}$ sia costituita da un unico vettore riga $\mathbf{r}_1^T$ l'applicazione della precedente conduce a:

     $$\mathbf{r}_1^T \mathbf{B} = a_{11}\mathbf{s}_1^T + a_{12}\mathbf{s}_2^T + \dots a_{1n}\mathbf{s}_n^T = \sum_{i \le n} a_{1i}\mathbf{s}_i^T \tag{1.62}$$

     In questo caso il prodotto delle due matrici è una somma di vettori riga, o meglio è una combinazione lineare delle righe di $\mathbf{B}$ con coefficienti le coordinate di $\mathbf{A} = \mathbf{r}_1^T$.
     Da questo deriva che se $\mathbf{e}_i^T$ è un vettore riga coordinato ($i \le m$), il prodotto $\mathbf{e}_i^T \mathbf{B}$ coincide con la i-esima riga di $\mathbf{B}$;

- specularmente se $\mathbf{B}$ è un vettore colonna $\mathbf{b}_1$

$$\mathbf{A}\mathbf{b}_1 = \mathbf{a}_1 b_{11} + \mathbf{a}_2 b_{21} + \ldots + \mathbf{a}_n b_{n1} = \sum_{i \leq n} \mathbf{a}_i b_{i1} \tag{1.63}$$

  si ha che $\mathbf{A}\mathbf{b}_1$ è una somma di vettori colonna, o meglio combinazione lineare delle colonne di $\mathbf{A}$ con coefficienti le coordinate di $\mathbf{b}_1$.
  Da questo deriva che se $\mathbf{e}_j$ è un vettore coordinato ($j \leq n$), il prodotto $\mathbf{A}\mathbf{e}_j$ coincide con la j-esima colonna di $\mathbf{A}$;

3. considerando $\mathbf{A}$ come unico blocco e decomponendo $\mathbf{B}$ in colonne, ricavando

$$\mathbf{A}\begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2 & \ldots & \mathbf{b}_p\end{bmatrix} = \begin{bmatrix}\mathbf{A}\mathbf{b}_1 & \mathbf{A}\mathbf{b}_2 & \ldots & \mathbf{A}\mathbf{b}_p\end{bmatrix} \tag{1.64}$$

   Ossia nel prodotto (righe per colonne) $\mathbf{AB}$ di due matrici, la j-esima colonna di $\mathbf{AB}$ coincide con il prodotto $\mathbf{A}\mathbf{b}_j$ ossia di $\mathbf{A}$ per la j-esima colonna $\mathbf{b}_j$ di $\mathbf{B}$;

4. specularmente considerando $\mathbf{B}$ come unico blocco e usando le righe di $\mathbf{A}$

$$\begin{bmatrix}\mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \ldots \\ \mathbf{r}_m^T\end{bmatrix}\mathbf{B} = \begin{bmatrix}\mathbf{r}_1^T\mathbf{B} \\ \mathbf{r}_2^T\mathbf{B} \\ \ldots \\ \mathbf{r}_m^T\mathbf{B}\end{bmatrix} \tag{1.65}$$

   Ossia nel prodotto (righe per colonne) $\mathbf{AB}$ di due matrici, la i-esima riga di $\mathbf{AB}$ coincide con il prodotto $\mathbf{r}_i^T\mathbf{B}$ ossia della i-esima riga $\mathbf{r}_i^T$ di $\mathbf{A}$ per $\mathbf{B}$.

1.5 Eliminazione di Gauss

L'algoritmo di eliminazione di Gauss (EG) serve per la risoluzione di un sistema in m equazioni per n incognite:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{cases} \tag{1.66}$$

L'idea base è eliminare progressivamente nelle equazioni del sistema sempre più incognite, ottenendo un sistema equivalente a quello di partenza, ma che si può risolvere facilmente mediante sostituzione all'indietro.

1.5.1 L'algoritmo

Definizione 1.5.1 (Rappresentazione in termini matriciali)**.** Un sistema di m equazioni per n incognite tipo 1.66 si può rappresentare, in forma matriciale, come:

$$\mathbf{Ax} = \mathbf{b} \tag{1.67}$$

ossia tramite l'uguaglianza del *vettore dei termini noti* $\mathbf{b}$ con il vettore $\mathbf{Ax}$, ottenuto dal prodotto righe per colonne della *matrice dei coefficienti* $\mathbf{A}$ del sistema per il *vettore delle incognite* $\mathbf{x}$.

Definizione 1.5.2 (Matrice dei coefficienti del sistema)**.**

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{1.68}$$

Definizione 1.5.3 (Vettore delle incognite)**.**

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \tag{1.69}$$

Definizione 1.5.4 (Vettore dei termini noti)**.**

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix} \tag{1.70}$$

Definizione 1.5.5 (Matrice aumentata)**.** È la matrice a blocchi $[\mathbf{Ab}]$.

Osservazione 37 (Funzionamento dell'algoritmo). L'obiettivo è applicare ripetutamente i principi di equivalenza concentrandosi solo sui "numeri" del sistema (ovvero sulla sua matrice aumentata), al fine di eliminare incognite (ossia rendere il relativo coefficiente della matrice pari a 0) sino a che si possa operare per sostituzione a ritroso.

Osservazione 38 (Principi di equivalenza e funzionamento dell'algoritmo). Sono i medesimi dei sistemi di equazioni:

1. moltiplicazione di una equazione (riga) per uno scalare non nullo (in $\mathbb{C}$: ricordarsi che si può moltiplicare per un reale, ma anche per immaginari derivati da i, o veri e propri complessi);
2. sostituzione di una equazione (riga) con la sua somma membro a membro con un'altra equazione (riga) del sistema;
3. scambio fra loro due equazioni (righe)

Osservazione 39. Per convenzione, nello svolgimento dell'algoritmo, lo scambio di righe serve per tenere in ordine la matrice su cui si opera, ossia in configurazione a scala per righe.

Definizione 1.5.6 (Configurazione a scala per righe)**.** Ad ogni riga a partire dalla seconda si abbiano un numero di zeri iniziali superiori alla riga precedente.

Esempio 1.5.1 (Sistema determinato)**.** Il sistema

$$\begin{cases} 2x_1 + 6x_2 + 3x_3 = 4 \\ x_1 - 2x_2 + \frac{1}{2}x_3 = 1 \\ -x_1 + x_2 - \frac{1}{2}x_3 = \frac{2}{5} \end{cases}$$

ha come matrice aumentata

$$\left[\begin{array}{ccc|c} 2 & 6 & 3 & 4 \\ 1 & -2 & 1/2 & 1 \\ -1 & 1 & -1/2 & 2/5 \end{array}\right]$$

Operiamo su essa come detto

$$\left[\begin{array}{ccc|c} 1 & 3 & 3/2 & 2 \\ 1 & -2 & 1/2 & 1 \\ -1 & 1 & -1/2 & 2/5 \end{array}\right]; \left[\begin{array}{ccc|c} 1 & 3 & 3/2 & 2 \\ 0 & -5 & -1 & -1 \\ 0 & 4 & 1 & 12/5 \end{array}\right]; \left[\begin{array}{ccc|c} 1 & 3 & 3/2 & 2 \\ 0 & 1 & 1/5 & 1/5 \\ 0 & 4 & 1 & 12/5 \end{array}\right]$$
$$\left[\begin{array}{ccc|c} 1 & 3 & 3/2 & 2 \\ 0 & 1 & 1/5 & 1/5 \\ 0 & 0 & 1/5 & 8/5 \end{array}\right]; \left[\begin{array}{ccc|c} 1 & 3 & 3/2 & 2 \\ 0 & 1 & 1/5 & 1/5 \\ 0 & 0 & 1 & 8 \end{array}\right] = [\mathbf{Uc}]$$

L'ultima matrice corrisponde al sistema

$$\begin{cases} x_1 + 3x_2 + \frac{3}{2}x_3 = 2 \\ x_2 + \frac{1}{5}x_3 = \frac{1}{5} \\ x_3 = 8 \end{cases}$$

Per cui a questo punto la EG è terminata e le soluzioni del sistema si possono facilmente trovare per sostituzione all'indietro, ottenendo

$$x_3 = 8, \quad x_2 = -\frac{7}{5} \quad x_1 = -\frac{29}{5}$$

Definizione 1.5.7 (Forma ridotta della matrice $\mathbf{A}$)**.** Si chiama così la matrice uni-triangolare superiore $\mathbf{U}$, ottenuta da $\mathbf{A}$ al termine della EG.

Definizione 1.5.8 (Forma ridotta della matrice aumentata)**.** Si chiama così la matrice $[\mathbf{Uc}]$ ottenuta da $[\mathbf{Ab}]$ al termine della EG

La risoluzione di un sistema con $\mathbf{A}$ quadrata (numero equazioni uguale al numero incognite) si effettua mediante `solve`:

```
A <- rmatrix(c(1,  3 , 3/2,
               1, -2 , 1/2,
               -1, 1, -1/2), nrow = 3)
colnames(A) <- paste0('x', 1:ncol(A))
b <- c(2, 1, 2/5)
solve(A, b)

##   x1   x2   x3
## -5.8 -1.4  8.0

## check risultati
-29/5; -7/5

## [1] -5.8
## [1] -1.4
```

In maxima:

- `triangularize` effettua l'EG ma non normalizza il pivot di ciascuna riga ad essere 1
- `echelon` EG e normalizza il pivot porta riduce la matrice in row echelon form (ossia a scala per righe con pivot unitari)

```
##
## M:matrix([3,4,a],[3,3,b])
##                                [ 3  4  a ]
##                                [         ]
##                                [ 3  3  b ]
## echelon(M)
##                              [    4    a   ]
##                              [ 1  -    -   ]
##                              [    3    3   ]
##                              [             ]
##                              [ 0  1  a - b ]
```

Esempio 1.5.2 (Sistema indeterminato)**.** Il sistema

$$\begin{cases} 2x_1 + 6x_2 + 3x_3 + 2x_4 = 4 \\ x_1 - 2x_2 + \frac{1}{2}x_3 + \frac{9}{4}x_4 = 1 \\ -x_1 + x_2 - \frac{1}{2}x_3 - x_4 = \frac{2}{5} \end{cases}$$

conduce alla matrice aumentata e all'applicazione di EG come segue

$$\left[\begin{array}{cccc|c} 2 & 6 & 3 & 2 & 4 \\ 1 & -2 & \frac{1}{2} & \frac{9}{4} & 1 \\ -1 & 1 & -\frac{1}{2} & -1 & \frac{2}{5} \end{array}\right]; \left[\begin{array}{cccc|c} 1 & 3 & 3/2 & 1 & 2 \\ 1 & -2 & \frac{1}{2} & \frac{9}{4} & 1 \\ -1 & 1 & -\frac{1}{2} & -1 & \frac{2}{5} \end{array}\right]; \left[\begin{array}{cccc|c} 1 & 3 & 3/2 & 1 & 2 \\ 0 & -5 & -1 & \frac{5}{4} & -1 \\ 0 & 4 & 1 & 0 & \frac{12}{5} \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & 3 & 3/2 & 1 & 2 \\ 0 & 1 & 1/5 & -\frac{1}{4} & 1/5 \\ 0 & 4 & 1 & 0 & \frac{12}{5} \end{array}\right]; \left[\begin{array}{cccc|c} 1 & 3 & 3/2 & 1 & 2 \\ 0 & 1 & 1/5 & -\frac{1}{4} & 1/5 \\ 0 & 0 & 1/5 & 1 & \frac{8}{5} \end{array}\right]; \left[\begin{array}{cccc|c} 1 & 3 & 3/2 & 1 & 1 \\ 0 & 1 & 1/5 & -1/4 & 1/5 \\ 0 & 0 & 1 & 5 & 8 \end{array}\right] = [\mathbf{Vc}]$$

L'ultima riga di $[\mathbf{Vc}]$ corrisponde all'equazione

$$x_3 + 5x_4 = 8$$

Dando all'incognita x_4 un qualunque valore h risulta

$$x_4 = h, \quad x_3 = 8 - 5h, \quad x_2 = -\frac{7}{5} + \frac{5}{4}h, \quad x_1 = -\frac{29}{5} + \frac{11}{4}h$$

Pertanto il sistema ha infinite soluzioni, dipendenti dal parametro h, che possono essere scritte compattamente come:

$$\begin{bmatrix} -29/5 \\ -7/5 \\ 8 \\ 0 \end{bmatrix} + h \begin{bmatrix} 11/4 \\ 5/4 \\ -5 \\ 1 \end{bmatrix}$$

Esempio 1.5.3 (Sistema impossibile)**.** Il sistema

$$\begin{cases} 2x_1 + 6x_2 + 3x_3 = 4 \\ x_1 - 2x_2 + \frac{1}{2}x_3 = 1 \\ -x_1 + x_2 - \frac{7}{10}x_3 = \frac{2}{5} \end{cases}$$

conduce a ha come matrice aumentata

$$\left[\begin{array}{ccc|c} 2 & 6 & 3 & 4 \\ 1 & -2 & 1/2 & 1 \\ -1 & 1 & -7/10 & 2/5 \end{array}\right]; \ldots ; \left[\begin{array}{ccc|c} 1 & 3 & 3/2 & 2 \\ 0 & 1 & 1/5 & 1/5 \\ 0 & 0 & 0 & 1 \end{array}\right] = [\mathbf{U}'\mathbf{c}]$$

Con l'ultima riga corrispondente all'equazione

$$0x_3 = 1$$

che evidentemente non ha soluzione. Pertanto il sistema non ha soluzione.

1.5.2 Soluzioni e matrice ridotta

Osservazione 40. Il sistema può avere nessuna, una o infinite soluzioni, a seconda della matrice in forma ridotta cui si giunge al termine della EG. Risulta utile introdurre la seguente definizione.

Definizione 1.5.9 (Colonna dominante)**.** In una matrice in forma a scala per righe si dice di una colonna che contiene il primo coefficiente non nullo di qualche riga.

Esempio 1.5.4. Nella

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 0 \end{bmatrix} \tag{1.71}$$

sono dominanti la prima e la seconda colonna, in quanto contengono 1 e 4.

Osservazione 41. Per individuarle, nella matrice in forma a scala seguire ciascuna riga sino a che si trova il primo coefficiente non nullo, dopodiché salire marcare la colonna come dominante.

Esempio 1.5.5. Sia

$$\mathbf{A}_\alpha = \begin{bmatrix} i & 0 & -i & i\alpha \\ 1 & \alpha^2+4 & 0 & \alpha \\ 1 & \alpha^2+4 & 0 & 2\alpha \end{bmatrix}$$

con $\alpha \in \mathbb{C}$. Per ogni $\alpha \in \mathbb{C}$ si trovi una forma ridotta di Gauss $\mathbf{U}_\alpha$ di $\mathbf{A}_\alpha$ e si dica quali sono le colonne dominanti e quali no.

$$\begin{bmatrix} i & 0 & -i & i\alpha \\ 1 & \alpha^2+4 & 0 & \alpha \\ 1 & \alpha^2+4 & 0 & 2\alpha \end{bmatrix} \begin{bmatrix} i & 0 & -i & i\alpha \\ 1 & \alpha^2+4 & 0 & \alpha \\ 0 & 0 & 0 & \alpha \end{bmatrix} \begin{bmatrix} 1 & \alpha^2+4 & 0 & \alpha \\ 0 & \alpha^2+4 & i & \alpha(1-i) \\ 0 & 0 & 0 & \alpha \end{bmatrix}$$

Se $\alpha^2 + 4 = 0 \iff \alpha = \pm 2i$ sono dominanti la prima, la terza e la quarta colonna; se $\alpha = 0$ dominante prima e seconda; altrimenti sono dominanti la prima seconda e la quarta.

Definizione 1.5.10 (Variabili dominanti)**.** Variabili (di $\mathbf{A}$) corrispondenti alle colonne dominanti

Definizione 1.5.11 (Variabili libere)**.** Variabili rimanenti (di $\mathbf{A}$) rispetto alle dominanti

Proposizione 1.5.1 (Soluzioni di un sistema)**.** *Se* $\mathbf{A}$ *è una matrice* $m \times n$ *non nulla, considerando la forma ridotta* $[\mathbf{Uc}]$ *della matrice aumentata* $[\mathbf{Ab}]$*:*

- *se l'ultima colonna è dominante, il sistema* ***non ammette soluzioni****. La forma ridotta assume una forma del tipo*

$$[\mathbf{Uc}] = \left[\begin{array}{ccccccc|c} \mathbf{0}^T & 1 & \dots & + & \dots & + & \dots & c_1 \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & + & \dots & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & c_k \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & c_{k+1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 \end{array}\right]$$

  con "+" a indicare coefficienti arbitrari e $c_{k+1} \neq 0$*.*
  Questo caso può accadere indipendentemente dalle dimensioni delle matrici considerate (ovvero dal numero di equazioni ed incognite nella matrice $\mathbf{A}$*)*

- *se tutte le colonne sono dominanti ad eccezione dell'ultima (colonna dei termini noti), il sistema ammette* ***una e una sola soluzione****. La forma ridotta è del tipo:*

$$[\mathbf{Uc}] = \left[\begin{array}{cccccc|c} 1 & + & + & \dots & + & + & c_1 \\ 0 & 1 & + & \dots & + & + & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & + & c_{n-1} \\ 0 & 0 & 0 & \dots & 0 & 1 & c_n \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{array}\right]$$

  Il presente caso può accadere solo se $m \geq n$ *(ovvero il numero di equazioni del sistema è superiore o uguale al numero di incognite).*

  Definizione 1.5.12 (Pivot)**.** Nella EG si possono anche non eseguire le moltiplicazioni che rendono uguali a 1 i primi coefficienti non nulli di ogni riga; in tal caso i numeri $d_i \neq 0$ che compaiono come primi coefficienti non nulli di ogni riga si chiamano pivot

- *se l'ultima colonna non è dominante ed esiste almeno un'altra colonna non dominante, il sistema ammette* ***infinite soluzioni*** *che dipendono da tanti* *parametri quante sono le variabili libere. La forma ridotta è del tipo*

$$[\mathbf{Uc}] = \left[\begin{array}{ccccccc|c} \mathbf{0}^T & 1 & \dots & + & \dots & + & \dots & c_1 \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & + & \dots & c_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 1 & \dots & c_k \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 & \mathbf{0}^T & 0 \end{array}\right]$$

dove compare effettivamente qualcuno dei vettori riga $\mathbf{0}^T$ in qualcuna delle prime k righe (ovvero ci sono variabili libere) e ci sono $k \leq n$ righe non nulle.

Esempio 1.5.6. Sia

$$\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \alpha & \alpha \\ 0 & 1 & 1 \\ 1 & \alpha+1 & \alpha+1 \end{bmatrix}$$

Si discuta, al variare di $\alpha \in \mathbb{C}$ il sistema omogeneo $\mathbf{A}_\alpha \mathbf{x} = \mathbf{0}$, determinandone tutte le soluzioni $\forall \alpha \in \mathbb{C}$.
Si ha:

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & \alpha & \alpha & 0 \\ 0 & 1 & 1 & 0 \\ 1 & \alpha+1 & \alpha+1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & \alpha & \alpha & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & \alpha-1 & \alpha-1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

da cui

$$\begin{cases} x_2 + x_3 = 0 \\ (\alpha-1)x_2 + (\alpha-1)x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases} \quad \begin{cases} x_1 = 0 \\ x_2 + x_3 = 0 \\ (\alpha-1)(x_2+x_3) = 0 \end{cases} \quad \begin{cases} x_1 = 0 \\ x_2 = -x_3 \\ (\alpha-1)(x_2+x_3) = 0 \end{cases}$$

Ora $\forall \alpha \in \mathbb{C}$, si ha che $x_1 = 0$, poniamo $x_3 = h$, si ha $x_2 = -h$. Le soluzioni pertanto sono

$$\begin{bmatrix} 0 \\ -h \\ h \end{bmatrix} = \left\langle \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\rangle$$

Esempio 1.5.7. Risolvere il sistema lineare $\mathbf{Ax} = \mathbf{b}$ con

$$\mathbf{A} = \begin{bmatrix} 2 & -2 & 6 & 4 \\ 1 & -1 & 7 & 4 \\ 3 & -3 & 9 & 6 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

Si ha:

$$\begin{bmatrix} 2 & -2 & 6 & 4 & 3 \\ 1 & -1 & 7 & 4 & 2 \\ 3 & -3 & 9 & 6 & 1 \end{bmatrix} \quad r_3 = r_3 - r_1$$

$$\begin{bmatrix} 2 & -2 & 6 & 4 & 3 \\ 1 & -1 & 7 & 4 & 2 \\ 1 & -1 & 3 & 2 & -2 \end{bmatrix} \quad r_3 = r_3 - r_2$$

$$\begin{bmatrix} 2 & -2 & 6 & 4 & 3 \\ 1 & -1 & 7 & 4 & 2 \\ 0 & 0 & -4 & -2 & -4 \end{bmatrix} \quad r_1 = r_1 - 2r_2,\ r_3 = r_3/(-2)$$

$$\begin{bmatrix} 0 & 0 & -8 & -4 & -1 \\ 1 & -1 & 7 & 4 & 2 \\ 0 & 0 & 2 & 1 & 2 \end{bmatrix} \quad r_3 = 4r_3 + r_1$$

$$\begin{bmatrix} 0 & 0 & -8 & -4 & -1 \\ 1 & -1 & 7 & 4 & 2 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix}$$

Da cui risulterebbe $0x_1 + 0x_2 + 0x_3 + 0x_4 = 7$ per cui il sistema è impossibile.

Esempio 1.5.8. Risolvere il sistema lineare $\mathbf{A}_\alpha \mathbf{x} = \mathbf{b}_\alpha$ dove

$$\mathbf{A}_\alpha = \begin{bmatrix} 2 & 2\alpha & 2 \\ 1 & \alpha+1 & \alpha+1 \\ 1 & \alpha & \alpha \\ 0 & 2 & 2\alpha \end{bmatrix}, \mathbf{b}_\alpha = \begin{bmatrix} 2\alpha \\ \alpha+1 \\ \alpha \\ \alpha^2+1 \end{bmatrix}$$

Si ha:

$$\begin{bmatrix} 2 & 2\alpha & 2 & 2\alpha \\ 1 & \alpha+1 & \alpha+1 & \alpha+1 \\ 1 & \alpha & \alpha & \alpha \\ 0 & 2 & 2\alpha & \alpha^2+1 \end{bmatrix} \quad r_1 = r_1 - r_3,\ r_2 = r_2 - r_3$$

$$\begin{bmatrix} 1 & \alpha & 2-\alpha & \alpha \\ 0 & 1 & 1 & 1 \\ 1 & \alpha & \alpha & \alpha \\ 0 & 2 & 2\alpha & \alpha^2+1 \end{bmatrix} \quad r_1 = r_1 - r_3$$

$$\begin{bmatrix} 0 & 0 & 2-\alpha & 0 \\ 0 & 1 & 1 & 1 \\ 1 & \alpha & \alpha & \alpha \\ 0 & 2 & 2\alpha & \alpha^2+1 \end{bmatrix} \quad r_4 = r_4 - 2r_2$$

$$\begin{bmatrix} 0 & 0 & 2-\alpha & 0 \\ 0 & 1 & 1 & 1 \\ 1 & \alpha & \alpha & \alpha \\ 0 & 0 & 2\alpha-2 & \alpha^2-1 \end{bmatrix} \quad r_1 = r_1 + r_4$$

$$\begin{bmatrix} 0 & 0 & 0 & \alpha^2-1 \\ 0 & 1 & 1 & 1 \\ 1 & \alpha & \alpha & \alpha \\ 0 & 0 & 2\alpha-2 & \alpha^2-1 \end{bmatrix} \quad \ldots \text{riordino}$$

$$\begin{bmatrix} 1 & \alpha & \alpha & \alpha \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 2\alpha-2 & \alpha^2-1 \\ 0 & 0 & 0 & \alpha^2-1 \end{bmatrix}$$

A questo punto i casi rilevanti sono $\alpha = 1$, $\alpha = -1$ e i casi rimanenti.
Se $\alpha = -1$ il sistema è determinato e ha come soluzione:

$$\begin{bmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{cases} -4x_3 = 0 \\ x_2 + x_3 = 1 \\ x_1 - x_2 - x_3 = -1 \end{cases} \quad \begin{cases} x_3 = 0 \\ x_2 = 1 \\ x_1 - 1 = -1 \end{cases} \quad \begin{cases} x_1 = 0 \\ x_2 = 1 \\ x_3 = 0 \end{cases} \rightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Se $\alpha = 1$ il sistema è indeterminato e le soluzioni sono

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{cases} x_3 = k \\ x_2 + k = 1 \\ x_1 + x_2 + k = 1 \end{cases} \quad \begin{cases} x_3 = k \\ x_2 = 1 - k \\ x_1 + 1 - k + k = 1 \end{cases} \quad \begin{cases} x_1 = 0 \\ x_2 = 1 - k \\ x_3 = k \end{cases} \rightarrow \begin{bmatrix} 0 \\ 1-k \\ k \end{bmatrix}, k \in \mathbb{C}$$

Nei rimanenti casi è impossibile poiché dovrebbe essere

$$0x_1 + 0x_2 + 0x_3 = \alpha^2 - 1$$

ma dato che $\alpha \neq \pm 1$, $\alpha^2 - 1 \neq 0$, e quindi l'equazione non può essere soddisfatta.

1.5.3 Rango

Definizione 1.5.13 (Rango)**.** Numero di righe non nulle nella forma ridotta $\mathbf{U}$.

Osservazione 42. Il rango di $\mathbf{U}$ coincide col numero delle sue colonne dominanti; nel caso di un sistema determinato è n.

Osservazione 43. Seguirà da quanto vedremo sugli spazi vettoriali che ogni forma ridotta cui si perviene mediante la EG da una data matrice ha sempre lo stesso rango.

In maxima si usa la funzione `rank`

```
##
## M:matrix([1,2,3],[2,4,6],[4,9,7])
##                                [ 1  2  3 ]
##                                [         ]
##                                [ 2  4  6 ]
##                                [         ]
##                                [ 4  9  7 ]
## triangularize(M)
##                               [ 1  2   3  ]
##                               [           ]
##                               [ 0  1  - 5 ]
##                               [           ]
##                               [ 0  0   0  ]
## rank(M)
##                                     2
```

1.6 Matrici inverse e pseudo-inverse

Osservazione 44. L'inverso di un numero a, $\frac{1}{a}$, è quel numero che moltiplicato per il primo risulta 1. Vogliamo trovare un equivalente di a^{-1} nell'insieme delle matrici. Il problema è più complesso poiché la moltiplicazione nelle matrici non gode della proprietà commutativa e ciò porta alle seguenti definizioni.

Osservazione 45. L'esistenza di inverse della matrice $\mathbf{A}$ nel sistema $\mathbf{Ax} = \mathbf{b}$, come si avrà modo di vedere, determina il numero di soluzioni del sistema.

Definizione 1.6.1 (Inversa sinistra)**.** Sia $\mathbf{A}$ matrice di tipo $m \times n$: si chiama inversa sinistra di $\mathbf{A}$ una matrice $\mathbf{L}$ (da left) di tipo $n \times m$ tale che:

$$\mathbf{LA} = \mathbf{I}_n$$

Definizione 1.6.2 (Inversa destra)**.** Sia $\mathbf{A}$ matrice di tipo $m \times n$: si chiama inversa destra di $\mathbf{A}$ la matrice $\mathbf{R}$ di tipo $n \times m$ tale che:

$$\mathbf{AR} = \mathbf{I}_m$$

Osservazione 46. Sia $\mathbf{L}$ che $\mathbf{R}$ sono matrici $n \times m$. Ciò è necessario per la fattibilità del prodotto e al contempo fa sì che il risultato sia una matrice quadrata ($\mathbf{I}$, appunto).

Osservazione 47. Una matrice può ammettere anche più di una inversa destra o sinistra.

Definizione 1.6.3 (Matrice inversa)**.** Si definisce così (o anche inversa bilatera), e si indica con $\mathbf{A}^{-1}$, una matrice che è contemporaneamente sia inversa destra che sinistra di una matrice data $\mathbf{A}$.

Definizione 1.6.4 (Matrice invertibile (o non singolare))**.** $\mathbf{A}$ si dice invertibile (o non singolare) se esiste la sua inversa $\mathbf{A}^{-1}$.

Definizione 1.6.5 (Matrice singolare)**.** Sinonimo di non invertibile.

Osservazione 48. Come si vedrà, la matrice inversa $\mathbf{A}^{-1}$ può esistere solo se $\mathbf{A}$ è quadrata.

Osservazione 49. La matrice inversa se esiste è unica come provato dalla seguente proposizione.

Proposizione 1.6.1. *Sia* $\mathbf{A}$ *una matrice* $m \times n$*. Se* $\mathbf{A}$ *ha sia inversa destra* $\mathbf{R}$ *che sinistra* $\mathbf{L}$*, allora* $\mathbf{R} = \mathbf{L}$*. Quindi* $\mathbf{R} = \mathbf{L}$ *è l'unica inversa (destra, sinistra, bilatera) di* $\mathbf{A}$*.*

Dimostrazione.

$$\mathbf{L} = \mathbf{L}\mathbf{I}_m = \mathbf{L}(\mathbf{A}\mathbf{R}) = (\mathbf{L}\mathbf{A})\mathbf{R} = \mathbf{I}_n\mathbf{R} = \mathbf{R}$$

dove ci si è limitato ad effettuare sostituzioni (giustificate dal fatto che $\mathbf{A}$ abbia opportune inverse) e a sfruttare le proprietà del prodotto tra matrici. □

Osservazione 50. Nel caso delle matrici quadrate si ha una situazione particolare in merito ad inverse destre e sinistra.
Premettiamo il seguente risultato

Lemma 1.6.2. *Data una qualunque matrice* $\mathbf{A}$*:*

1. *esiste una matrice invertibile* $\mathbf{E}$ *tale che il prodotto* $\mathbf{EA}$ *ha la seguente forma bordata:*
$$\mathbf{EA} = \begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix}$$
2. *esiste una matrice invertibile* $\mathbf{F}$ *tale che* $\mathbf{FA} = \mathbf{U}$*, dove* $\mathbf{U}$ *è una forma ridotta di* $\mathbf{A}$

Dimostrazione. Per ora prendiamoli per buoni: il primo verrà giustificato in 1.8.10, il secondo nelle fattorizzazioni dei teoremi 1.8.7 ($\mathbf{A} = \mathbf{LU}$ da cui $\mathbf{F} = \mathbf{L}^{-1}$) e 1.8.8 ($\mathbf{A} = \mathbf{P}^{-1}\mathbf{LU}$ da cui $\mathbf{F} = (\mathbf{P}^{-1}\mathbf{L})^{-1}$). □

Proposizione 1.6.3. *Sia* $\mathbf{A}$ *una matrice quadrata di ordine* n*. Allora una inversa destra di* $\mathbf{A}$ *è anche inversa sinistra, e viceversa.*

Osservazione 51. In altre parole, se $\mathbf{A}$ ha inversa destra, allora questa è anche inversa sinistra (o viceversa); e quindi è inversa bilaterale (l'unica).

Dimostrazione. Dimostriamo innanzitutto che una inversa destra è anche sinistra. Sia $\mathbf{AR} = \mathbf{I}_n$ con $\mathbf{A}$ quadrata di ordine n e ragioniamo per induzione su n. Se $n = 1$ l'asserto è ovvio: ad esempio se $\mathbf{A} = \begin{bmatrix} a \end{bmatrix}$, allora $\mathbf{R} = \begin{bmatrix} a^{-1} \end{bmatrix}$ e $\mathbf{AR} = \mathbf{RA} = \begin{bmatrix} 1 \end{bmatrix}$. Sia allora $n > 1$ e supponiamo che l'asserto valga per $n - 1$.

Usiamo la forma bordata di $\mathbf{EA}$ data dal punto 1 del lemma 1.6.2, che vale per una opportuna matrice invertibile $\mathbf{E}$. Osserviamo che

$$\mathbf{AR} = \mathbf{I}_n \overset{(1)}{\iff} \mathbf{EAR} = \mathbf{E} \overset{(2)}{\iff} \mathbf{E}\underbrace{\mathbf{AR}}_{\mathbf{I}_n}\mathbf{E}^{-1} = \mathbf{EE}^{-1} = \mathbf{I}_n$$

dove in (1) abbiamo pre-moltiplicato per $\mathbf{E}$ entrambi i membri (dato che $\mathbf{A}$ e $\mathbf{I}_n$ sono conformabili con $\mathbf{E}$, avendo le stesse dimensioni), mentre in (2) abbiamo post-moltiplicato per $\mathbf{E}^{-1}$ dato che l'inversa di $\mathbf{E}$ esiste.
Decomponiamo anche $\mathbf{RE}^{-1}$ in forma bordata, ossia battezziamola come

$$\mathbf{RE}^{-1} = \begin{bmatrix} b & \mathbf{y}^T \\ \mathbf{z} & \mathbf{Y} \end{bmatrix}$$

Dall'uguaglianza di poco fa, scritta per esteso

$$\underbrace{\begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix}}_{\mathbf{EA}}\underbrace{\begin{bmatrix} b & \mathbf{y}^T \\ \mathbf{z} & \mathbf{Y} \end{bmatrix}}_{\mathbf{RE}^{-1}} = \underbrace{\begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}_{n-1} \end{bmatrix}}_{\mathbf{I}_n}$$

si ricavano le uguaglianze

$$ab + \mathbf{x}^T\mathbf{z} = 1, \quad a\mathbf{y}^T + \mathbf{x}^T\mathbf{Y} = \mathbf{0}^T, \quad \mathbf{Xz} = \mathbf{0}, \quad \mathbf{XY} = \mathbf{I}_{n-1}$$

Poiché le matrici $\mathbf{X}$ e $\mathbf{Y}$ sono quadrate e hanno ordine $n-1$ per l'ipotesi induttiva risulta $\mathbf{YX} = \mathbf{I}_{n-1}$ quindi $\mathbf{Y} = \mathbf{X}^{-1}$ (cioè abbiamo ipotizzato che per $n-1$ valga il fatto che per una matrice quadrata di ordine $n-1$ una inversa destra sia anche sinistra e qui $\mathbf{X}$ è quadrata e il loro prodotto è $\mathbf{I}$ quindi deve essere che $\mathbf{Y} = \mathbf{X}^{-1}$) e $\mathbf{z} = \mathbf{I}_{n-1}\mathbf{z} = \mathbf{YXz} = \mathbf{Y0} = \mathbf{0}$. Ne consegue che $ab = 1$ e $\mathbf{y}^T = -a^{-1}\mathbf{x}^T\mathbf{X}^{-1}$ quindi

$$\mathbf{RE}^{-1} = \begin{bmatrix} a^{-1} & -a^{-1}\mathbf{x}^T\mathbf{X} \\ \mathbf{0} & \mathbf{X}^{-1} \end{bmatrix}$$

Basta allora a questo punto eseguire la moltiplicazione $\mathbf{RA} = \mathbf{RE}^{-1}\mathbf{EA}$ a blocchi:

$$\underbrace{\begin{bmatrix} a^{-1} & -a^{-1}\mathbf{x}^T\mathbf{X}^{-1} \\ \mathbf{0} & \mathbf{X}^{-1} \end{bmatrix}}_{\mathbf{RE}^{-1}}\underbrace{\begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix}}_{\mathbf{EA}} = \begin{bmatrix} a^{-1}a & a^{-1}\mathbf{x}^T - a^{-1}\mathbf{x}^T\mathbf{X}^{-1}\mathbf{X} \\ \mathbf{0} & \mathbf{X}^{-1}\mathbf{X} \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{I}_{n-1} \end{bmatrix}}_{\mathbf{I}_n}$$

per ottenere che $\mathbf{RA} = \mathbf{I}_n$, ossia che inversa destra $\mathbf{R}$ è anche inversa sinistra di $\mathbf{A}$.
Per il fatto invece che inversa sinistra è anche destra, se $\mathbf{LA} = \mathbf{I}_n$, usando quanto appena visto e il rapporto tra trasposizione e prodotto si ricava:

$$\begin{aligned}\mathbf{LA} = \mathbf{I}_n &\iff (\mathbf{LA})^T = \mathbf{I}_n^T \iff \mathbf{A}^T\mathbf{L}^T = \mathbf{I}_n \overset{(1)}{\iff} \mathbf{L}^T\mathbf{A}^T = \mathbf{I}_n \iff (\mathbf{AL})^T = \mathbf{I}_n \\ &\iff \mathbf{AL} = \mathbf{I}_n\end{aligned}$$

dove in (1) abbiamo usato il fatto che dato che $\mathbf{L}^T$ è inversa destra di $\mathbf{A}^T$ allora è anche inversa di sinistra e si può scrivere $\mathbf{L}^T\mathbf{A}^T = \mathbf{I}_n$ □

Osservazione 52. Prima di procedere oltre vediamo che relazione c'è, in un sistema $\mathbf{Ax} = \mathbf{b}$, tra l'esistenza di inverse per $\mathbf{A}$ e la presenza di soluzioni per il sistema

Proposizione 1.6.4. *Il sistema* $\mathbf{Ax} = \mathbf{b}$ *ammette:*

1. almeno una *soluzione se* $\mathbf{A}$ *ammette inversa* destra*;*
2. al più una *soluzione se* $\mathbf{A}$ *ammette inversa* sinistra*;*
3. una e una sola *soluzione se* $\mathbf{A}$ *ammette inversa* bilatera.

Dimostrazione. Infatti:

1. se $\mathbf{A}$ ammette inversa destra $\mathbf{R}$, il vettore $\mathbf{Rb}$ risulta essere una soluzione (quindi ve ne è almeno una) poiché:

$$\mathbf{A}(\mathbf{Rb}) = (\mathbf{AR})\mathbf{b} = \mathbf{Ib} = \mathbf{b}$$

2. se $\mathbf{A}$ ammette inversa sinistra $\mathbf{L}$, e se il sistema ammette soluzioni, allora la soluzione è unica. Ipotizzando per assurdo che ve ne siano due soluzioni differenti $\mathbf{u}$ e $\mathbf{v}$, dovendo soddisfare il sistema si ha:

$$\mathbf{Au} = \mathbf{b} = \mathbf{Av}$$

Moltiplicando a sinistra ambo i membri per $\mathbf{L}$ si ricava:

$$\mathbf{u} = \mathbf{LAu} = \mathbf{Lb} = \mathbf{LAv} = \mathbf{v}$$

quindi si va a contraddire l'ipotesi che $\mathbf{v}$ e $\mathbf{u}$ fossero differenti;

3. la 3 (una e una sola soluzione se bilatera) segue da 1 (almeno una se destra) e 2 (al massimo uno se sinistra).
Nel caso, la soluzione è chiaramente $\mathbf{A}^{-1}\mathbf{b}$, in quanto soddisfa l'equivalenza:

$$\mathbf{A}(\mathbf{A}^{-1}\mathbf{b}) = \mathbf{b}$$
$$(\mathbf{AA}^{-1})\mathbf{b} = \mathbf{b}$$

□

Osservazione 53. La proposizione fornisce condizioni sufficienti

1.6.1 Caratterizzazione delle matrici con inversa

Osservazione 54. Sotto quali condizioni una matrice ammette inversa destra o sinistra?

Osservazione 55. I risultati seguenti mostrano come:

- una matrice può avere inversa destra solo se è "orizzontale", cioè con un numero di colonne non inferiore a quello delle righe
- specularmente una matrice può avere inversa sinistra solo se è "verticale"
- una matrice può avere inversa solo se quadrata.

Proposizione 1.6.5. *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *dotata di inversa destra. Allora* $m \leq n$.

Dimostrazione. Sia $\mathbf{AR} = \mathbf{I}_m$. Se fosse $m > n$ si potrebbero decomporre a blocchi $\mathbf{A}$ e $\mathbf{R}$ nel modo seguente

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \end{bmatrix}; \quad \mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \end{bmatrix}$$

con $\mathbf{A}_1$ e $\mathbf{R}_1$ blocchi quadrati di ordine n (da cui $\mathbf{A}_2$ di tipo $(m-n) \times n$ e $\mathbf{R}_2$ di tipo $n \times (m-n)$). Moltiplicando $\mathbf{A}$ e $\mathbf{R}$ a blocchi e uguagliando a $\mathbf{I}_m$ (decomposta a blocchi in modo conforme) si ottiene

$$\mathbf{AR} = \begin{bmatrix} \mathbf{A}_1\mathbf{R}_1 & \mathbf{A}_1\mathbf{R}_2 \\ \mathbf{A}_2\mathbf{R}_1 & \mathbf{A}_2\mathbf{R}_2 \end{bmatrix} = \mathbf{I}_m = \begin{bmatrix} \mathbf{I}_m & \mathbb{O}_1 \\ \mathbb{O}_2 & \mathbf{I}_{m-n} \end{bmatrix}$$

con $\mathbb{O}_1$ di tipo $n \times (m-n)$ e $\mathbb{O}_2$ di tipo $(m-n) \times n$. Quindi si ottiene

$$\mathbf{A}_1\mathbf{R}_1 = \mathbf{I}_n, \quad \mathbf{A}_1\mathbf{R}_2 = \mathbb{O}_1, \quad \mathbf{A}_2\mathbf{R}_1 = \mathbb{O}_2, \quad \mathbf{A}_2\mathbf{R}_2 = \mathbf{I}_n$$

Dalla proposizione 1.6.3, visto che $\mathbf{A}_1$ è quadrata, sappiamo che $\mathbf{A}_1\mathbf{R}_1 = \mathbf{R}_1\mathbf{A}_1 = \mathbf{I}_n$, quindi:

$$\mathbf{A}_2\mathbf{R}_1 = \mathbb{O}_2 \overset{(1)}{\iff} \mathbf{A}_2\mathbf{R}_1\mathbf{A}_1 = \mathbb{O}_2 \iff \mathbf{A}_2 = \mathbb{O}_2$$

dove in (1) abbiamo post-moltiplicato per $\mathbf{A}_1$ (di tipo $n \times n$) entrambi i membri dato che sono conformabili e sostituito con $\mathbf{I}_n$.
Se $\mathbf{A}_2 = \mathbb{O}_2$ non può essere vera l'uguaglianza $\mathbf{A}_2\mathbf{R}_2 = \mathbf{I}_{m-n}$ e quindi $\mathbf{R}$ non è inversa di $\mathbf{A}$ dato che il loro prodotto non restituisce $\mathbf{I}$. □

Proposizione 1.6.6. *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *dotata di inversa sinistra. Allora* $m \geq n$.

Dimostrazione. $\mathbf{A}$ (di tipo $m \times n$) ammette come inversa sinistra la matrice $\mathbf{L}$ (di tipo $n \times m$) se e solo se $\mathbf{A}^T$ (tipo $n \times m$) ammette come inversa destra la matrice $\mathbf{L}^T$ ($m \times n$). Ciò deriva dal fatto che se $\mathbf{LA} = \mathbf{I}$, applicando la trasposizione ad entrambi i membri e le relative proprietà di trasposizione e prodotto si ha:

$$\mathbf{LA} = \mathbf{I} \iff (\mathbf{LA})^T = \mathbf{I}^T \iff \mathbf{A}^T\mathbf{L}^T = \mathbf{I}^T$$

Si conclude applicando ad $\mathbf{A}^T$ la proposizione 1.6.5. □

Proposizione 1.6.7. *Congiuntamente se una matrice* $\mathbf{A}$ *ha inversa destra e inversa sinistra, allora è una matrice quadrata invertibile.*

Dimostrazione. Sia $\mathbf{A}$ $m \times n$ per la proposizione 1.6.5 risulta $m \leq n$, per la 1.6.6 $m \geq n$, quindi $m = n$. E dato che quadrata la proposizione 1.6.1 assicura che è invertibile (in quanto inversa destra e sinistra coincidono). □

Osservazione 56. Caratterizziamo ora le matrici che ammettono inversa destra.

Teorema 1.6.8. *Per una matrice* $\mathbf{A}$ *di tipo* $m \times n$ *le seguenti proprietà sono equivalenti:*

1. $\mathbf{A}$ *ammette inversa destra;*

2. *il sistema* $\mathbf{Ax} = \mathbf{b}$ *ammette almeno una soluzione per ogni scelta del vettore* $\mathbf{b}$*;*

3. $\mathbf{A}$ *ha rango* m

Dimostrazione. Rispettivamente:

- $1 \implies 2$: segue da proposizione 1.6.4, punto 1
- $2 \implies 3$: se per assurdo $\mathbf{A}$ avesse rango $< m$, una sua forma ridotta $\mathbf{U}$ avrebbe l'ultima riga nulla. Indicando con $\mathbf{e}_m$ l'm-esimo vettore coordinato con m coordinate, nella matrice $\begin{bmatrix}\mathbf{U}\mathbf{e}_m\end{bmatrix}$ l'ultima colonna sarebbe dominante quindi detto sistema non avrebbe soluzioni. Procedendo a ritroso, a partire dalla matrice $[\mathbf{U}\mathbf{e}_m]$ con le operazioni elementari inverse di quelle eseguite sulla matrice $\mathbf{A}$ per ottenere la forma ridotta $\mathbf{U}$, si per perviene alla matrice aumentata $[\mathbf{A}\mathbf{b}]$ di un sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$, equivalente al sistema $\mathbf{U}\mathbf{x} = \mathbf{e}_m$, che non ha soluzione, il che è assurdo (cioè avevamo ipotizzato avesse soluzioni e invece concludiamo per assurdo il contrario). Quindi $\mathbf{A}$ ha rango m.
- $3 \implies 1$: si tratta di provare che esiste una matrice $n \times m$ $\mathbf{R} = [\mathbf{r}_1\,\mathbf{r}_2\,\ldots\,\mathbf{r}_m]$ (della quale sono state messe in evidenza le colonne) tale che $\mathbf{A}\mathbf{R} = \mathbf{I}_m$; in altre parole, spezzando e vedendo colonna per colonna, deve essere $\mathbf{A}\mathbf{r}_j = \mathbf{e}_j$ per ogni $j \leq m$. Bisogna trovare quindi soluzioni per ciascuno dei sistemi $\mathbf{A}\mathbf{x}_j = \mathbf{e}_j$. La EG sulle matrici aumentate $[\mathbf{A}\mathbf{e}_j]$ produce le forme ridotte $[\mathbf{U}\mathbf{c}_j]$ in cui l'ultima colonna non è mai dominante, dato che l'ultima riga di $\mathbf{U}$ non è nulla (poiché rango è m). Pertanto i sistemi considerati hanno soluzione e la matrice $\mathbf{R}$ esiste (ed è composta dalle colonne soluzione dei vari sistemi).

□

Osservazione 57. L'equivalenza di 1 e 3 nel teorema 1.6.8 e la proposizione 1.6.3 hanno la seguente conseguenza immediata.

Corollario 1.6.9. *Una matrice quadrata di ordine* m *è invertibile se e solo se ha rango* m

Dimostrazione. Se ha inversa destra, questa coincide con la sinistra e con l'inversa tout court; quindi è invertibile e pertanto ha rango m (e viceversa). □

Osservazione 58. Veniamo ora alla caratterizzazione delle matrici che ammettono inversa sinistra.

Teorema 1.6.10. *Per una matrice* $\mathbf{A}$ $m \times n$ *le seguenti proprietà sono equivalenti:*

1. $\mathbf{A}$ *ammette inversa sinistra;*
2. *il sistema* $\mathbf{A}\mathbf{x} = \mathbf{0}$ *ammette solo la soluzione nulla*
3. $\mathbf{A}$ *ha rango* n
4. *la matrice* $\mathbf{A}^H\mathbf{A}$ *è invertibile*

Dimostrazione. Rispettivamente:

- $1 \implies 2$ segue dalla proposizione 1.6.4, punto 2;

- 2 $\iff$ 3: $\mathbf{A}$ ha rango $< n$ se e solo se la matrice aumentata $[\mathbf{A}\,\mathbf{0}]$ ha forma ridotta $[\mathbf{U}\,\mathbf{0}]$ con meno di n righe non nulle. Ciò equivale al fatto che almeno una colonna di $\mathbf{U}$ non è dominante, cioè il sistema $\mathbf{Ax} = \mathbf{0}$ ha infinite soluzioni, oltre a quella nulla (mentre se rango $= n$ il sistema è determinato e l'unica soluzione possibile è $\mathbf{0}$, ossia la soluzione nulla);

- 3 $\implies$ 4: essendo $\mathbf{A}^H\mathbf{A}$ matrice quadrata di ordine n, per il corollario 1.6.9 è sufficiente provare che ha rango n. Per l'implicazione 2 $\implies$ 3 applicata alla matrice $\mathbf{A}^H\mathbf{A}$ basta provare che il sistema $\mathbf{A}^H\mathbf{Ax} = \mathbf{0}$ ammette solo la soluzione nulla. Sia allora $\mathbf{v}$ una generica soluzione tale che $\mathbf{A}^H\mathbf{Av} = \mathbf{0}$; pre-moltiplicando entrambi i membri per $\mathbf{v}^H$ (deve essere vettore riga per essere conformabile e $\mathbf{v}^H$ lo è) e si ottiene $\mathbf{v}^H(\mathbf{A}^H\mathbf{Av}) = 0$ (occhio che quest'ultimo è scalare). Posto $\mathbf{Av} = \mathbf{y} = (y_1, \ldots, y_m)^T$ si ha

$$\mathbf{y}^H\mathbf{y} = (\mathbf{Av})^H\mathbf{Av} = \mathbf{v}^H\mathbf{A}^H\mathbf{Av} = 0$$

  Il fatto che $\mathbf{y}^H\mathbf{y} = \overline{y}_1 y_1 + \ldots \overline{y}_m y_m = 0$ implica che $y_i = 0$ per ogni i (perché $\overline{y}_i y_i$ è un numero reale ≥ 0 per ogni i) e quindi $\mathbf{y} = \mathbf{0}$. Ma allora $\mathbf{Av} = \mathbf{y} = \mathbf{0}$ e dalla 2 segue che $\mathbf{v} = \mathbf{0}$ (unica soluzione possibile) come desiderato (volevamo dimostrare che $\mathbf{A}^H\mathbf{Ax} = \mathbf{0}$ ammette solo $\mathbf{x} = \mathbf{v} = \mathbf{0}$ come soluzione);

- 4 $\implies$ 1 basta rilevare che una inversa sinistra di $\mathbf{A}$ è $(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$

□

Osservazione 59. Manca l'equivalente del punto 4 per la caratterizzazione delle matrici con inversa destra, che andiamo ad aggiungere ora, in quanto in grado di dimostrarlo.

Corollario 1.6.11. *Una matrice $m \times n$ $\mathbf{A}$ ha inversa destra se e solo se la matrice $\mathbf{AA}^H$ è invertibile*

Dimostrazione. Se $\mathbf{A}$ ha inversa destra allora $\mathbf{A}^H$ ha inversa sinistra: infatti se $\mathbf{AR} = \mathbf{I}$ allora $(\mathbf{AR})^H = \mathbf{I}^H$ da cui $\mathbf{R}^H\mathbf{A}^H = \mathbf{I}$ e quindi una inversa sinistra di $\mathbf{A}^H$ è $\mathbf{R}^H$. Per il teorema 1.6.10, punto 4, allora $(\mathbf{A}^H)^H\mathbf{A}^H = \mathbf{AA}^H$ è invertibile.
Viceversa se $\mathbf{AA}^H$ è invertibile, una inversa destra di $\mathbf{A}$ è $\mathbf{A}^H(\mathbf{AA}^H)^{-1}$ dato che $\mathbf{AA}^H(\mathbf{AA}^H)^{-1} = \mathbf{I}$. □

Osservazione 60. Per ciò che concerne le matrici quadrate, le caratterizzazioni di matrici con inversa destra e sinistra, assieme alla proposizione 1.6.3 porgono come conseguenza il seguente teorema. Il punto 5 impiega la nozione di determinante (introdotta in seguito) ed è pertanto posto qui a titolo informativo, per completezza.

Teorema 1.6.12. *Sia $\mathbf{A}$ una matrice quadrata di ordine n. Le seguenti proprietà sono equivalenti:*

1. *$\mathbf{A}$ è invertibile;*

2. *il sistema $\mathbf{Ax} = \mathbf{b}$ ammette almeno una soluzione per ogni scelta del vettore $\mathbf{b}$*

3. *il sistema* $\mathbf{Ax} = \mathbf{0}$ *ha come unica soluzione* $\mathbf{x} = \mathbf{0}$

4. $\mathbf{A}$ *ha rango* n

5. *il determinante di* $\mathbf{A}$ *è diverso da 0*

Dimostrazione. La 1 $\implies$ 2 deriva dal fatto che se $\mathbf{A}$ è invertibile, ha sia inversa destra che sinistra e queste coincidono; allora dato che ha inversa destra si giunge alla proprietà 2 dalla caratterizzazione delle matrici con inversa destra (punto 2).
La 1 $\implies$ 3 deriva dal fatto che $\mathbf{A}$ ha inversa sinistra, alla luce della caratterizzazione delle matrici con inversa sinistra (punto 2).
La 1 $\implies$ 4 deriva dalle caratterizzazioni di matrici con inversa destra e sinistra (punti 3).
La 5 è lasciata a titolo informativo, per ora. □

Osservazione 61. Il modo più semplice per vedere se una matrice quadrata è invertibile risulta quello di verificare il punto 4 del precedente teorema: si controlla che una forma ridotta della matrice non abbia righe nulle.

Proposizione 1.6.13 (Proprietà matrici quadrate invertibili di ordine n)**.** *Siano* $\mathbf{A}$ *e* $\mathbf{B}$ *matrici quadrate invertibili di ordine* n*. Allora:*

1. $\mathbf{AB}$ *è invertibile e* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

2. $\mathbf{A}^{-1}$ *è invertibile e* $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$

3. $\mathbf{A}^T$ *e* $\mathbf{A}^H$ *sono invertibili e* $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$, $(\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H$

Dimostrazione. Per il primo punto è facile verificare che

$$\mathbf{ABB}^{-1}\mathbf{A}^{-1} = \mathbf{AIA}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}$$
$$\mathbf{B}^{-1}\mathbf{A}^{-1}\mathbf{AB} = \mathbf{B}^{-1}\mathbf{IB} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$$

Quindi si può concludere che $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
Per il secondo punto, dato che $\mathbf{A}$ è invertibile si ha che $\mathbf{AA}^{-1} = \mathbf{I}$ e $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$; ma allora vedendola dall'altra angolazione anche $\mathbf{A}^{-1}$ è invertibile e la sua inversa $(\mathbf{A}^{-1})^{-1}$ è $\mathbf{A}$.
Per il terzo punto (mostriamo la dimostrazione per la trasposta) $\mathbf{A}^T$ è quadrata di ordine n; dalla $\mathbf{AA}^{-1} = \mathbf{I}$ traspongo entrambi i membri $(\mathbf{AA}^{-1})^{\mathbf{T}} = \mathbf{I}^T$ e per le proprietà della trasposizione si ha $(\mathbf{A}^{-1})^{\mathbf{T}}\mathbf{A}^T = \mathbf{I}$. Pertanto l'inversa della trasposta $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$. Idem per $\mathbf{A}^H$. □

1.6.2 Inversa di matrice 2×2

Osservazione 62. L'algoritmo di inversione presentato in seguito è il più generale, ma vediamo prima un esempio notevole.

Proposizione 1.6.14 (Inversa di matrice 2×2)**.** *La matrice* 2×2

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

ha inversa se e solo se $\Delta = ad - bc \neq 0$ *(si vedrà che* Δ*è il determinante della matrice* $\mathbf{A}$*), e in tal caso l'inversa è la matrice*

$$\mathbf{A}^{-1} = \Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Dimostrazione. Se $\Delta \neq 0$ mostriamo che $\mathbf{A}$ è invertibile; ad esempio, sfruttando la pre-moltiplicazione si vede come:

$$\Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} ad-bc & db-bd \\ -ca+ac & -cb+ad \end{bmatrix} = \frac{1}{ad-bc} \begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

e dunque $\mathbf{A}$ è invertibile con inversa $\Delta^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
Viceversa nell'ipotesi $\mathbf{A}$ sia invertibile dimostriamo che $\Delta \neq 0$; se

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \mathbf{I}_2$$

allora deve essere che $ax + bz = 1$, $cx + dz = 0$ e $cy + dw = 1$. Se:

- $c \neq 0$, sottraendo dalla prima uguaglianza moltiplicata per c la seconda moltiplicata per a si ha che è $(bc - ad)z = c$;

- $d \neq 0$, sottraendo dalla prima uguaglianza moltiplicata per d la seconda moltiplicata per b si ricava $(ad - bc)x = d$

Poiché non può essere che contemporaneamente $c = d = 0$ (visto che $cy+dw = 1$) in ogni caso risulta $\Delta \neq 0$. □

Esempio 1.6.1. Sia $\mathbf{A} = \begin{bmatrix} 3 & i \\ 1 & 1 \end{bmatrix}$. Si provi che $\mathbf{A}$ è non singolare (ossia invertibile) e si calcoli $\mathbf{A}^{-1}$.
Il determinante è $3 \cdot 1 - i = 3 - i \neq 0$, quindi la matrice è invertibile. L'inversa è

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{3-i} \begin{bmatrix} 1 & -i \\ -1 & 3 \end{bmatrix} = \frac{3+i}{(3-i)(3+i)} \begin{bmatrix} 1 & -i \\ -1 & 3 \end{bmatrix} = \frac{3+i}{10} \begin{bmatrix} 1 & -i \\ -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} \frac{3+i}{10} & \frac{1-3i}{10} \\ -\frac{3+i}{10} & \frac{9+3i}{10} \end{bmatrix} \end{aligned}$$

1.6.3 Inverse di matrici triangolari

Osservazione 63. È utile sapere cosa accade per le inverse di matrici triangolari e per matrici decomposte a blocchi.

Esempio 1.6.2. Una matrice triangolare superiore (risp. inferiore) $\mathbf{T} = (t_{ij})$ è invertibile se e solo se tutti i coefficienti diagonali t_{ii} sono diversi da zero. Se ciò accade, la matrice inversa $\mathbf{T}^{-1}$ è ancora triangolare superiore (risp. inferiore) e ha come coefficienti diagonali gli inversi dei corrispondenti coefficienti diagonali di $\mathbf{T}$. La dimostrazione si esegue facilmente per induzione sull'ordine della matrice.

Esempio 1.6.3. Lo stesso avviene per matrici triangolari a blocchi, con blocchi diagonali quadrati, sostituendo le condizioni sui coefficienti diagonali con analoghe condizioni sui blocchi diagonali. In particolare, data la matrice triangolare a blocchi

$$\mathbf{T} = \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbb{O} & \mathbf{V} \end{bmatrix}$$

con $\mathbf{X}$ e $\mathbf{V}$ blocchi quadrati invertibili risulta

$$\mathbf{T}^{-1} = \begin{bmatrix} \mathbf{X}^{-1} & -\mathbf{X}^{-1}\mathbf{Y}\mathbf{V}^{-1} \\ \mathbb{O} & \mathbf{V}^{-1} \end{bmatrix}$$

in quanto

$$\begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbb{O} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{X}^{-1} & -\mathbf{X}^{-1}\mathbf{Y}\mathbf{V}^{-1} \\ \mathbb{O} & \mathbf{V}^{-1} \end{bmatrix} = \begin{bmatrix} \mathbf{X}\mathbf{X}^{-1} & -\mathbf{X}\mathbf{X}^{-1}\mathbf{Y}\mathbf{V}^{-1} + \mathbf{Y}\mathbf{V}^{-1} \\ \mathbb{O} & \mathbf{V}\mathbf{V}^{-1} \end{bmatrix}$$

dove l'ultima matrice a blocchi è evidentemente la matrice identità.

1.6.4 Pseudo-inversa di Moore-Penrose

Osservazione 64. Una matrice ha inversa (bilatera) solo in casi particolari (quadrate). Esiste una nozione che generalizza quella di matrice inversa ed è applicabile a qualunque matrice.

Definizione 1.6.6 (Pseudo-inversa)**.** Data una qualunque matrice $\mathbf{A}$ di tipo $m \times n$ si chiama pseudo-inversa di $\mathbf{A}$ una matrice $\mathbf{A}^+$ che soddisfa quattro condizioni:

1. $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$
2. $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$
3. $\mathbf{A}\mathbf{A}^+ = (\mathbf{A}\mathbf{A}^+)^H$
4. $\mathbf{A}^+\mathbf{A} = (\mathbf{A}^+\mathbf{A})^H$

Osservazione 65. La pseudo-inversa ha necessariamente dimensioni $n \times m$. Si pone il problema di esistenza ed unicità di tale matrice; partiamo da questo secondo problema.

Proposizione 1.6.15 (Unicità della matrice pseudo-inversa)**.** *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *e siano* $\mathbf{B}$ *e* $\mathbf{C}$ *due matrici che soddisfano le quattro condizioni che deve soddisfare una matrice pseudo-inversa. Allora* $\mathbf{B} = \mathbf{C}$.

Dimostrazione. Facciamo uso delle proprietà delle matrici H-trasposte (indicate con (h) e delle quattro condizioni ((1), (2), (3), (4))

$$\begin{aligned}
\mathbf{B} &\overset{(2)}{=} \mathbf{BAB} \overset{(4)}{=} (\mathbf{BA})^H\mathbf{B} \overset{(h)}{=} \mathbf{A}^H\mathbf{B}^H\mathbf{B} \overset{(1)}{=} (\mathbf{ACA})^H\mathbf{B}^H\mathbf{B} \\
&\overset{(h)}{=} \mathbf{A}^H(\mathbf{AC})^H\mathbf{B}^H\mathbf{B} \overset{(h)}{=} \mathbf{A}^H\mathbf{C}^H\mathbf{A}^H\mathbf{B}^H\mathbf{B} \overset{(h)}{=} (\mathbf{CA})^H(\mathbf{BA})^H\mathbf{B} \\
&\overset{(4)}{=} \mathbf{CABAB} \overset{(2)}{=} \mathbf{CAB} \overset{(2)}{=} \mathbf{CACAB} \overset{(3)}{=} \mathbf{C}(\mathbf{AC})^H(\mathbf{AB})^H \\
&\overset{(h)}{=} \mathbf{CC}^H\mathbf{A}^H\mathbf{B}^H\mathbf{A}^H \overset{(h)}{=} \mathbf{CC}^H(\mathbf{ABA})^H \overset{(1)}{=} \mathbf{CC}^H\mathbf{A}^H \overset{(h)}{=} \mathbf{C}(\mathbf{AC})^H \\
&\overset{(3)}{=} \mathbf{CAC} \overset{(2)}{=} \mathbf{C}
\end{aligned}$$

□

Teorema 1.6.16 (Esistenza della pseudo-inversa)**.** *Data una qualunque matrice* $\mathbf{A}$ *la sua pseudo-inversa* $\mathbf{A}^+$ *esiste ed è unica.*

Osservazione 66. Proviamo questo teorema suddividendo la casistica in matrici che ammettono inversa, inversa destra, inversa sinistra e matrici generiche (che possono non ammettere inverse).

Proposizione 1.6.17. *Sia* $\mathbf{A}$ *matrice quadrata invertibile.* $\mathbf{A}^{-1}$ *è la sua pseudo-inversa, ossia* $\mathbf{A}^+ = \mathbf{A}^{-1}$

Dimostrazione. È immediato verificare che $\mathbf{A}^{-1}$ verifica tutte le quattro condizioni della pseudo-inversa. □

Osservazione 67. Pertanto nelle matrici quadrate, l'inversa coincide con la pseudo-inversa.

Proposizione 1.6.18. *Sia* $\mathbf{A}$ *matrice non quadrata che ammette inversa destra. La matrice ha pseudo-inversa:*

$$\mathbf{A}^+ = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}$$

Dimostrazione. Sia $\mathbf{A}$ di tipo $m \times n$ con $m < n$. La caratterizzazione delle matrici con inversa destra (nello specifico il punto 4 presentato in teorema 1.6.11) assicura che esista la matrice $\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}$ che è ovviamente inversa destra di $\mathbf{A}$; vogliamo verificare che essa sia anche pseudo-inversa ossia $\mathbf{A}^+ = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}$. Poiché ogni inversa destra soddisfa banalmente le prime tre delle condizioni richieste delle pseudo-inverse basta controllare che verifichi anche la quarta ossia che $\mathbf{A}^+\mathbf{A} = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}\mathbf{A}$ sia hermitiana. Si ha che:

$$(\mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}\mathbf{A})^H \overset{(1)}{=} \mathbf{A}^H((\mathbf{A}\mathbf{A}^H)^{-1})^H\mathbf{A} \overset{(2)}{=} \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}\mathbf{A}$$

dove in (1) si sono applicate le proprietà dell'h-trasposizione in rapporto al prodotto; in (2) si ha che $\mathbf{A}\mathbf{A}^H$ è quadrata $m \times m$ ed essendo quadrata ed invertibile possiamo "invertire" gli esponenti (punto 3 di proposizione 1.6.13), quindi $((\mathbf{A}\mathbf{A}^H)^{-1})^H = ((\mathbf{A}\mathbf{A}^H)^H)^{-1}$ e per proprietà della H-trasposta si ha $(\mathbf{A}\mathbf{A}^H)^H = (\mathbf{A}^H)^H\mathbf{A}^H = \mathbf{A}\mathbf{A}^H$ da cui $((\mathbf{A}\mathbf{A}^H)^H)^{-1} = (\mathbf{A}\mathbf{A}^H)^{-1}$.
Pertanto anche la quarta proprietà è rispettata e $\mathbf{A}^+ = \mathbf{A}^H(\mathbf{A}\mathbf{A}^H)^{-1}$, ossia anche questo tipo di matrice ha pseudo-inversa.
Si può poi verificare o dedurre dalla 1.6.15 che questa è l'unica tra le infinite inverse destre $\mathbf{R}$ di $\mathbf{A}$ a soddisfare il fatto che $\mathbf{R}\mathbf{A}$ sia hermitiana (quarto requirement). □

Proposizione 1.6.19. *Sia* $\mathbf{A}$ *matrice non quadrata che ammette inversa sinistra. La matrice ha pseudo-inversa:*

$$\mathbf{A}^+ = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$$

Dimostrazione. Sia $\mathbf{A}$ di tipo $m \times n$ con $m > n$. Con ragionamento speculare (ma fatto con riferimento alla caratterizzazione di matrici aventi l'inversa sinistra) qualsiasi inversa sinistra soddisfa le condizioni 1, 2 e 4 della definizione di pseudo-inversa e si ha che $(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ verifica anche la terza. Pertanto $\mathbf{A}^+ = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ □

Osservazione 68. L'esistenza della matrice pseudo-inversa nel caso di una matrice qualsiasi si deduce dai due precedenti casi, una volta che si sappia decomporre la matrice in modo opportuno

Proposizione 1.6.20. *Sia* $\mathbf{A}$ *una matrice* $m \times n$ *e sia* $\mathbf{A} = \mathbf{BC}$ *una sua fattorizzazione tale che* $\mathbf{B}$ *ammetta inversa sinistra e* $\mathbf{C}$ *ammetta inversa destra. Allora* $\mathbf{A}^+$ *esiste e risulta:*

$$\mathbf{A}^+ = \mathbf{C}^+\mathbf{B}^+ = \mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^{-1}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H \tag{1.72}$$

Osservazione 69. Come si vedrà nell'osservazione 95 una generica matrice $\mathbf{A}$ ha sempre fattorizzazioni soddisfacenti l'ipotesi di proposizione 1.6.20 (*dette decomposizioni a rango pieno*), pertanto qualsiasi matrice ha pseudo-inversa.

Dimostrazione. Dai casi di matrice con inversa destra e sinistra visti in precedenza sappiamo che (dato che $\mathbf{B}$ ha inversa sinistra e $\mathbf{C}$ inversa destra):

$$\mathbf{B}^+ = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H, \quad \mathbf{C}^+ = \mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^{-1}$$

Nell'ipotesi che $\mathbf{A} = \mathbf{BC}$ e $\mathbf{A}^+ = \mathbf{C}^+\mathbf{B}^+$ verifichiamo le uguaglianze (punti 1 e 2) richieste dalla definizione di pseudo-inversa :

$$\underbrace{\mathbf{BC}}_{\mathbf{A}} = \underbrace{\mathbf{BC}}_{\mathbf{A}}\underbrace{\mathbf{C}^+\mathbf{B}^+}_{\mathbf{A}^+}\underbrace{\mathbf{BC}}_{\mathbf{A}}, \qquad \underbrace{\mathbf{C}^+\mathbf{B}^+}_{\mathbf{A}^+} = \underbrace{\mathbf{C}^+\mathbf{B}^+}_{\mathbf{A}^+}\underbrace{\mathbf{BC}}_{\mathbf{A}}\underbrace{\mathbf{C}^+\mathbf{B}^+}_{\mathbf{A}^+}$$

Le uguaglianze di sopra risultano verificate essendo $\mathbf{B}^+ = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ inversa sinistra di $\mathbf{B}$ e $\mathbf{C}^+ = \mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^{-1}$ inversa destra di $\mathbf{C}$.
La verifica dei punti 3 e 4 della definizione di pseudo-inversa, ossia che sono hermitiane le matrici

$$\mathbf{A}\mathbf{A}^+ = \mathbf{B}\mathbf{C}\mathbf{C}^+\mathbf{B}^+ = \mathbf{B}\mathbf{B}^+, \qquad \mathbf{A}^+\mathbf{A} = \mathbf{C}^+\mathbf{B}^+\mathbf{B}\mathbf{C} = \mathbf{C}^+\mathbf{C}$$

segue dal fatto che $\mathbf{B}\mathbf{B}^+$ e $\mathbf{C}\mathbf{C}^+$ sono hermitiane, come visto nelle dimostrazioni delle proposizioni 1.6.18 e 1.6.19 □

Proposizione 1.6.21 (Proprietà della matrice pseudo-inversa)**.** *Si ha:*

(pi_1) $(\mathbf{A}^+)^+ = \mathbf{A}$

(pi_2) $(\mathbf{A}^+)^H = (\mathbf{A}^H)^+$

(pi_3) $\mathbb{O}^+_{mn} = \mathbb{O}_{nm}$

(pi_4) $(\alpha\mathbf{A})^+ = \alpha^{-1}\mathbf{A}^+$ *per ogni scalare* $\alpha \neq 0$

(pi_5) $\mathbf{A}^H = \mathbf{A}^+\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}\mathbf{A}^+$

(pi_6) *se* $\mathbf{A} = \mathbf{BC}$ *con* $\mathbf{B}$ *soddisfacente l'uguaglianza* $\mathbf{B}^H\mathbf{B} = \mathbf{I}$*, allora* $\mathbf{A}^+ = \mathbf{C}^+\mathbf{B}^H$

(pi_7) *se* $\mathbf{A} = \mathbf{BC}$ *con* $\mathbf{C}$ *soddisfacente l'uguaglianza* $\mathbf{C}\mathbf{C}^H = \mathbf{I}$*, allora* $\mathbf{A}^+ = \mathbf{C}^H\mathbf{B}^+$

Dimostrazione. Verifica lasciata per esercizio □

TODO: fixme

Esempio 1.6.4. Si consideri la matrice 2×3

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}$$

È immediato verificare che può essere fattorizzata come segue

$$\mathbf{A} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 3 \end{bmatrix}$$

dove $\mathbf{B} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ ha inversa sinistra e $\mathbf{C} = \begin{bmatrix} 2 & 0 & 1 & 3 \end{bmatrix}$ ha inversa destra. Un facile calcolo mostra che

$$\mathbf{B}^+ = (\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H = \left(\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 & 1 \end{bmatrix} = \begin{bmatrix} 2/5 & 1/5 \end{bmatrix}$$

$$\mathbf{C}^+ = \mathbf{C}^H(\mathbf{C}\mathbf{C}^H)^{-1} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix} \left(\begin{bmatrix} 2 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix} \right)^{-1} = \begin{bmatrix} 2/14 \\ 0 \\ 1/14 \\ 3/14 \end{bmatrix}$$

Pertanto la pseudo-inversa di $\mathbf{A}$ risulta essere:

$$\mathbf{A}^+ = \begin{bmatrix} 2/14 \\ 0 \\ 1/14 \\ 3/14 \end{bmatrix} \begin{bmatrix} 2/5 & 1/5 \end{bmatrix} = \begin{bmatrix} 4/70 & 2/70 \\ 0 & 0 \\ 2/70 & 1/70 \\ 6/70 & 3/70 \end{bmatrix} = 70^{-1}\mathbf{A}^T$$

La pseudo inversa si calcola in R mediante `MASS::ginv`:

```
A <- rmatrix(c(4, 0, 2, 6,
               2, 0, 1, 3), nrow = 2)
(pseudoinv <- MASS::ginv(A))

##             [,1]       [,2]
## [1,] 0.05714286 0.02857143
## [2,] 0.00000000 0.00000000
## [3,] 0.02857143 0.01428571
## [4,] 0.08571429 0.04285714

## verifica
pseudoinv * 70

##      [,1] [,2]
## [1,]    4    2
## [2,]    0    0
## [3,]    2    1
## [4,]    6    3
```

In maxima occorre utilizzare la funzione `moore_penrose_pseudoinverse` del pacchetto `linearalgebra`:

```
##
## load(linearalgebra)
## moore_penrose_pseudoinverse(matrix([4,0,2,6],[2,0,1,3]))
##                                    [ 2   1  ]
##                                    [ --  -- ]
##                                    [ 35  35 ]
##                                    [        ]
##                                    [ 0   0  ]
##                                    [        ]
##                                    [ 1   1  ]
##                                    [ --  -- ]
##                                    [ 35  70 ]
##                                    [        ]
##                                    [ 3   3  ]
##                                    [ --  -- ]
##                                    [ 35  70 ]
```

1.7 Calcolo delle matrici inverse

Osservazione 70. La ricerca di una inversa destra (o bilatera, nel caso quadrato) di una matrice viene ricondotta alla soluzione di una serie di sistemi lineari: tramite l'algoritmo di Gauss-Jordan (che perfeziona la EG) queste serie di sistemi viene trattata in blocco e viene restituita l'inversa.

1.7.1 Inversa di matrice quadrata

Osservazione 71. Vediamo il funzionamento dell'algoritmo con un esempio.

Esempio 1.7.1. Data la matrice 3×3

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 1 \\ -1 & 0 & -2 \end{bmatrix}$$

si consideri la matrice a blocchi 3×6 ottenuta affiancandogli $\mathbf{I}_3$

$$[\mathbf{A}|\mathbf{I}_3] = \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ -1 & 0 & -2 & 0 & 0 & 1 \end{array}\right]$$

Lo scopo dell'algoritmo Gauss-Jordan è di modificare questa matrice sino ad ottenerne una del tipo $[\mathbf{I}_3|\mathbf{B}]$, con $\mathbf{B}$ che risulterà essere la matrice inversa di $\mathbf{A}$. Si esegue l'usuale EG sulla matrice $[\mathbf{A}|\mathbf{I}_3]$ pervenendo alla matrice

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 2 & 1 & 2 \end{array}\right]$$

Il blocco di sinistra è una forma ridotta di $\mathbf{A}$: ora inizia la parte nuova dell'algoritmo. Si procede con quella che si chiama eliminazione all'indietro che, tramite le solite operazioni elementari pone degli zeri al di sopra degli elementi diagonali, partendo dall'ultima riga e risalendo fino alla prima.
Usando come pivot il coefficiente di posto $(3,3)$ si sostituisce alla seconda riga la sua somma con la terza moltiplicata per $-1/2$ e alla prima riga la sua somma con la terza moltiplicata per -2; si ottiene la matrice

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 0 & -3 & -2 & -4 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 2 & 1 & 2 \end{array}\right]$$

Ora usando il coefficiente di posto $(2,2)$ si sostituisce alla prima riga la sua somma con la seconda moltiplicata per 1, ottenendo la matrice

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & -2 & -5 \\ 0 & 1 & 0 & -1 & 0 & -1 \\ 0 & 0 & 1 & 2 & 1 & 2 \end{array}\right]$$

La matrice ha assunto quindi la forma desiderata $[\mathbf{I}_3|\mathbf{B}]$ con

$$\mathbf{B} = \begin{bmatrix} -4 & -2 & -5 \\ -1 & 0 & -1 \\ 2 & 1 & 2 \end{bmatrix}$$

e $\mathbf{B}$ coincide con l'inversa $\mathbf{A}^{-1}$ (facile verificare che $\mathbf{AB} = \mathbf{I}_3$)

In R la funzione `solve` può essere utilizzata su matrici quadrate anche per il calcolo dell'inversa, se si fornisce una matrice identità come parametro `b` (o anche senza).

```
A <- rmatrix(c( 1, -1,  2,
                0,  2,  1,
               -1,  0, -2), nrow = 3)
b <- diag(3)
solve(A, b) # solve(A) è equivalente

##      [,1] [,2] [,3]
## [1,]   -4   -2   -5
## [2,]   -1    0   -1
## [3,]    2    1    2
```

In maxima si usa `invert` o `^^(-1)` per invertire una matrice:

```
##
## A:matrix([1,-1,2],[0,2,1],[-1,0,-2])
## invert(A)
##                               [ - 4  - 2  - 5 ]
##                               [               ]
##                               [ - 1   0   - 1 ]
##                               [               ]
##                               [  2    1    2  ]
## A . A^^(-1)
##                                  [ 1  0  0 ]
##                                  [         ]
##                                  [ 0  1  0 ]
##                                  [         ]
##                                  [ 0  0  1 ]
```

Osservazione 72. Vediamo perché l'algoritmo funziona.

Proposizione 1.7.1 (Inversa di matrice quadrata)**.** *Sia* $\mathbf{A}$ *una matrice quadrata* $n \times n$*. Se* $\mathbf{A}$ *è invertibile, la sua inversa* $\mathbf{A}^{-1}$ *si ottiene a partire dalla matrice a blocchi* $[\mathbf{A}|\mathbf{I}_n]$ *operando su di essa con l'algoritmo di Gauss-Jordan fino a trasformarla nella forma* $[\mathbf{I}_n|\mathbf{B}]$*. Allora la matrice inversa* $\mathbf{A}^{-1}$ *coincide con la matrice* $\mathbf{B}$*.*

Dimostrazione. Mettiamo in evidenza le colonne della matrice $\mathbf{A}^{-1}$

$$\mathbf{A}^{-1} = [\mathbf{v}_1\, \mathbf{v}_2\, \dots\, \mathbf{v}_n]$$

L'eguaglianza $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n$ si può scrivere come una serie di n uguaglianze

$$\mathbf{A}\mathbf{v}_i = \mathbf{e}_i \qquad (i = 1, 2, \dots, n)$$

dato che $\mathbf{A}\mathbf{v}_i$ è la i-esima colonna della matrice prodotto $\mathbf{A}\mathbf{A}^{-1}$ ed $\mathbf{e}_i$ è la i-esima colonna della matrice identità. Perciò trovare $\mathbf{A}^{-1}$ significa risolvere gli n sistemi lineari, ciascuno in n equazioni ed n incognite $\mathbf{A}\mathbf{x} = \mathbf{e}_i$ $(i = 1, 2, \dots, n)$. Questi sistemi sono tra loro "parenti" poiché condividono la matrice dei coefficienti, mentre i loro termini noti sono gli n vettori coordinati $\mathbf{e}_1, \dots, \mathbf{e}_n$. Le operazioni elementari che si eseguono sulle matrici aumentate degli n sistemi

$$[\mathbf{A}|\mathbf{e}_1], [\mathbf{A}|\mathbf{e}_2], \dots, [\mathbf{A}|\mathbf{e}_n]$$

sono sempre le stesse, per cui è conveniente considerare una unica matrice, detta *matrice pluri-aumentata* di $\mathbf{A}$:

$$[\mathbf{A}|\mathbf{e}_1\, \mathbf{e}_2\, \dots\, \mathbf{e}_n]$$

Se si eseguono in un sol colpo su tale matrice le operazioni elementari che si eseguirebbero per gli n sistemi, quindi se si opera la EG cui si fanno seguire la eliminazione all'indietro, si ottiene la matrice

$$[\mathbf{I}_n|\mathbf{B}] = [\mathbf{I}_n|\mathbf{b}_1\, \dots\, \mathbf{b}_n]$$

dove i vettori $\mathbf{b}_i$ sono le colonne della matrice $\mathbf{B}$. Ciò significa che gli n sistemi $\mathbf{A}\mathbf{x} = \mathbf{e}_i$ sono stati trasformati negli n sistemi equivalenti

$$\mathbf{I}_n\mathbf{x} = \mathbf{b}_i$$

dove evidentemente le soluzioni sono proprio i vettori $\mathbf{b}_i$, perciò per ogni $i = 1, 2, \ldots, n$ risulta $\mathbf{A}\mathbf{b}_i = \mathbf{e}_i$. Ne consegue che $\mathbf{v}_i = \mathbf{b}_i$ per ogni $i = 1, 2, \ldots, n$ pertanto $\mathbf{A}^{-1} = \mathbf{B}$ □

Osservazione 73 (Algoritmo applicato a matrice non invertibile). Se si applica l'algoritmo di Gauss-Jordan ad una matrice non invertibile $n \times n$, la EG modificherà la matrice pluri-aumentata $[\mathbf{A}|\mathbf{I}_n]$ in una matrice $[\mathbf{U}|\mathbf{C}]$ dove $\mathbf{U}$ non sarà più triangolare superiore con elementi diagonali non nulli, perché altrimenti completando l'algoritmo si otterrebbe l'inversa che non esiste.
La matrice $\mathbf{U}$ sarà solo a scala per righe con almeno una riga nulla; perciò la successiva parte dell'algoritmo, cioè l'eliminazione all'indietro, non potrà in alcun modo trasformare $\mathbf{U}$ in $\mathbf{I}_n$.

Osservazione 74 (Efficienza algoritmo). L'algoritmo fornisce un modo semplice ed efficiente per vedere se una matrice quadrata ha inversa e, nel caso, di calcolarla.

Esempio 1.7.2. Sia

$$\mathbf{A}_\alpha = \begin{bmatrix} 1 & \alpha & \alpha+3 \\ 0 & \alpha & -2\alpha \\ 1 & \alpha & 2\alpha+4 \end{bmatrix}$$

con $\alpha \in \mathbb{C}$. Si usi il metodo di Gauss-Jordan per calcolare gli α per cui $\mathbf{A}_\alpha$ è non singolare e, per tali valori si trovi $\mathbf{A}_\alpha^{-1}$.
Effettuando la riduzione

$$\left[\begin{array}{ccc|ccc} 1 & \alpha & \alpha+3 & 1 & 0 & 0 \\ 0 & \alpha & -2\alpha & 0 & 1 & 0 \\ 1 & \alpha & 2\alpha+4 & 0 & 0 & 1 \end{array}\right] r_3 = r_3 - r_1$$

$$\left[\begin{array}{ccc|ccc} 1 & \alpha & \alpha+3 & 1 & 0 & 0 \\ 0 & \alpha & -2\alpha & 0 & 1 & 0 \\ 0 & 0 & \alpha+1 & -1 & 0 & 1 \end{array}\right]$$

La matrice è invertibile solo se ha rango 3 ossia 3 colonne non tutte nulle ossia deve essere

$$\begin{cases} \alpha \neq 0 \\ -2\alpha \neq 0 \\ \alpha + 1 \neq 0 \end{cases} \quad \begin{cases} \alpha \neq 0 \\ \alpha \neq -1 \end{cases}$$

Ora sebbene abbia provato, non son riuscito a riportare la parte sinistra della matrice ad $\mathbf{I}$ (per ottenere l'inversa dalla parte destra). Tuttavia possiamo ugualmente sfruttare l'idea alla base del metodo per cui procediamo risolvendo i sistemi lineari derivanti dai vari $[\mathbf{A}|\mathbf{e}_i]$ ottenendo le colonne dell'inversa, per andare a comporre l'inversa in un secondo tempo accostandole.
Concentrandoci sulla prima colonna dell'inversa il sistema è:

$$\begin{cases} x_3(\alpha+1) = -1 \\ x_2\alpha + x_3(-2\alpha) = 0 \\ x_1 + \alpha x_2 + (\alpha+3)x_3 = 1 \end{cases} \quad \begin{cases} x_3 = -1/(1+\alpha) \\ x_2\alpha + \frac{2\alpha}{1+\alpha} = 0 \\ x_1 + \alpha x_2 - \frac{3+\alpha}{1+\alpha} = 1 \end{cases} \quad \begin{cases} x_3 = -1/(1+\alpha) \\ x_2 = -\frac{2}{1+\alpha} \\ x_1 - \frac{2\alpha}{1+\alpha} - \frac{3+\alpha}{1+\alpha} = 1 \end{cases} \quad \begin{cases} x_1 = 4 \\ x_2 = -\frac{2}{1+\alpha} \\ x_3 = -1/(1+\alpha) \end{cases}$$

Per la seconda colonna

$$\begin{cases} x_3(\alpha+1)=0 \\ x_2\alpha + x3(-2\alpha)=1 \\ x_1+\alpha x_2+(\alpha+3)x_3=0 \end{cases} \quad \begin{cases} x_3=0 \\ x_2=\frac{1}{\alpha} \\ x_1+1=0 \end{cases} \quad \begin{cases} x_1=-1 \\ x_2=\frac{1}{\alpha} \\ x_3=0 \end{cases}$$

Per la terza colonna

$$\begin{cases} x_3(\alpha+1)=1 \\ x_2(\alpha)+x_3(-2\alpha)=0 \\ x_1+\alpha x_2+(\alpha+3)x_3=0 \end{cases} \quad \begin{cases} x_3=\frac{1}{1+\alpha} \\ x_2\cdot\alpha+\frac{-2\alpha}{1+\alpha}=0 \\ x_1+\alpha x_2+\frac{\alpha+3}{1+\alpha}=0 \end{cases} \quad \begin{cases} x_3=\frac{1}{1+\alpha} \\ x_2=\frac{2}{1+\alpha} \\ x_1+\frac{2\alpha}{1+\alpha}\alpha+\frac{\alpha+3}{1+\alpha}=0 \end{cases} \quad \begin{cases} x_1=-3 \\ x_2=\frac{2}{1+\alpha} \\ x_3=\frac{1}{1+\alpha} \end{cases}$$

Pertanto l'inversa sarà

$$\begin{bmatrix} 4 & -1 & -3 \\ -\frac{2}{1+\alpha} & \frac{1}{\alpha} & \frac{2}{1+\alpha} \\ -\frac{1}{1+\alpha} & 0 & \frac{1}{1+\alpha} \end{bmatrix}$$

1.7.2 Inversa destra di matrice orizzontale

Osservazione 75. Generalizziamo il metodo per calcolare matrici inverse destre di una matrice $m \times n$ di rango $m < n$, qualora esse si presentino in una forma opportuna.

Proposizione 1.7.2 (Inversa destra di matrice orizzontale)**.** *Sia* $\mathbf{A} = [\mathbf{B}|\mathbf{C}]$ *una matrice* $m \times n$*, con* $m < n$*, decomposta in due blocchi: il blocco* $\mathbf{B}$ *quadrato* $m \times m$ *invertibile, e il blocco* $\mathbf{C}$ $m \times (n-m)$*. Allora le matrici inverse destre della matrice* $\mathbf{A}$ *sono tutte e sole le matrici* $\mathbf{R}(\mathbf{Y})$ *che si ottengono nel modo seguente: si sceglie una arbitraria matrice* $\mathbf{Y}$ *del tipo* $(n-m) \times m$ *e si pone:*

$$\mathbf{R}(\mathbf{Y}) = \begin{bmatrix} \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} \\ \mathbf{Y} \end{bmatrix}$$

Dimostrazione. Il prodotto a blocchi di $\mathbf{A}$ con $\mathbf{R}(\mathbf{Y})$ porge

$$\begin{aligned} \mathbf{A}\cdot\mathbf{R}(\mathbf{Y}) &= [\mathbf{B}|\mathbf{C}] \begin{bmatrix} \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} \\ \mathbf{Y} \end{bmatrix} \\ &= \mathbf{B}\mathbf{B}^{-1} - \mathbf{B}\mathbf{B}^{-1}\mathbf{C}\mathbf{Y} + \mathbf{C}\mathbf{Y} \\ &= \mathbf{I}_m - \mathbf{C}\mathbf{Y} + \mathbf{C}\mathbf{Y} \\ &= \mathbf{I}_m \end{aligned}$$

Pertanto tutte le matrici $\mathbf{R}(\mathbf{Y})$ definite in questo modo sono inverse destre di $\mathbf{A}$. Viceversa, se $\mathbf{R}$ è inversa destra di $\mathbf{A}$, decomponendola a blocchi nella forma

$$\mathbf{R} = \begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix}$$

con $\mathbf{X}$ e $\mathbf{Y}$ blocchi di dimensioni $m \times m$ e $(n-m)\times m$ rispettivamente. Andiamo a determinare $\mathbf{X}$ e $\mathbf{Y}$; si ha:

$$\begin{aligned} \mathbf{A}\mathbf{R} = \mathbf{I}_m &\iff \mathbf{B}\mathbf{X} + \mathbf{C}\mathbf{Y} = \mathbf{I}_m \overset{(1)}{\iff} \mathbf{B}^{-1}\mathbf{B}\mathbf{X} + \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} = \mathbf{B}^{-1} \\ &\iff \mathbf{X} = \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} \end{aligned}$$

dove in (1) si è pre-moltiplicato per $\mathbf{B}^{-1}$. Pertanto $\mathbf{R} = \mathbf{R}(\mathbf{Y})$ □

Esempio 1.7.3. Per trovare tutte le inverse destre della matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 3 & 5 & 2 \end{bmatrix}$$

decomponiamo $\mathbf{A}$ a blocchi $\mathbf{A} = [\mathbf{B}|\mathbf{C}]$ dove

$$\mathbf{B} = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Dato che $\mathbf{B}$ è invertibile (determinante $\neq 0$) con

$$\mathbf{B}^{-1} = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}$$

si può applicare la proposizione 1.7.2. Sia quindi $\mathbf{Y} = [a\, b]$ una qualunque matrice 1×2; le inverse destre della matrice $\mathbf{A}$ sono allora le matrici

$$\begin{bmatrix} \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} \\ \mathbf{Y} \end{bmatrix}$$

dove

$$\begin{aligned} \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{C}\mathbf{Y} &= \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} - \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} \\ &= \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} - \begin{bmatrix} -a & -b \\ a & b \end{bmatrix} \\ &= \begin{bmatrix} 5+a & -3+b \\ -3-a & 2-b \end{bmatrix} \end{aligned}$$

Pertanto (tutte e sole) le matrici inverse destre della matrice $\mathbf{A}$ sono della forma

$$\begin{bmatrix} 5+a & -3+b \\ -3-a & 2-b \\ a & b \end{bmatrix}$$

Osservazione 76. Per il calcolo dell'inversa destra di una generica matrice $m \times n$ di rango $m < n$ (ovvero non necessariamente con $\mathbf{B}$ invertibile) occorre dapprima trasformarla in una matrice che soddisfi le ipotesi di cui al teorema 1.7.2, procedere come illustrato, e una volta trovare l'inversa destra risalire a quella della matrice di partenza. Questo richiede alcuni risultati sulle matrici di permutazione e sarà visto in seguito.

Osservazione 77. In alternativa mostriamo, mediante un esempio, un modo semplice per il calcolo delle inverse destre mediante la soluzione di sistemi lineari.

Esempio 1.7.4. Per trovare tutte le inverse destre della matrice

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & -1 & 0 & 2 \\ 0 & 0 & 2 & 2 & -1 \\ -1 & 2 & 0 & -1 & -2 \end{bmatrix}$$

È facile verificare che il blocco di $\mathbf{A}$ formato dalle prime tre colonne non è invertibile: effettuandone la riduzione, il rango non è 3 ma 2:

$$\begin{bmatrix} 1 & -2 & -1 \\ 0 & 0 & 2 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 2 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Perciò la proposizione 1.7.2 non è applicabile. Si tratta di trovare tutte le matrici 5×3 $\mathbf{R} = [\mathbf{r}_1\, \mathbf{r}_2\, \mathbf{r}_3]$ tali che $\mathbf{AR} = \mathbf{I}_3$ o, equivalentemente, si tratta di risolvere i tre sistemi lineari $\mathbf{Ax} = \mathbf{e}_j$ $(j = 1, \ldots, 3)$. Conviene a tal fine considerare la matrice pluri-aumentata

$$[\mathbf{A}|\mathbf{I}_3] = \left[\begin{array}{ccccc|ccc} 1 & -2 & -1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & -1 & 0 & 1 & 0 \\ -1 & 2 & 0 & -1 & -2 & 0 & 0 & 1 \end{array}\right]$$

La EG su tale matrice produce la forma ridotta

$$\left[\begin{array}{ccccc|ccc} 1 & -2 & -1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -\frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 & -1 & -2 \end{array}\right]$$

Con facili calcoli si vede che il sistema $\mathbf{Ax} = \mathbf{e}_1$ ha le infinite soluzioni dipendenti dai parametri h_1, k_1:

$$x_5 = -2, \quad x_4 = h_1, \quad x_3 = -1 - h_1 \quad x_2 = k_1 \quad x_1 = 4 - h_1 + 2k_1$$

Analogamente, il sistema $\mathbf{Ax} = \mathbf{e}_2$ ha le infinite soluzioni dipendenti dai parametri h_2, k_2

$$x_5 = -1, \quad x_4 = h_2, \quad x_3 = -h_2, \quad x_2 = k_2, \quad x_1 = 2 - h_2 + 2k_2$$

Infine il sistema $\mathbf{Ax} = \mathbf{e}_3$ ha le infinite soluzioni dipendenti dai parametri h_3, k_3:

$$x_5 = -2, \quad x_4 = h_3, \quad x_3 = -1 - h_3, \quad x_2 = k_3, \quad x_1 = 3 - h_3 + 2k_3$$

In definitiva, le infinite inverse destre della matrice $\mathbf{A}$ sono tutte e sole le matrici dipendenti da 6 parametri

$$\begin{bmatrix} 4 - h_1 + 2k_1 & 2 - h_2 + 2k_2 & 3 - h_3 + 2k_3 \\ k_1 & k_2 & k_3 \\ -1 - h_1 & -h_2 & -1 - h_3 \\ h_1 & h_2 & h_3 \\ -2 & -1 & -2 \end{bmatrix}$$

Esempio 1.7.5. Trovare tute le inverse destre della matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Si ha che il determinante di $\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ è $1(-1) - 1(-1) = 0$ pertanto la matrice non è invertibile. Procediamo allora mediante sistema lineare:

$$\left[\begin{array}{ccc|cc} 1 & 1 & -1 & 1 & 0 \\ -1 & -1 & 2 & 0 & 1 \end{array}\right] \left[\begin{array}{ccc|cc} 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 0 \end{array}\right] \left[\begin{array}{ccc|cc} 1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

A questo punto considerando a turno la quarta colonna si ha:

$$\begin{cases} x_3 = 1 \\ x_2 = h \\ x_1 + h - 1 = 1 \end{cases} \qquad \begin{cases} x_1 = 2 - h \\ x_2 = h \\ x_3 = 1 \end{cases}$$

e considerando la quinta

$$\begin{cases} x_3 = 1 \\ x_2 = k \\ x_1 + k - 1 = 0 \end{cases} \qquad \begin{cases} x_1 = 1 - k \\ x_2 = k \\ x_3 = 1 \end{cases}$$

Pertanto

$$\mathbf{R}(\mathbf{A}) = \begin{bmatrix} 2-h & 1-k \\ h & k \\ 1 & 1 \end{bmatrix}$$

con $h, k \in \mathbb{C}$

1.7.3 Inversa sinistra di matrice verticale

Osservazione 78. Infine, per trovare le inverse sinistre di una matrice $m \times n$ $\mathbf{A}$ di rango n, con $m > n$, basta osservare che una matrice $\mathbf{L}$ è inversa sinistra di $\mathbf{A}$ se e solo se la matrice $\mathbf{L}^T$ è inversa destra della matrice $\mathbf{A}^T$, in quanto

$$\mathbf{LA} = \mathbf{I} \overset{(1)}{\iff} (\mathbf{LA})^T = \mathbf{I}^T \iff \mathbf{A}^T\mathbf{L}^T = \mathbf{I}$$

dove in (1) si sono trasposti entrambi i membri.
Pertanto si procede determinano le inverse destre di $\mathbf{A}^T$ col procedimento sopra descritto e poi le si traspone, ottenendo le inverse sinistra di $\mathbf{A}$.

Esempio 1.7.6. Si trovino tutte le inverse destre della matrice

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

e tutte le inverse sinistre di $\mathbf{A}^T$.
Il determinante di $\begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = 0 - 1 \neq 0$ quindi la matrice è invertibile e si può applicare il metodo senza eliminazione di Gauss. Si ha:

$$\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} a & b \end{bmatrix}, \quad \mathbf{B}^{-1} = -1 \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix}$$

Per cui:

$$\begin{aligned} \mathbf{B}^{-1} - \mathbf{B}^{-1}\mathbf{CY} &= \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} - \begin{bmatrix} 3 \\ -6 \end{bmatrix} \begin{bmatrix} a & b \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 \\ 1 & -2 \end{bmatrix} - \begin{bmatrix} 3a & 3b \\ -6a & -6b \end{bmatrix} = \begin{bmatrix} -3a & 1-3b \\ 1+6a & -2+6b \end{bmatrix} \end{aligned}$$

Pertanto

$$\mathbf{R}(\mathbf{A}) = \begin{bmatrix} -3a & 1-3b \\ 1+6a & -2+6b \\ a & b \end{bmatrix}$$

Per l'inversa sinistra di $\mathbf{A}^T$ basta trasporre $R(\mathbf{A})$, ossia è

$$\begin{bmatrix} -3a & 1+6a & a \\ 1-3b & -2+6b & b \end{bmatrix}$$

1.8 Matrici elementari e decomposizione LU

1.8.1 Matrici elementari

Definizione 1.8.1 (Matrice elementare)**.** Si tratta di una matrice ottenuta a partire da una matrice identità mediante una singola operazione su una riga (o equivalentemente su una colonna) a scelta tra scambio di righe, moltiplicazione per uno scalare non nullo, rimpiazzo di una riga con la somma di questa con il multiplo di una riga che le sta sopra.

Osservazione 79. Esistono tre tipologie di matrici elementari; queste matrici derivano la loro importanza dal fatto che ogni singola operazione elementare che si esegue nel corso della EG corrisponde alla pre-moltiplicazione per una matrice elementare.

Definizione 1.8.2 (Matrici $\mathbf{E}_{ij}(\alpha)$)**.** Una matrice elementare $m \times m$ $\mathbf{E}_{ij}(\alpha)$ differisce dalla matrice identità $\mathbf{I}_m$ solo nel coefficiente di posto (i, j), con $i \neq j$ dove anziché 0 compare lo scalare $\alpha \neq 0$. Scritte a blocchi riga e colonna:

$$
\mathbf{E}_{ij}(\alpha) = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_i^T + \alpha \mathbf{e}_j^T \\ \dots \\ \mathbf{e}_m^T \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_j + \alpha \mathbf{e}_i & \dots & \mathbf{e}_m \end{bmatrix}
$$

dove la riga $\mathbf{e}_i^T + \alpha \mathbf{e}_j^T$ compare all'i-esimo posto, mentre la colonna $\mathbf{e}_j + \alpha \mathbf{e}_i$ compare al posto j-esimo

Esempio 1.8.1.

$$
\mathbf{E}_{31}(-1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}
$$

Osservazione 80. Queste matrici corrispondono all'operazione di sommare una riga i con un multiplo α di un'altra riga. Tale proprietà viene dimostrata in seguito; qui la vediamo all'opera con un esempio.

Esempio 1.8.2. Pre-moltiplicando per $\mathbf{E}_{31}(-1)$ una data matrice si toglie alla terza riga la prima. Post moltiplicando si toglie alla prima colonna la terza. Un esempio con matrici 3×3:

```
m <- rmatrix(1:9, nrow = 3)

elematrix(n = 3, i = 3, j = 1, alpha = -1)

##      [,1] [,2] [,3]
## [1,]    1    0    0
## [2,]    0    1    0
## [3,]   -1    0    1

elematrix(n = 3, i = 3, j = 1, alpha = -1) %*% m
```

```
##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    4    5    6
## [3,]    6    6    6
```

```r
m %*% elematrix(n = 3, i = 3, j = 1, alpha = -1)
```

```
##      [,1] [,2] [,3]
## [1,]   -2    2    3
## [2,]   -2    5    6
## [3,]   -2    8    9
```

Definizione 1.8.3 (Matrici elementari $\mathbf{E}_{ij}$, matrici di trasposizione)**.** Una matrice elementare $m \times m$ $\mathbf{E}_{ij}$ (anche detta matrice di trasposizione) differisce dalla matrice identità $\mathbf{I}_m$ per lo scambio tra loro della riga (equivalentemente colonna) i-esima e la j-esima

Esempio 1.8.3.

$$E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Osservazione 81. Corrispondono (attraverso la pre-moltiplicazione) all'operazione di scambio di due righe nella EG; se utilizzate in post-moltiplicazione scambiano due colonne tra loro.

Esempio 1.8.4. Per scambiare la prima e la terza riga (o colonna, rispettivamente) della matrice `m` si ha:

```r
elematrix(n = 3, i = 3, j = 1)
```

```
##      [,1] [,2] [,3]
## [1,]    0    0    1
## [2,]    0    1    0
## [3,]    1    0    0
```

```r
elematrix(n = 3, i = 3, j = 1) %*% m
```

```
##      [,1] [,2] [,3]
## [1,]    7    8    9
## [2,]    4    5    6
## [3,]    1    2    3
```

```r
m %*% elematrix(n = 3, i = 3, j = 1)
```

```
##      [,1] [,2] [,3]
## [1,]    3    2    1
## [2,]    6    5    4
## [3,]    9    8    7
```

Definizione 1.8.4 (Matrici elementari $\mathbf{E}_i(\alpha)$)**.** Sono matrici diagonali che hanno sulla diagonale tutti 1 tranno che nel posto i-esimo, dove compare il coefficiente $\alpha \neq 0$

Esempio 1.8.5.

$$\mathbf{E}_2(5) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Osservazione 82. Corrispondono, tramite pre-moltiplicazione, all'operazione di moltiplicazione della riga i-esima per α. In post-moltiplicazione alla moltiplicazione della colonna i-esima per α

Esempio 1.8.6.
```
elematrix(n = 3, i = 2, alpha = 5)

##      [,1] [,2] [,3]
## [1,]    1    0    0
## [2,]    0    5    0
## [3,]    0    0    1

elematrix(n = 3, i = 2, alpha = 5) %*% m

##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]   20   25   30
## [3,]    7    8    9

m %*% elematrix(n = 3, i = 2, alpha = 5)

##      [,1] [,2] [,3]
## [1,]    1   10    3
## [2,]    4   25    6
## [3,]    7   40    9
```

Osservazione 83. Vediamo alcune proprietà delle matrici elementari.

Proposizione 1.8.1. *Tutte le matrici elementari sono invertibili e si ha:*

$$\begin{aligned} &\mathbf{E}_{ij}(\alpha)^{-1} = \mathbf{E}_{ij}(-\alpha) \\ &\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij} \\ &\mathbf{E}_i(\alpha)^{-1} = \mathbf{E}_i(\alpha^{-1}) \end{aligned}$$

In particolare l'inversa è ancora una matrice elementare.

Dimostrazione. Per la verifica della prima uguaglianza, a titolo di esempio:

$$\mathbf{E}_{ij}(\alpha)\mathbf{E}_{ij}(-\alpha) = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_i^T + \alpha\mathbf{e}_j^T \\ \dots \\ \mathbf{e}_m^T \end{bmatrix} \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_j - \alpha\mathbf{e}_i & \dots & \mathbf{e}_m \end{bmatrix}$$

dove la riga $\mathbf{e}_i^T + \alpha\mathbf{e}_j^T$ compare all'i-esimo posto, mentre la colonna $\mathbf{e}_j - \alpha\mathbf{e}_i$ al j-esimo posto. Vogliamo provare che questo prodotto coincide con $\mathbf{I}_m$. Tenuto conto che $\mathbf{e}_i^T\mathbf{e}_i = 1$ (così abbiamo sistemato la diagonale) e che $\mathbf{e}_i^T\mathbf{e}_j = 0$ se $i \neq j$ si vede che il generico coefficiente di posto (i, j) (sempre con $i \neq j$) del prodotto risulta uguale a

$$(\mathbf{e}_i^T + \alpha\mathbf{e}_j^T)(\mathbf{e}_j - \alpha\mathbf{e}_i) = -\alpha + \alpha = 0$$

□

Proposizione 1.8.2. *Per le trasposte delle matrici elementari si ha*

$$\begin{aligned}
&\mathbf{E}_{ij}(\alpha)^T = \mathbf{E}_{ji}(\alpha)\\
&\mathbf{E}_{ij}^T = \mathbf{E}_{ij}\\
&\mathbf{E}_i(\alpha)^T = \mathbf{E}_i(\alpha)
\end{aligned}$$

Dimostrazione. Verifiche immediate. □

Proposizione 1.8.3. *Un prodotto di matrici elementari del tipo $\mathbf{E}_{ij}(\alpha)$ con $i > j$ è una matrice uni-triangolare inferiore*

Dimostrazione. Deriva dal fatto che se $i > j$ ogni matrice $\mathbf{E}_{ij}(\alpha)$ è uni-triangolare inferiore e il prodotto di queste rimane uni-triangolare inferiore (esempio 1.2.7). □

Proposizione 1.8.4. *Un prodotto di matrici elementari del tipo $\mathbf{E}_{ij}$ è una matrice che si ottiene dalla matrice identità, scambiandone le righe (o le colonne) in modo opportuno*

Dimostrazione. Immediata conseguenza delle proprietà di queste matrici □

Definizione 1.8.5 (Matrici di permutazione)**.** Matrici che si ottengono moltiplicando tra loro matrici elementari del tipo $\mathbf{E}_{ij}$.

Esempio 1.8.7. Le permutazioni di n oggetti (ad esempio righe) sono $n!$. Per cui le possibili matrici di permutazione 3×3 sono 3! (tutti i modi con cui si possono permutare le righe:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}, \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}, \begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix},$$
$$\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}, \begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}, \begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}.$$

Osservazione 84. Le matrici di permutazione sono caratterizzate dall'avere in ogni riga e in ogni colonna tutti gli elementi nulli tranne uno (uguale a 1). Inoltre sono invertibili (in quanto prodotto di matrici invertibili).

Osservazione 85. Dimostriamo ora gli effetti della pre-moltiplicazione di una generica matrice per le matrici elementari

Proposizione 1.8.5. *Sia $\mathbf{A}$ una matrice $m \times n$;*

1. *la pre-moltiplicazione di* $\mathbf{A}$ *per la matrice* $m \times m$ $\mathbf{E}_{ij}(\alpha)$ *modifica la sola riga i-esima di* $\mathbf{A}$*, che diventa uguale alla somma della riga* $i-$*esima con la riga* j*-esima moltiplicata per* α

2. *la pre-moltiplicazione di* $\mathbf{A}$ *per la matrice* $m \times m$ $\mathbf{E}_{ij}$ *scambia tra loro la riga i-esima e la riga* j*-esima di* $\mathbf{A}$ *e lascia invariate le altre righe*

3. *la pre-moltiplicazione di* $\mathbf{A}$ *per la matrice* $m \times m$ $\mathbf{E}_i(\alpha)$ *moltiplica la riga i-esima di* $\mathbf{A}$ *per* α *e lascia invariate le altre righe*

Dimostrazione. Per il primo punto eseguiamo il prodotto $\mathbf{E}_{ij}(\alpha)\mathbf{A}$ decomponendo $\mathbf{E}_{ij}(\alpha)$ in blocchi riga

$$\mathbf{E}_{ij}(\alpha)\mathbf{A} = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_i^T + \alpha\mathbf{e}_j^T \\ \dots \\ \mathbf{e}_m^T \end{bmatrix} \mathbf{A} = \begin{bmatrix} \mathbf{e}_1^T\mathbf{A} \\ \mathbf{e}_2^T\mathbf{A} \\ \vdots \\ (\mathbf{e}_i^T + \alpha\mathbf{e}_j^T)\mathbf{A} \\ \dots \\ \mathbf{e}_m^T\mathbf{A} \end{bmatrix}$$

e si conclude osservando che $\mathbf{e}_i^T\mathbf{A}$ coincide con la i-esima riga di $\mathbf{A}$.
Per il secondo punto la dimostrazione è analoga alla precedente, tenendo conto che la i-esima riga di $\mathbf{E}_{ij}$ coincide con $\mathbf{e}_j^T$ e la j-esima riga coincide con $\mathbf{e}_i^T$.
Per la terza, poiché $\mathbf{E}_i(\alpha)$ è una matrice diagonale, $\mathbf{E}_i(\alpha)\mathbf{A}$ coincide con $\mathbf{A}$ tranne che per la i-esima riga, che si ottiene dalla i-esima riga di $\mathbf{A}$ moltiplicandola per α. □

Osservazione 86. Vediamo ora gli effetti della post moltiplicazione: vale un risultato analogo alla proposizione 1.8.5 che riguarda non più le righe ma bensì le colonne di $\mathbf{A}$

Proposizione 1.8.6. *Sia* $\mathbf{A}$ *una matrice* $m \times n$*:*

- *la post-moltiplicazione di* $\mathbf{A}$ *per la matrice* $n \times n$ $\mathbf{E}_{ij}(\alpha)$ *modifica la sola colonna* j*-esima di* $\mathbf{A}$*, che diventa uguale alla somma della colonna* j*-esima con la colonna i-esima moltiplicata per* α*;*

- *la post-moltiplicazione di* $\mathbf{A}$ *per la matrice* $n \times n$ $\mathbf{E}_{ij}$ *scambia tra loro la colonna i-esima e la colonna* j*-esima di* $\mathbf{A}$ *e lascia invariate le altre colonne;*

- *la post-moltiplicazione di* $\mathbf{A}$ *per la matrice* $n \times n$ $\mathbf{E}_i(\alpha)$ *moltiplica la colonna i-esima di* $\mathbf{A}$ *per* α *e lascia invariate la altre colonne*

Dimostrazione. Analoga a quella della proposizione 1.8.5 considerando che la j-esima colonna di $\mathbf{AB}$ coincide con il prodotto $\mathbf{Ab}_j$, ossia di $\mathbf{A}$ per la j-esima colonna $\mathbf{b}_j$ di $\mathbf{B}$ □

1.8.2 Applicazioni

1.8.2.1 Fattorizzazione LU

Osservazione 87. La proposizione 1.8.5 consente di tradurre in fattorizzazione matriciale la EG eseguita su una matrice.

Teorema 1.8.7. *Sia* $\mathbf{A}$ *una matrice* $m \times n$*. La EG su* $\mathbf{A}$*, qualora non necessiti di scambi di righe, produce una decomposizione* $\mathbf{A} = \mathbf{LU}$ *con* $\mathbf{U}$ *forma ridotta di* $\mathbf{A}$ *e* $\mathbf{L}$ *matrice triangolare inferiore invertibile*

Dimostrazione. Se non intervengono swap di riga, la EG consiste nell'applicare un certo numero di operazioni elementari corrispondenti alla pre-moltiplicazione per matrici elementari di tipo $\mathbf{E}_{ij}(\alpha)$ o $\mathbf{E}_i(\alpha)$. Se denotiamo con $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_r$ tali matrici elementari, nell'ordine in cui sono applicate (pre-moltiplicate), si avrà

$$\mathbf{E}_r \dots \mathbf{E}_2 \mathbf{E}_1 \mathbf{A} = \mathbf{U}$$

con $\mathbf{U}$ una forma ridotta di $\mathbf{A}$. Sia

$$\mathbf{E} = \mathbf{E}_r \dots \mathbf{E}_2 \mathbf{E}_1$$

il prodotto delle matrici elementari usate; $\mathbf{E}$, in quanto prodotto di matrici triangolari inferiori[3] invertibili (utilizzando matrici elementari), è triangolare inferiore (per esempio 1.2.7) ed invertibile (per proposizione 1.6.13). Poniamo $\mathbf{L} = \mathbf{E}^{-1}$; tale matrice è ancora triangolare inferiore (esempio 1.6.2) e risulta $\mathbf{A} = \mathbf{LU}$ □

Esempio 1.8.8. Riprendendo l'esempio 1.5.2 e considerando solo la matrice dei coefficienti

$$\mathbf{A} = \begin{bmatrix} 2 & 6 & 3 & 2 \\ 1 & -2 & \frac{1}{2} & \frac{9}{4} \\ -1 & 1 & -\frac{1}{2} & -1 \end{bmatrix}$$

Le operazioni elementari applicate sono state, nell'ordine:

1. moltiplicare la prima riga per 1/2 (corrisponde alla pre-moltiplicazione per $\mathbf{E}_1 = \mathbf{E}_1(1/2)$)

2. sommare alla seconda riga la prima moltiplicata per -1 ($\mathbf{E}_2 = \mathbf{E}_{21}(-1)$)

3. sommare alla terza riga la prima moltiplicata per 1 ($\mathbf{E}_3 = \mathbf{E}_{31}(1)$)

4. moltiplicare la seconda riga per $-1/5$ ($\mathbf{E}_4 = \mathbf{E}_2(-1/5)$)

5. sommare alla terza riga la seconda moltiplicata per -4 ($\mathbf{E}_5 = \mathbf{E}_{32}(-4)$)

6. moltiplicare la terza riga per 5 ($\mathbf{E}_6 = \mathbf{E}_3(5)$)

La forma ridotta cui si perviene al termine (qui rinominata in $\mathbf{U}$) è

$$\mathbf{U} = \begin{bmatrix} 1 & 3 & 3/2 & 1 \\ 0 & 1 & 1/5 & -1/4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

[3] $E_i(\alpha)$ sono diagonali e quindi triangolari; per quanto riguarda $E_{ij}(\alpha)$ ci limitiamo ad utilizzare matrici con $i > j$, triangolari inferiori; se non servono swap di riga, come si vedrà nell'esempio seguente, per arrivare alla forma ridotta ci basta di fatto modificare le righe in basso utilizzando quelle più in alto.

Un calcolo diretto mostra che

$$
\begin{aligned}
\mathbf{E} &= \mathbf{E}_6\mathbf{E}_5\mathbf{E}_4\mathbf{E}_3\mathbf{E}_2\mathbf{E}_1 = \mathbf{E}_3(5)\mathbf{E}_{32}(-4)\mathbf{E}_2(-1/5)\mathbf{E}_{31}(1)\mathbf{E}_{21}(-1)\mathbf{E}_1(1/2) \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1/5 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1/2 & 0 & 0 \\ 1/10 & -1/5 & 0 \\ 1/2 & 4 & 5 \end{bmatrix}
\end{aligned}
$$

Il calcolo con l'algoritmo di Gauss-Jordan di $\mathbf{L} = \mathbf{E}^{-1}$ porge

$$
\mathbf{L} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -5 & 0 \\ -1 & 4 & 1/5 \end{bmatrix}
$$

e come si può verificare $\mathbf{A} = \mathbf{LU}$.

Osservazione 88. Si osservi come nell'esempio precedente si ha che

$$
\mathbf{L} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{E}_3^{-1}\mathbf{E}_4^{-1}\mathbf{E}_5^{-1}\mathbf{E}_6^{-1} = \mathbf{E}_1(2)\mathbf{E}_{21}(1)\mathbf{E}_{31}(-1)\mathbf{E}_2(-5)\mathbf{E}_{32}(4)\mathbf{E}_3(1/5)
$$

dove i coefficienti non nulli che compaiono ordinatamente dall'alto verso il basso nella prima colonna di $\mathbf{L}$, poi nella seconda e nella terza coincidono con gli scalari che compaiono tra parentesi (da sinistra verso destra) nelle matrici elementari. Questo pone la domanda se si possano descrivere facilmente gli elementi della matrice $\mathbf{L}$ a partire dalle operazioni elementari eseguite. Prima però vediamo il caso in cui nella EG intercorrano permutazioni di righe.

Osservazione 89. Se nella EG avvengono permutazioni di righe, ordinatamente $\mathbf{E}_{i_1,j_1}, \mathbf{E}_{i_2,j_2}, \ldots, \mathbf{E}_{i_k,j_k}$, indichiamo con $\mathbf{P}$ la matrice di permutazione per la pre-moltiplicazione di $\mathbf{A}$, ossia

$$
\mathbf{P} = \mathbf{E}_{i_k,j_k} \ldots \mathbf{E}_{i_2,j_2}\mathbf{E}_{i_1,j_1}
$$

La matrice $\mathbf{P}$, in quanto prodotto di matrici invertibili, è invertibile. L'idea qui è applicare la EG senza scambi di riga (con la fattorizzazione $\mathbf{LU}$ che ne segue), in seguito allo scambio di righe necessarie, ossia alla pre-moltiplicazione di $\mathbf{A}$ per $\mathbf{P}$.

Teorema 1.8.8. *Sia* $\mathbf{A}$ *una matrice* $m \times n$. *Qualora la EG su* $\mathbf{A}$ *necessiti di scambi di righe, si possono eseguire subito tali scambi su* $\mathbf{A}$ *ottenendo la nuova matrice* $\mathbf{PA}$, *ove* $\mathbf{P}$ *denota una matrice di permutazione. La EG su* $\mathbf{PA}$ *produce una fattorizzazione* $\mathbf{PA} = \mathbf{LU}$ *dove* $\mathbf{U}$ *è una forma ridotta di* $\mathbf{A}$ *e* $\mathbf{L}$ *matrice triangolare inferiore invertibile. Pertanto, pre-moltiplicando entrambi i membri per* $\mathbf{P}^{-1}$, *si può scomporre* $\mathbf{A}$ *come*

$$
\mathbf{A} = \mathbf{P}^{-1}\mathbf{LU}
$$

Dimostrazione. La EG sulla matrice $\mathbf{PA}$ non necessita di scambi di righe (Libro pag 52 per dimostrazione, nel caso). Ovvero se si eseguono subito su $\mathbf{A}$ ordinatamente tutti gli scambi di righe che si usano nel corso della EG e si procede poi sulla matrice ottenuta con la EG, non sono più necessari scambi di righe. □

TODO: fixme

Esempio 1.8.9. Considerando la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 1 & 2 & 2 & 2 & 2 \\ 2 & 5 & -1 & 3 & 4 \\ 1 & 2 & 5 & 1 & 2 \end{bmatrix}$$

Sommando alla seconda riga l'opposto della prima, alla terza riga la prima moltiplicata per -2, e alla quarta riga l'opposto della prima si ricava

$$\begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 3 & 0 & 1 \end{bmatrix}$$

Eseguiamo lo scambio di seconda a terza riga, poi nella matrice ottenuta quello di terza e quarta, ottenendo

$$\begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

dalla quale moltiplicando la terza riga per $1/3$ si ottiene la forma ridotta

$$\begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

Le operazioni eseguite corrispondono alla pre-moltiplicazione di $\mathbf{A}$ come segue

$$\mathbf{E}_3(1/3)\mathbf{E}_{34}\mathbf{E}_{23}\mathbf{E}_{41}(-1)\mathbf{E}_{31}(-2)\mathbf{E}_{21}(-1)\mathbf{A}$$

Pre-moltiplicando $\mathbf{A}$ per la matrice di permutazione

$$\mathbf{P} = \mathbf{E}_{34}\mathbf{E}_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

la nuova matrice

$$\mathbf{PA} = \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 2 & 5 & -1 & 3 & 4 \\ 1 & 2 & 5 & 1 & 2 \\ 1 & 2 & 2 & 2 & 2 \end{bmatrix}$$

non avrà bisogno, nell'eseguire la EG di alcuno scambio di righe. In effetti, la EG su $\mathbf{PA}$ porge

$$\mathbf{E}_3(1/3)\mathbf{E}_{41}(-1)\mathbf{E}_{31}(-1)\mathbf{E}_{21}(-2)\mathbf{PA} = \mathbf{U}$$

Quindi si ha la decomposizione $\mathbf{PA} = \mathbf{LU}$ dove $\mathbf{L}$ triangolare inferiore è data da

$$\mathbf{L} = \mathbf{E}_{21}(2)\mathbf{E}_{31}(1)\mathbf{E}_{41}(1)\mathbf{E}_3(3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Ne consegue che per $\mathbf{A}$ si ha la fattorizzazione

$$\begin{aligned}
\mathbf{A} &= \mathbf{P}^{-1}\mathbf{L}\mathbf{U} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 1 & 1 \\ 0 & 1 & -5 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1/3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}
\end{aligned}$$

Si noti che il primo fattore nell'ultimo prodotto non è una matrice triangolare inferiore; resta però una matrice invertibile.

Osservazione 90. Si noti che nel teorema 1.8.8 la matrice $\mathbf{P}^{-1}\mathbf{L}$ è invertibile (prodotto di matrici invertibili) e che, se $\mathbf{P} = \mathbf{E}_{i_k,j_k} \dots \mathbf{E}_{i_2,j_2}\mathbf{E}_{i_1,j_1}$ allora:

$$\begin{aligned}
\mathbf{P}^{-1} &= (\mathbf{E}_{i_k,j_k} \dots \mathbf{E}_{i_1,j_1})^{-1} = \mathbf{E}_{i_1,j_1}^{-1} \dots \mathbf{E}_{i_k,j_k}^{-1} = \mathbf{E}_{i_1,j_1} \dots \mathbf{E}_{i_k,j_k} \\
&= \mathbf{E}_{i_1,j_1}^T \dots \mathbf{E}_{i_k,j_k}^T = (\mathbf{E}_{i_k,j_k} \dots \mathbf{E}_{i_1,j_1})^T \\
&= \mathbf{P}^T
\end{aligned}$$

in virtù del fatto che $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij} = \mathbf{E}_{ij}^T$ per ogni $\mathbf{E}_{ij}$.

Osservazione 91. Dato che la matrice $\mathbf{P}^{-1}\mathbf{L}$ è invertibile, il teorema 1.8.8 fornisce la dimostrazione del punto 2 di lemma 1.6.2

Osservazione 92. Soprattutto se si fanno i calcoli a mano può essere utile un modo di determinare facilmente/direttamente i coefficienti della matrice triangolare inferiore $\mathbf{L}$ che fattorizza $\mathbf{A}$ (senza passare dal prodotto delle rispettive matrici elementari). Consideriamo solo il caso in cui non si necessitino scambi di righe (se ce ne fossero faremmo la stessa cosa per $\mathbf{PA}$). Vale la seguente proposizione (che di fatto descrive quello che è stato visto nell'osservazione 88 inerente l'esempio 1.8.8.

Proposizione 1.8.9. *Sia* $\mathbf{A}$ *una matrice per cui la EG non necessiti di scambi di righe. La matrice triangolare inferiore invertibile* $\mathbf{L}$ *ottenuta mediante la fattorizzazione* $\mathbf{A} = \mathbf{LU}$ *ha come coefficiente di posto* (i, j)*:*

- *lo scalare* $-\alpha$ *se* $i > j$ *dove* $\mathbf{E}_{ij}(\alpha)$ *è la matrice elementare usata per porre 0 al posto* (i, j)*, oppure 0 se non si è eseguita alcuna operazione elementare*
- *lo scalare* α^{-1} *se* $i = j$ *dove* $\mathbf{E}_i(\alpha)$ *è la matrice elementare usata per porre 1 al posto* (i, i)*, oppure 1 se non si è eseguita alcuna operazione elementare*

Dimostrazione. Salce pag 55-56 □

Osservazione 93. Se nella fattorizzazione $\mathbf{LU}$ si desidera che gli elementi l_{ii} sulla diagonale di $\mathbf{L}$ sian tutti uguali a 1 (non è detto che lo siano, a differenza della diagonale di $\mathbf{U}$) si può (effettuando dell'algebra basica):

- dividere ogni colonna di $\mathbf{L}$ per il suo coefficiente diagonale, ottenendo una nuova matrice $\mathbf{L}_0$ che ha i primi elementi non nulli di ogni colonna a 1;

- post moltiplicare per $\mathbf{D} = \mathbf{Diag}(l_{11}, \ldots, l_{nn})$.

Così facendo si fattorizza anche $\mathbf{L} = \mathbf{L}_0\mathbf{D}$ e la decomposizione $\mathbf{A} = \mathbf{LU}$ pertanto diviene $\mathbf{A} = \mathbf{L}_0\mathbf{DU}$.

Esempio 1.8.10. La decomposizione **LU** della matrice **B** dell'esempio 1.8.8 è

$$\mathbf{B} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & -5 & 0 \\ -1 & 4 & 1/5 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3/2 & 1 \\ 0 & 1 & 1/5 & -1/4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

mentre la $\mathbf{B} = \mathbf{L}_0\mathbf{DU}$ è

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ -1/2 & 4/5 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1/5 \end{bmatrix} \begin{bmatrix} 1 & 3 & 3/2 & 1 \\ 0 & 1 & 1/5 & -1/4 \\ 0 & 0 & 1 & 5 \end{bmatrix}$$

1.8.2.2 Decomposizioni a rango pieno

Si è chiamato *rango* il numero k di righe non nulle di una forma ridotta **U** della matrice $m \times n$ **A**.
Se la EG su **A** non richiede scambi di righe possiamo ottenere una decomposizione $\mathbf{A} = \mathbf{LU}$ con **U** che ha $m - k$ righe nulle (se k, rango sono le non nulle). Queste righe nulle nel prodotto della fattorizzazione vengono moltiplicate per le ultime $m - k$ colonne di **L**:

$$\mathbf{A} = \mathbf{LU} = \begin{bmatrix} \mathbf{l}_1 & \mathbf{l}_2 & \ldots & \mathbf{l}_m \end{bmatrix} \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \ldots \\ \mathbf{u}_m^T \end{bmatrix} = \mathbf{l}_1\mathbf{u}_1^T + \ldots + \mathbf{l}_k\mathbf{u}_k^T + \mathbf{l}_{k+1}\mathbf{0}^T + \ldots \mathbf{l}_m\mathbf{0}^T$$

Quindi ignorando ciò che si perde a causa dei vettori nulli si può fattorizzare **A** ugualmente come $\mathbf{A} = \mathbf{BC}$ dove

$$\mathbf{B} = \begin{bmatrix} \mathbf{l}_1 & \mathbf{l}_2 & \ldots & \mathbf{l}_k \end{bmatrix} \qquad \mathbf{C} = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \ldots \\ \mathbf{u}_k^T \end{bmatrix}$$

Osservazione 94. Queste decomposizioni $\mathbf{A} = \mathbf{BC}$ si chiamano decomposizioni a *rango pieno* e sono caratterizzate dal fatto che il numero di colonne del primo fattore (**B**) coincide con il numero di righe del secondo fattore (**C**) ed è uguale al rango di **A** (ossia k).

Osservazione 95. Questa decomposizione testimonia che ogni matrice **A** può essere fattorizzata in maniera tale da soddisfare le ipotesi della proposizione 1.6.20, che assicura l'esistenza della matrice pseudo-inversa; questo perché è immediato verificare che **B** ha inversa sinistra (essendo formata dalle prime k colonne di una matrice invertibile, vedi esercizio 1.27), e che **C** ha inversa destra (essendo matrice in forma ridotta senza righe nulle).

Esempio 1.8.11. Per un esempio di decomposizione a rango pieno, data la matrice

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 4 \\ 2 & -2/3 & 5/3 \\ -6 & 2 & -8 \end{bmatrix}$$

la EG, non essendo necessari scambi di righe, produce

$$\mathbf{E}_2(-1)\mathbf{E}_{31}(6)\mathbf{E}_{21}(-2)\mathbf{E}_1(1/3)\mathbf{A} = \begin{bmatrix} 1 & -1/3 & 4/3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \mathbf{U}$$

Quindi si ha la decomposizione

$$\begin{aligned} \mathbf{A} = \mathbf{LU} &= \mathbf{E}_1(3)\mathbf{E}_{21}(2)\mathbf{E}_{31}(-6)\mathbf{E}_2(-1)\mathbf{U} \\ &= \begin{bmatrix} 3 & 0 & 0 \\ 2 & -1 & 0 \\ -6 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1/3 & 4/3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Ne consegue che una decomposizione a rango pieno di $\mathbf{A}$ è data da

$$\mathbf{A} = \begin{bmatrix} 3 & 0 \\ 2 & -1 \\ -6 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1/3 & 4/3 \\ 0 & 0 & 1 \end{bmatrix}$$

1.8.2.3 Altro

Osservazione 96. Possiamo dimostrare il lemma 1.6.2, prima parte

Lemma 1.8.10. *Data una matrice* $\mathbf{A}$*, esiste una matrice invertibile* $\mathbf{E}$ *tale che*

$$\mathbf{EA} = \begin{bmatrix} a & \mathbf{x}^T \\ \mathbf{0} & \mathbf{X} \end{bmatrix}$$

Dimostrazione. Se la prima colonna di $\mathbf{A}$ è nulla, si ponga $\mathbf{E} = \mathbf{I}$, altrimenti $\mathbf{E}$ non è altro che il prodotto delle matrici elementari usate per mettere 0 sotto al primo pivot a. □

Proposizione 1.8.11. *Una matrice* $\mathbf{A}$ *è invertibile se e solo se è prodotto di matrici elementari.*

Dimostrazione. Tentativo di dimostrazione:

- sia $\mathbf{A}$ il prodotto di matrici elementari $\mathbf{E}_1, \ldots, \mathbf{E}_n$ ossia

  $$\mathbf{A} = \mathbf{E}_1 \cdot \ldots \cdot \mathbf{E}_n$$

  tutte le matrici elementari sono invertibili (per proposizione 1.8.1), quindi anche $\mathbf{A}$, essendo prodotto di matrici invertibili, lo è (avrà inversa $\mathbf{E}_n^{-1} \cdot \ldots \cdot \mathbf{E}_1^{-1}$, per proposizione 1.6.13)

- sia viceversa $\mathbf{A}$ invertibile e proviamo che è il prodotto di matrici elementari:

  – se non sono necessari scambi di riga $\mathbf{A}$ puù essere fattorizzata come $\mathbf{LU}$ con $\mathbf{U}$ forma ridotta di $\mathbf{A}$ e $\mathbf{L}$ matrice triangolare inferiore.
    Nella dimostrazione del teorema 1.8.7 si mostra che $\mathbf{L} = \mathbf{E}_1^{-1} \ldots \mathbf{E}_r^{-1}$, dove $\mathbf{E}_i$ sono le matrici elementari utilizzate per ridurre $\mathbf{A}$ in forma ridotta; ma dato che l'inversa di una matrice elementare è ancora una matrice elementare, le $\mathbf{E}_i^{-1}$ sono matrici elementari.

Venendo invece ad $\mathbf{U}$ esso può essere ricondotto al prodotto di matrici elementari svolgendo quanto fatto nel "procedimento all'indietro" dell'eliminazione di Gauss per la determinazione dell'inversa ossia mediante premoltiplicazioni per opportune matrici elementari che riconducano la matrice all'identità (dalla quale si procede poi alla determinazione dell'inversa attraverso risoluzione dei sistemi multipli):

$$\mathbf{E}_{r+1} \dots \mathbf{E}_{r+n}\mathbf{U} = \mathbf{I}$$

da cui

$$\mathbf{U} = \mathbf{I}(\mathbf{E}_{r+1} \dots \mathbf{E}_{r+n})^{-1} = \mathbf{E}_{r+n}^{-1} \dots \mathbf{E}_{r+1}^{-1}$$

pertanto si conclude che $\mathbf{A}$ è fattorizzabile nel prodotto di matrici elementari.

– se sono necessari scambi di riga $\mathbf{A}$ può esser fattorizzata come $\mathbf{P}^{-1}\mathbf{LU}$ con $\mathbf{LU}$ che viene analizzato come in precedenza e con $\mathbf{P}^{-1}$ che (teorema 1.8.8) è l'inversa del prodotto di matrici $\mathbf{E}_{ij}$, quindi ancora il prodotto di matrici elementari.

□

Capitolo 2

Spazi vettoriali

Contents

2.1 Spazi vettoriali e sottospazi

2.1.1 Introduzione

Definizione 2.1.1 (Spazio vettoriale (complesso))**.** Uno spazio vettoriale (complesso) V è un insieme, i cui elementi vengono chiamati *vettori*, sul quale sono definite le operazioni (interne all'insieme stesso):

1. addizione tra vettori: funzione $V \times V \to V$ definita come $(\mathbf{u}, \mathbf{v}) \to \mathbf{u} + \mathbf{v}$;
2. prodotto di vettore per scalare complesso: funzione $\mathbb{C} \times V \to V$ definita come $(\alpha, \mathbf{v}) \to \alpha\mathbf{v}$.

che soddisfano i seguenti assiomi, o *regole di calcolo* negli spazi vettoriali (qui $\mathbf{v}, \mathbf{w} \in V$ e $\alpha, \beta \in \mathbb{C}$):

(A_0) Proprietà commutativa[1] dell'addizione:

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v} \tag{2.1}$$

(A_1) Proprietà associativa dell'addizione:

$$(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}) \tag{2.2}$$

(A_2) Esistenza dell'elemento neutro dell'addizione: esiste un vettore $\mathbf{0} \in V$ detto vettore nullo, tale che:

$$\mathbf{v} + \mathbf{0} = \mathbf{v} \tag{2.3}$$

(A_3) Esistenza dell'opposto di un vettore dato: esiste un vettore $\mathbf{w} \in V$ detto opposto di $\mathbf{v}$ tale che:

$$\mathbf{v} + \mathbf{w} = \mathbf{0} \tag{2.4}$$

(M_1) Proprietà associativa del prodotto per uno scalare:

$$\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v} \tag{2.5}$$

(M_2) Proprietà distributiva del prodotto per uno scalare rispetto alla somma di scalari:

$$(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v} \tag{2.6}$$

(M_3) Proprietà distributiva del prodotto per uno scalare rispetto alla somma di vettori:

$$\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w} \tag{2.7}$$

(M_4) Normalizzazione del prodotto per uno scalare:

$$1\mathbf{v} = \mathbf{v} \tag{2.8}$$

Osservazione 97. Vediamo alcuni risultati immediati degli spazi vettoriali complessi.

[1] In alcune trattazioni (ad esempio Salce) questa proprietà è fatta derivare dalle altre e pertanto non costituisce assioma. Per semplicità qui la si considera come assioma

Proposizione 2.1.1. *L'elemento neutro dell'addizione è unico.*

Dimostrazione. Ipotizziamo per assurdo che oltre a $\mathbf{0}$ vi sia un altro elemento neutro dell'addizione $\mathbf{z} \in V$, allora per definizione

$$\mathbf{v} + \mathbf{z} = \mathbf{v}$$

e deve essere che la proprietà valga anche se al posto di $\mathbf{v}$ sostituiamo $\mathbf{0}$, dato che $\mathbf{0} \in V$

$$\mathbf{0} + \mathbf{z} = \mathbf{0}$$

Ma essendo $\mathbf{z} \in V$ si ha anche che

$$\mathbf{z} + \mathbf{0} = \mathbf{z}$$

Quindi eguagliando queste ultime due

$$\mathbf{0} = \mathbf{0} + \mathbf{z} = \mathbf{z} + \mathbf{0} = \mathbf{z}$$

e si conclude considerando primo e quarto membro dell'uguaglianza che $\mathbf{z} = \mathbf{0}$ □

Osservazione 98. Occhio a notare la differenza tra 0 ($\in \mathbb{C}$) e $\mathbf{0}$ ($\in V$)

Proposizione 2.1.2. *L'elemento opposto dell'addizione è unico.*

Dimostrazione. Supponiamo per assurdo che esistano due elementi $\mathbf{w}$ e $\mathbf{w}'$ opposti di $\mathbf{v}$. Si ha:

$$\mathbf{v} + \mathbf{w} = \mathbf{0}, \qquad \mathbf{v} + \mathbf{w}' = \mathbf{0}$$

Considerando allora l'elemento

$$\mathbf{u} = (\mathbf{w} + \mathbf{v}) + \mathbf{w}' = \mathbf{w} + (\mathbf{v} + \mathbf{w}')$$

applichiamo varie regole alle ultime due parti dell'uguaglianza

$$\mathbf{u} = (\mathbf{w} + \mathbf{v}) + \mathbf{w}' = \mathbf{0} + \mathbf{w}' = \mathbf{w}'$$
$$\mathbf{u} = \mathbf{w} + (\mathbf{v} + \mathbf{w}') = \mathbf{w} + \mathbf{0} = \mathbf{w}$$

Per cui $\mathbf{w} = \mathbf{w}'$ □

Osservazione 99 (Notazione opposto). D'ora in avanti denotiamo l'opposto del vettore $\mathbf{v}$ come $-\mathbf{v}$, in quanto

$$\mathbf{v} + (-\mathbf{v}) = 1\mathbf{v} + (-1)\mathbf{v} = (1 + (-1))\mathbf{v} = 0\mathbf{v} = \mathbf{0}$$

Proposizione 2.1.3. *Per ogni* $\mathbf{v} \in V$ *si ha* $0\mathbf{v} = \mathbf{0}$

Dimostrazione. Poniamo $\mathbf{u} = 0\mathbf{v}$; dato che $0 + 0 = 0$

$$\mathbf{u} = 0\mathbf{v} = (0 + 0)\mathbf{v} = 0\mathbf{v} + 0\mathbf{v} = \mathbf{u} + \mathbf{u}$$

Perciò possiamo scrivere

$$\mathbf{0} = \mathbf{u} + (-\mathbf{u}) = (\mathbf{u} + \mathbf{u}) + (-\mathbf{u}) = \mathbf{u} + (\mathbf{u} + (-\mathbf{u})) = \mathbf{u} + \mathbf{0} = \mathbf{u}$$

e quindi $\mathbf{u} = \mathbf{0}$ □

Proposizione 2.1.4. *Se* $\alpha \in \mathbb{C}$, $\mathbf{v} \in V$ *e* $\alpha\mathbf{v} = \mathbf{0}$ *si ha* $\alpha = 0$ *oppure* $\mathbf{v} = \mathbf{0}$

Dimostrazione. Se $\alpha = 0$ si verifica la 2.1.3 e non c'è nulla da dimostrare. Supponiamo allora $\alpha \neq 0$; allora

$$\mathbf{0} = \alpha^{-1}\mathbf{0} = \alpha^{-1}(\alpha\mathbf{v}) = (\alpha^{-1}\alpha)\mathbf{v} = 1\mathbf{v} = \mathbf{v}$$

da cui $\mathbf{v} = \mathbf{0}$. □

Proposizione 2.1.5. *Per ogni* $\alpha \in \mathbb{C}$ *si ha* $\alpha\mathbf{0} = \mathbf{0}$

Dimostrazione. Poniamo $\mathbf{u} = \alpha\mathbf{0}$; per la proprietà fondamentale di $\mathbf{0}$ che $\mathbf{0}+\mathbf{0} = \mathbf{0}$, abbiamo

$$\mathbf{u} = \alpha\mathbf{0} = \alpha(\mathbf{0}+\mathbf{0}) = \alpha\mathbf{0} + \alpha\mathbf{0} = \mathbf{u} + \mathbf{u}$$

perciò come prima possiamo scrivere

$$\mathbf{0} = \mathbf{u} + (-\mathbf{u}) = (\mathbf{u}+\mathbf{u}) + (-\mathbf{u}) = \mathbf{u} + (\mathbf{u} + (-\mathbf{u})) = \mathbf{u} + \mathbf{0} = \mathbf{u}$$

e possiamo concludere che $\mathbf{u} = \alpha\mathbf{0} = \mathbf{0}$ □

Proposizione 2.1.6. *Se* $\mathbf{v}, \mathbf{w} \in V$*, allora* $\mathbf{w} = -\mathbf{v}$ *se e solo se* $\mathbf{v} + \mathbf{w} = \mathbf{0}$

Dimostrazione. Se $\mathbf{w} = -\mathbf{v}$ ($\mathbf{w}$ è l'opposto di $\mathbf{v}$) allora per definizione $\mathbf{v}+\mathbf{w} = \mathbf{0}$. Viceversa se $\mathbf{0} = \mathbf{v} + \mathbf{w}$ possiamo sommare ad ambo i membri $-\mathbf{v}$:

$$-\mathbf{v} = \mathbf{v} + \mathbf{w} - \mathbf{v} = \mathbf{w} + (\mathbf{v} - \mathbf{v}) = \mathbf{w} + \mathbf{0} = \mathbf{w}$$

per cui $\mathbf{w} = -\mathbf{v}$ □

Proposizione 2.1.7. *Per ogni* $\mathbf{v} \in V$ *e* $\alpha \in \mathbb{C}$ *si ha*

$$-(\alpha\mathbf{v}) = (-\alpha)\mathbf{v} = \alpha(-\mathbf{v}) \tag{2.9}$$

Dimostrazione. Dimostriamo innanzitutto che $-(\alpha\mathbf{v}) = (-\alpha)\mathbf{v}$: ci basta mostrare che $\alpha\mathbf{v} + (-\alpha)\mathbf{v} = \mathbf{0}$:

$$\alpha\mathbf{v} + (-\alpha)\mathbf{v} = (\alpha + (-\alpha))\mathbf{v} = 0\mathbf{v} = \mathbf{0}$$

per cui per la proposizione 2.1.6 abbiamo che $-(\alpha\mathbf{v}) = (-\alpha)\mathbf{v}$. Analogamente

$$\alpha\mathbf{v} + \alpha(-\mathbf{v}) = \alpha(\mathbf{v} + (-\mathbf{v})) = \alpha\mathbf{0} = \mathbf{0}$$

e sempre per la 2.1.6 si conclude che $-(\alpha\mathbf{v}) = \alpha(-\mathbf{v})$. □

2.1.2 Esempi di spazi vettoriali

Esempio 2.1.1 (Matrici complesse: $M_{m\times n}(\mathbb{C})$)**.** Per ogni $m, n \geq 1$ l'insieme delle matrici $m \times n$ a coefficienti complessi $M_{m\times n}(\mathbb{C})$ costituisce uno spazio vettoriale, adottando le operazioni di addizione fra matrici e prodotto di matrice per uno scalare complesso come definite in precedenza e nonché matrice composta unicamente da zeri come $\mathbf{0}$ (ossia si può verificare che sotto tali premesse le regole di calcolo proprie della definizione di spazio vettoriale sono rispettate).

Esempio 2.1.2 (Vettori complessi: $\mathbb{C}^m$)**.** Più specificamente per ogni $m \geq 1$ l'insieme dei vettori a coefficienti complessi $M_{m\times 1}(\mathbb{C})$ è uno spazio vettoriale. È indicato anche con $\mathbb{C}^m$.

Esempio 2.1.3 (Matrici e vettori reali: $M_{m\times n}(\mathbb{R})$ e $\mathbb{R}^m$)**.** Analogamente $M_{m\times n}(\mathbb{R})$ e $M_{m\times 1}(\mathbb{R}) = \mathbb{R}^m$ costituiscono spazi vettoriali.

Esempio 2.1.4 (Funzioni continue sull'intervallo $[a, b]$: $\mathscr{C}([a, b])$)**.** L'insieme delle funzioni reali continue sull'intervallo $[a, b]$, indicato con $\mathscr{C}([a, b])$ costituisce uno spazio vettoriale se, per $f, g \in \mathscr{C}([a, b])$ e $t \in [a, b]$, $\alpha \in \mathbb{R}$, si definiscono:

$$(f + g)(t) = f(t) + g(t)$$
$$(\alpha f)(t) = \alpha f(t)$$

e $\mathbf{0}$ come la funzione che vale costantemente 0; infatti sia $f + g$ che αf sono ancora continue in $[a, b]$ e le regole di calcolo degli spazi vettoriali sono rispettate.

Definizione 2.1.2 (Polinomio a coefficienti reali)**.** Espressione della forma:

$$a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$$

per opportuni $a_0, a_1, \ldots, a_n \in \mathbb{R}$ detti coefficienti. La X non indica (necessariamente) un numero, ma solo un simbolo che obbedisce alle regole $X^aX^b = X^{a+b}$.

Definizione 2.1.3 (Grado polinomio)**.** È l'esponente massimo che compare tra le X del polinomio.

Definizione 2.1.4 (Polinomio nullo e suo grado)**.** È detto nullo un polinomio in cui tutti i coefficienti sono nulli, come ad esempio

$$0 + 0X^1 + 0X^2$$

Il polinomio nullo ha per convenzione grado $-\infty$.

Osservazione 100. Una costante $k \in \mathbb{C}$ ha grado 0, in quanto può essere riscritta come

$$k = kX^0 = k \cdot 1$$

Esempio 2.1.5 (Polinomi a coefficienti reali: $\mathscr{P}(\mathbb{R})$)**.** I polinomi formano uno spazio vettoriale, indicato con $\mathscr{P}(\mathbb{R})$, con l'usuale addizione e moltiplicazione per scalari (definite come in algebra) e il polinomio nullo a fungere da $\mathbf{0}$ (sotto tali condizioni le regole della definizione di spazio vettoriale sono rispettate).

Esempio 2.1.6 (Polinomi di grado $< n$: $\mathscr{P}_n(\mathbb{R})$)**.** Dati due polinomi f, g di grado $< n$, il grado di $f + g$ (o di αf) non è maggiore del grado di f e di g (o di f), la somma e la moltiplicazione per costante è operazione interna all'insieme. Pertanto l'insieme dei polinomi di grado $< n$, indicato con $\mathscr{P}_n(\mathbb{R})$, costituisce uno spazio vettoriale.

Osservazione 101. In particolare $\mathscr{P}_0(\mathbb{R}) = \{0\}$ (ossia contiene solo il polinomio nullo, di grado $-\infty$), mentre possiamo considerare $\mathscr{P}_1(\mathbb{R}) = \mathbb{R}$.

Esempio 2.1.7 (Polinomi a coefficienti complessi: $\mathscr{P}(\mathbb{C})$ e $\mathscr{P}_n(\mathbb{C})$)**.** Quanto detto per i polinomi coefficienti reali vale specularmente per polinomi a coefficienti complessi.

Esempio 2.1.8 (Spazio vettoriale nullo: $\{\mathbf{0}\}$)**.** Un insieme composto dal solo elemento $\mathbf{0}$ è uno spazio vettoriale in quanto sono definite la somma tra vettori $\mathbf{0}+\mathbf{0} = \mathbf{0}$ e il prodotto per scalare $\alpha\mathbf{0} = \mathbf{0}$ e sono rispettate le relative proprietà.

2.1.3 Sottospazi

Definizione 2.1.5 (Sottospazio (vettoriale))**.** Un sottoinsieme U dello spazio vettoriale V si dice sottospazio vettoriale se U è uno spazio vettoriale a sua volta, ossia:

(S_1) il vettore nullo $\mathbf{0} \in U$

(S_2) U è chiuso rispetto alla somma: $\mathbf{v} + \mathbf{w} \in U$ per ogni $\mathbf{v}, \mathbf{w} \in U$

(S_3) U è chiuso rispetto al prodotto per uno scalare: $t\mathbf{v} \in U$, per ogni $t \in \mathbb{C}, \mathbf{v} \in U$

Osservazione 102. L'esistenza dell'opposto di un elemento di un dato sottospazio è garantito dalla proprietà di prodotto per scalare, alla luce del fatto che $-\mathbf{v} = (-1)\mathbf{v}$.

Osservazione 103. Modo alternativo di verificare se un certo sottoinsieme U dello spazio vettoriale V è un sottospazio è dato dalla seguente.

Proposizione 2.1.8. *Sia V uno spazio vettoriale e sia U un sottoinsieme di V. Allora U è sottospazio di V se e solo se:*

(V_1) $U \neq \emptyset$

(V_2) si ha $\beta_1\mathbf{u}_1 + \beta\mathbf{u}_2 \in U$ per ogni $\mathbf{u}_1, \mathbf{u}_2 \in U$ e $\beta_1, \beta_2 \in \mathbb{C}$

Dimostrazione. Mostriamo innanzitutto che da (V_1) (V_2) derivano $(S_1), (S_2), (S_3)$. Dato che $U \neq \emptyset$ esiste almeno un elemento $\mathbf{u}_0 \in U$; allora possiamo applicare (V_2) con $\beta_1 = 1, \beta_2 = -1$ ottenendo che

$$1\mathbf{u}_0 + (-1)\mathbf{u}_0 = (1-1)\mathbf{u}_0 = \mathbf{0} \in U$$

e quindi vale la (S_1). La (S_2) si prova scegliendo $\beta_1 = \beta_2 = 1$. Infine la (S_3) prendendo $\beta_1 = \alpha$, $\beta_2 = 0$.
Mostriamo ora il viceversa: il fatto che $U \neq \emptyset$ è evidente (quanto meno per la presenza di $\mathbf{0}$); per (S_3) si ha che $\beta_1\mathbf{u}_1, \beta_2\mathbf{u}_2 \in U$, quindi per (S_2) anche $\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2$ (da cui è dimostrata la (V_2)) □

Esempio 2.1.9 (Insieme dei multipli scalari)**.** Sia V uno spazio vettoriale; definiamo l'insieme:

$$\langle \mathbf{v} \rangle = \{\alpha\mathbf{v} : \alpha \in \mathbb{C}, \mathbf{v} \in V\}$$

di tutti i *multipli scalari* di $\mathbf{v}$.
$\langle \mathbf{v} \rangle$ è uno sottospazio di V. Infatti $\mathbf{0} = 0\mathbf{v} \in \langle \mathbf{v} \rangle$. Inoltre se $\mathbf{v}_1, \mathbf{v}_2 \in \langle \mathbf{v} \rangle$, allora per qualche scalare α_1, α_2 deve essere $\mathbf{v}_1 = \alpha_1\mathbf{v}$ e $\mathbf{v}_2 = \alpha_2\mathbf{v}$; dunque

$$\mathbf{v}_1 + \mathbf{v}_2 = \alpha_1\mathbf{v} + \alpha_2\mathbf{v} = (\alpha_1 + \alpha_2)\mathbf{v} \in \langle \mathbf{v} \rangle$$

Infine se $\gamma \in \mathbb{C}$ e $\mathbf{v}_1 = \alpha_1\mathbf{v} \in \langle v \rangle$, allora

$$\gamma\mathbf{v}_1 = \gamma(\alpha_1\mathbf{v}) = (\gamma\alpha_1)\mathbf{v} \in \langle \mathbf{v} \rangle$$

Esempio 2.1.10 (Nucleo della matrice A)**.** Considerando una matrice $\mathbf{A} \in M_{m\times n}(\mathbb{C})$, definiamo il seguente sottoinsieme di $\mathbb{C}^n$:

$$\mathrm{N}(\mathbf{A}) = \{\mathbf{v} \in \mathbb{C}^n : \mathbf{A}\mathbf{v} = \mathbf{0}\}$$

Tale insieme è detto *nucleo della matrice* e corrisponde all'insieme delle soluzioni del sistema $\mathbf{A}\mathbf{x} = \mathbf{0}$ associato ad $\mathbf{A}$.
Il nucleo di matrice $m \times n$ è sottospazio vettoriale di $\mathbb{C}^n$. Infatti è ovvio che $\mathbf{0} \in \mathrm{N}(\mathbf{A})$ dato che $\mathbf{A}\mathbf{0} = \mathbf{0}$; inoltre dati due elementi $\mathbf{u}, \mathbf{v} \in \mathrm{N}(\mathbf{A})$, anche $\mathbf{u} + \mathbf{v} \in \mathrm{N}(\mathbf{A})$ dato che:

$$\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{A}\mathbf{u} + \mathbf{A}\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

e infine anche $\alpha\mathbf{v} \in \mathrm{N}(\mathbf{A})$ dato che:

$$\mathbf{A}(\alpha\mathbf{v}) = \alpha\mathbf{A}\mathbf{v} = \alpha\mathbf{0} = \mathbf{0}$$

Esempio 2.1.11. Per ogni $n \geq 0$, $\mathscr{P}_n(\mathbb{C})$ è un sottospazio vettoriale di $\mathscr{P}(\mathbb{C})$.

Esempio 2.1.12. Se W un sottospazio di U e U è un sottospazio di V, allora W è un sottospazio di V.

Esempio 2.1.13. Dato uno spazio vettoriale V, lo spazio nullo $\{\mathbf{0}\}$ e lo spazio V stesso sono sottospazi di V.

Esempio 2.1.14. Sia $V = M_2(\mathbb{C})$ lo spazio vettoriale delle matrici 2×2 complesse. Sia W l'insieme delle matrici 2×2 complesse simmetriche: si verifichi che W è un sottospazio vettoriale di V.
Si ha che $\mathbb{O} = \mathbb{O}^T$ pertanto $\mathbb{O} \in W$; inoltre date due matrici simmetriche $\mathbf{A} = \mathbf{A}^T$ e $\mathbf{B} = \mathbf{B}^T$ si ha che $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T = \mathbf{A} + \mathbf{B}$ pertanto $\mathbf{A} + \mathbf{B} \in W$. Infine se $\mathbf{A} = \mathbf{A}^T$, $(\alpha\mathbf{A})^T = \alpha\mathbf{A}^T = \alpha\mathbf{A}$ per cui $\alpha\mathbf{A} \in W$.

Esempio 2.1.15. Sia V lo spazio vettoriale delle funzioni $f : \mathbb{R} \to \mathbb{R}$. Sia $W = \{f \in V : f(1) = 0 \vee f(4) = 0\}$. Si dica se W è un sottospazio vettoriale di V.
Si ha che $\mathbf{0} \in W$ (dove $\mathbf{0}$ è la funzione costantemente nulla). Tuttavia basta un controesempio per mostrare che $f + g \notin W$ e pertanto W non p un sottospazio vettoriale; ad esempio basta scegliere una funzione f tale che $f(1) = 0, f(4) = \alpha \neq 0$ e una g tale che $g(1) = \alpha, g(4) = 0$ e si ottiene che $f + g(1) = \alpha \neq 0$ e $f + g(4) = \alpha \neq 0$.

Definizione 2.1.6 (Sottospazi propri)**.** Un sottospazio U di V si dice proprio se $U \neq V$ (ossia esiste un elemento $\mathbf{v}$ tale che $\mathbf{v} \in V$ ma $\mathbf{v} \notin U$).

Definizione 2.1.7 (Intersezione e somma di sottospazi)**.** Supponiamo che U_1 e U_2 siano sottospazi di V. Possiamo definire altri due sottospazi:

- $U_1 \cap U_2$; è il più grande sottospazio di V contenuto in U_1 e U_2
- $U_1 + U_2 = \{\mathbf{u}_1 + \mathbf{u}_2 : \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2\}$; è il più piccolo sottospazio di V che contiene sia U_1 che U_2

Definizione 2.1.8 (Somma diretta di sottospazi)**.** Se $U_1 \cap U_2 = \{\mathbf{0}\}$, scriveremo $U_1 + U_2 = U_1 \oplus U_2$ e si parla di *somma diretta di sottospazi.*

2.1.4 Insieme finito di vettori

Definizione 2.1.9 (Insieme finito di vettori)**.** Un insieme finito di vettori dello spazio vettoriale V

$$\mathscr{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n\} \tag{2.10}$$

è una lista finita di elementi di V che ammette che il singolo elemento sia ripetuto (pertanto non si tratta di un insieme vero e proprio).
Nel caso di $n = 0$ parliamo di insieme di vettori vuoto, indicato con $\emptyset$.

Osservazione 104 (Sottoinsieme finito di vettori)*.* Quando parleremo di sottoinsieme di un insieme finito di vettori si intenderà che se un certo vettore compare k volte nel sottoinsieme, esso dovrà comparire *almeno* k volte nell'insieme originario.

Osservazione 105. Quando nel seguito si fa riferimento ad insieme di vettori, indicati con $\mathscr{A}, \mathscr{B}, \mathscr{C}$, salvo indicazione contraria si sottintende insieme finito.

Esempio 2.1.16. Una matrice $\mathbf{A} \in M_{m\times n}(\mathbb{C})$ definisce un insieme di vettori in $\mathbb{C}^m$; scrivendo $\mathbf{A} = [\mathbf{a}_1\, \mathbf{a}_2 \dots \mathbf{a}_n]$ associamo ad $\mathbf{A}$ l'insieme di n vettori colonna che la compongono.

Definizione 2.1.10 (Unione di insiemi)**.** Possiamo unire due insiemi $\mathscr{A}, \mathscr{B}$ ponendo il secondo di seguito al primo e mantenendo le ripetizioni degli elementi. Indichiamo l'insieme così ottenuto con $\mathscr{A} \sqcup \mathscr{B}$.

2.1.5 Combinazioni lineari

Definizione 2.1.11 (Combinazione lineare di $\mathscr{A}$)**.** Dato un insieme di vettori $\mathscr{A} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ dello spazio vettoriale V e delle costanti $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{C}$, si chiama combinazione lineare di $\mathscr{A}$ l'elemento

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \dots + \alpha_n\mathbf{v}_n \in V \tag{2.11}$$

Osservazione 106. Una combinazione lineare si ottiene sfruttando somma e prodotto per scalare su elementi dello spazio vettoriale pertanto, per definizione, il risultato appartiene al medesimo spazio (essendo l'insieme chiuso rispetto a somma e prodotto per scalare).

Definizione 2.1.12 (Sottospazio generato da $\mathscr{A}$)**.** Dato un insieme di vettori $\mathscr{A} = \{\mathbf{v}_1; \mathbf{v}_2; \dots, \mathbf{v}_n\}$ dello spazio vettoriale V, si definisce sottospazio generato da $\mathscr{A}$ l'insieme di tutte le combinazioni lineari dei vettori dell'insieme, indicato con $\langle \mathscr{A} \rangle$ o con $\langle \mathbf{v}_1; \mathbf{v}_2; \dots; \mathbf{v}_n \rangle$

Osservazione 107. Per convenzione, nel caso dell'insieme vuoto si pone

$$\langle \emptyset \rangle = \{\mathbf{0}\} \tag{2.12}$$

Osservazione 108. Il sottospazio generato dall'insieme $\{\mathbf{u}; \mathbf{v}\}$ è lo stesso del sottospazio generato da $\{\mathbf{v}; \mathbf{u}\}$ a causa del fatto che l'addizione di vettori è commutativa.

Proposizione 2.1.9. *Dato l'insieme di vettori $\mathscr{A}$ dello spazio vettoriale V, $\langle \mathscr{A} \rangle$ è effettivamente un sottospazio di V.*

Dimostrazione. Il fatto è evidente, per definizione se $\mathscr{A} = \emptyset$ (poiché $\langle\emptyset\rangle = \{\mathbf{0}\}$, lo spazio nullo, è sottospazio di V); supponiamo allora che $\mathscr{A} = \{\mathbf{v}_1; \mathbf{v}_2; \ldots; \mathbf{v}_n\}$ e verifichiamo che $\langle\mathscr{A}\rangle$ è sottospazio (lo si fa seguendo il metodo di proposizione 2.1.8).
$\mathbf{0} \in \langle\mathscr{A}\rangle$ (quindi $\langle\mathscr{A}\rangle \neq \emptyset$), in quanto possiamo scrivere $\mathbf{0}$ come combinazione dei vettori di $\mathscr{A}$ con coefficienti tutti nulli:

$$0\mathbf{v}_1 + \ldots + 0\mathbf{v}_n = \mathbf{0}$$

Se $\mathbf{u}_1, \mathbf{u}_2 \in \langle\mathscr{A}\rangle$, per opportuni α_i e α_i' deve essere

$$\mathbf{u}_1 = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i, \quad \mathbf{u}_2 = \sum_{i=1}^{n} \alpha_i' \mathbf{v}_i,$$

e per $\beta_1, \beta_2 \in \mathbb{C}$

$$\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 = \beta_1 \sum_{i=1}^{n} \alpha_i \mathbf{v}_i + \beta_2 \sum_{i=1}^{n} \alpha_i' \mathbf{v}_i = \sum_{i=1}^{n} (\beta_1\alpha_i + \beta_2\alpha_i')\mathbf{v}_i \in \langle\mathscr{A}\rangle$$

e la verifica è completa. □

Esempio 2.1.17 (Combinazione lineare e sistema lineare)**.** Considerando l'insieme di vettori in $\mathbb{C}^3$

$$\mathscr{A} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}; \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}; \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}; \right\}$$

una combinazione lineare dei vettori di $\mathscr{A}$ sarà un vettore della forma

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \begin{bmatrix} \alpha_1 + 2\alpha_2 - \alpha_3 \\ -\alpha_1 + \alpha_2 + 4\alpha_3 \\ 0 \end{bmatrix}$$

Il concetto di combinazione lineare e di sistema lineare sono strettamente legati; infatti, ogni vettore della forma:

$$\mathbf{b} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ 0 \end{bmatrix}$$

può esser scritto come combinazione lineare dei vettori di $\mathscr{A}$: basta impostare il sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ con:

$$\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3] = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & 4 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

e risolvere per $\mathbf{x}$ (ossia determinare i coefficienti $\alpha_1, \alpha_2, \alpha_3$ per i quali vanno moltiplicati $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ al fine di ottenere $\mathbf{b}$).

Proposizione 2.1.10. *Se $\mathscr{A} \subset \mathscr{B}$, allora $\langle\mathscr{A}\rangle$ è un sottospazio di $\langle\mathscr{B}\rangle$.*

Dimostrazione. Ancora se $\mathscr{A} = \emptyset$ la cosa è ovvia. Supponiamo allora che $\mathscr{A}$ abbia $m > 0$ elementi, mentre $\mathscr{B}$ $n > m$. Possiamo scrivere:

$$\mathscr{A} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$$
$$\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\}$$

Ci basta verificare che ogni vettore che sia combinazione lineare dei vettori di $\mathscr{A}$ è anche combinazione lineare dei vettori di $\mathscr{B}$. È ciò semplicemente si fa impostando a 0 i coefficienti dei vettori da $\mathbf{v}_{m+1}$ in poi, ossia:

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_m\mathbf{v}_m = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_m\mathbf{v}_m + 0\mathbf{v}_{m+1} + \ldots + 0\mathbf{v}_n$$

□

Osservazione 109. La proposizione 2.1.10 (e soprattutto l'idea alla base della sua dimostrazione) giustifica il fatto che si ponga $\langle\emptyset\rangle = \{\mathbf{0}\}$ (ossia dato che l'insieme di vettori vuoto è sottoinsieme di qualsiasi altro insieme di vettori $\mathscr{A}$, è come se facesse la combinazione lineare con coefficienti nulli dei vettori di $\mathscr{A}$).

2.2 Insiemi di generatori e indipendenza lineare

2.2.1 Insiemi di generatori

Esempio 2.2.1. Considerando lo spazio vettoriale $\mathbb{C}^n$ e l'insieme finito di vettori ottenuti dalla matrice identità $\mathbf{I}_n = [\mathbf{e}_1\ \mathbf{e}_2\ \ldots\ \mathbf{e}_n]$, l'insieme $\mathscr{E} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ è caratterizzato dal fatto che se prendiamo un qualsiasi $\mathbf{v} = [\alpha_1\, \alpha_2\, \ldots\, \alpha_n]^T \in \mathbb{C}^n$ si può facilmente verificare che

$$\mathbf{v} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_n\mathbf{e}_n$$

e quindi in generale l'intero spazio vettoriale coincide con il sottospazio generato da $\mathscr{E}$:

$$\mathbb{C}^n = \langle\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\rangle$$

dal momento che ogni vettore di $\mathbb{C}^n$ si può scrivere come combinazione lineare dei vettori dell'insieme $\mathscr{E}$. Generalizziamo ora questa idea.

Definizione 2.2.1 (Insieme di generatori di V)**.** Un insieme finito di vettori $\mathscr{A}$ nello spazio vettoriale V si dice insieme di generatori di V se $V = \langle\mathscr{A}\rangle$, cioè ogni vettore di V si può scrivere come combinazione lineare dei vettori dell'insieme $\mathscr{A}$.

Esempio 2.2.2. Lo spazio vettoriale nullo $\{\mathbf{0}\}$ ammette come insieme di generatori sia $\emptyset$ che l'insieme $\{\mathbf{0}\}$.

Esempio 2.2.3 (Legame tra insiemi di generatori e sistemi lineari)**.** Considerando i seguenti vettori di $\mathbb{C}^2$

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

mostriamo che $\mathscr{A} = \{\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3\}$ è un insieme di generatori di $\mathbb{C}^2$. Dato un qualsiasi vettore $\mathbf{v} = [r\ s]^T$ dobbiamo poterlo scrivere come combinazione lineare dei tre vettori dati ossia trovare gli $\alpha_1, \alpha_2, \alpha_3$ tali che

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3$$

ossia

$$\begin{cases} 2\alpha_1 - \alpha_2 + \alpha_3 = r \\ \alpha_1 + 3\alpha_2 - 2\alpha_3 = s \end{cases}$$

e si verifica facilmente che tale sistema (indeterminato) ha soluzioni (multiple) per qualsiasi coppia $r, s \in \mathbb{C}$.
Analogamente si può osservare come anche $\{\mathbf{v}_1, \mathbf{v}_2\}$ sia un insieme di generatori di $\mathbb{C}^2$ (sistema determinato, una soluzione unica per ogni coppia r, s), mentre nessun insieme con un solo elemento può costituirlo.

Esempio 2.2.4. Si dica se $S = \left\{ \begin{bmatrix} 1 \\ i \\ 1 \end{bmatrix}, \begin{bmatrix} 1+i \\ i \\ 1-i \end{bmatrix}, \begin{bmatrix} i \\ 0 \\ 2 \end{bmatrix} \right\}$ è un insieme di generatori di $\mathbb{C}^3$.
S è un insieme di generatori se e solo se ha almeno una soluzione il sistema $\mathbf{Ax} = \mathbf{d}$ con

$$\mathbf{A} = \begin{bmatrix} 1 & 1+i & i \\ i & i & 0 \\ 1 & 1-i & 2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \ \forall a, b, c \in \mathbb{C}$$

Si ha

$$\begin{bmatrix} 1 & 1+i & i & a \\ i & i & 0 & b \\ 1 & 1-i & 2 & c \end{bmatrix} \qquad r_3 = r_3 - r_1$$

$$\begin{bmatrix} 1 & 1+i & i & a \\ i & i & 0 & b \\ 0 & 0 & 2-i & c-a \end{bmatrix} \qquad r_2 = r_2 \cdot i$$

$$\begin{bmatrix} 1 & 1+i & i & a \\ -1 & -1 & 0 & bi \\ 0 & 0 & 2-i & c-a \end{bmatrix} \qquad r_1 = r_1 + r_2$$

$$\begin{bmatrix} 0 & i & i & a+bi \\ -1 & -1 & 0 & bi \\ 0 & 0 & 2-i & c-a \end{bmatrix} \qquad \text{e riordino}$$

$$\begin{bmatrix} -1 & -1 & 0 & bi \\ 0 & i & i & a+bi \\ 0 & 0 & 2-i & c-a \end{bmatrix}$$

Il sistema è determinato e quindi si tratta di un insieme di generatori.

Esempio 2.2.5. Sia $V = M_2(\mathbb{C})$ lo spazio vettoriale delle matrici 2×2 complesse. Sia W l'insieme delle matrici 2×2 complesse simmetriche: si provi che

$$S = \left\{ \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \right\}$$

è un insieme di generatori per W.
S è un insieme di generatori se $\forall \alpha_1, \ldots, \alpha_5$

$$\alpha_1 \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 & 3 \\ 3 & 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + \alpha_4 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} + \alpha_5 \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \in W$$

e ciò accade se e solo se sono uguali gli elementi di posto (2,1) e quello di posto (1,2). Come si può vedere entrambi sono uguali a $\alpha_1 2 + \alpha_2 3 + \alpha_3 1 + \alpha_4 0 + \alpha_5 1$ per cui S è un insieme di generatori

Definizione 2.2.2 (Spazio vettoriale finitamente generato)**.** Uno spazio vettoriale V che ammette un insieme finito di generatori.

Osservazione 110. Nel seguito salvo eccezioni si considerano solo spazi finitamente generati.

Proposizione 2.2.1. *Lo spazio vettoriale delle matrici $M_{m\times n}(\mathbb{C})$ è finitamente generato.*

Dimostrazione. Data la definizione di prodotto di matrice per scalare è immediato verificare che le mn matrici della base canonica possono essere utilizzate in una combinazione lineare. Ad esempio per le matrici 2×2:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

□

Proposizione 2.2.2. *Se $\mathscr{A}$ e $\mathscr{B}$ sono insiemi di vettori nello spazio vettoriale V tali che $\mathscr{A} \subseteq \mathscr{B}$ e e $\mathscr{A}$ è un insieme di generatori di V. Allora anche $\mathscr{B}$ è un insieme di generatori.*

Dimostrazione. Dalle ipotesi segue che $V = \langle \mathscr{A} \rangle \subseteq \langle \mathscr{B} \rangle \subseteq V$ e quindi $V = \langle \mathscr{B} \rangle$. $\langle \mathscr{A} \rangle \subseteq \langle \mathscr{B} \rangle$ deriva dalla 2.1.10, mentre $\langle \mathscr{B} \rangle \subseteq V$ dal fatto che una combinazione lineare di vettori di V (ossia quelli di $\mathscr{B}$) appartiene ancora a V. □

Osservazione 111. Ogni volta che un vettore di un insieme è combinazione lineare degli altri, lo possiamo togliere dall'insieme senza cambiare lo spazio generato, come mostra il seguente risultato.

Proposizione 2.2.3. *Se $\{\mathbf{v}_1; \mathbf{v}_2; \ldots; \mathbf{v}_n\}$ è un insieme di generatori dello spazio vettoriale V e $\mathbf{v}_n \in \langle \mathbf{v}_1; \mathbf{v}_2; \ldots; \mathbf{v}_{n-1} \rangle$, allora anche $\{\mathbf{v}_1; \mathbf{v}_2; \ldots; \mathbf{v}_{n-1}\}$ è un insieme di generatori di V.*

Dimostrazione. Sia $\mathbf{v} \in V$ un generico vettore; possiamo scriverlo come combinazione lineare degli n generatori:

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

per opportuni coefficienti α_i. Se l'ultimo generatore è però una combinazione dei $n - 1$ precedenti si può scrivere:

$$\mathbf{v}_n = \sum_{i=1}^{n-1} \beta_i \mathbf{v}_i$$

sempre per opportuni coefficienti β_i. Basta allora sostituire quest'ultima nella prima uguaglianza, ottenendo:

$$\mathbf{v} = \sum_{i=1}^{n-1} \alpha_i \mathbf{v}_i + \alpha_n \mathbf{v}_n = \sum_{i=1}^{n-1} \alpha_i \mathbf{v}_i + \alpha_n \sum_{i=1}^{n-1} \beta_i \mathbf{v}_i = \sum_{i=1}^{n-1} (\alpha_i + \alpha_n \beta_i) \mathbf{v}_i$$

per mostrare, come richiesto, che un generico $\mathbf{v} \in V$ può essere generato anche facendo a meno di $\mathbf{v}_n$. $\square$

Esempio 2.2.6. Considerando la matrice in forma ridotta

$$\mathbf{U} = [\mathbf{u}_1\, \mathbf{u}_2\, \mathbf{u}_3\, \mathbf{u}_4\, \mathbf{u}_5] = \begin{bmatrix} 1 & -2 & 2 & 1 & 3 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Possiamo scrivere $\mathbf{u}_4 = 3\mathbf{u}_1 + 0\mathbf{u}_2 - \mathbf{u}_3 + 0\mathbf{u}_5$; nel caso non ce ne fossimo accorti, per verificare che tale combinazione esiste è valida, basta risolvere il sistema $\mathbf{Ax} = \mathbf{u}_4$, con $\mathbf{A} = [\mathbf{u}_1\, \mathbf{u}_2\, \mathbf{u}_3\, \mathbf{u}_5]$ e notare come $\mathbf{x} = [3\,0\,-1\,0]^T$ sia una soluzione. Nello specifico la matrice aumentata del sistema $\mathbf{Ax} = \mathbf{u}_4$

$$\begin{bmatrix} 1 & -2 & 2 & 3 & 1 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{cases} x_5 = 0 \\ x_3 = -1 \\ x_2 = h \\ x_1 - 2h + 2(-1) + 3 \cdot 0 = 1 \end{cases} \qquad \begin{cases} x_5 = 0 \\ x_3 = -1 \\ x_2 = h \\ x_1 = 2h + 3 \end{cases}$$

da cui se $h = 0$ si ottiene appunto $\mathbf{x} = [3\,0\,-1\,0]^T$.
Se un vettore $\mathbf{v} \in \mathbb{C}^3$ è combinazione lineare di $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$, per esempio $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4\mathbf{u}_4 + \alpha_5\mathbf{u}_5$ abbiamo anche che è combinazione di $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_5\}$ essendo $\mathbf{u}_4$ combinazione lineare dei rimanenti:

$$\begin{aligned} \mathbf{v} &= \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \alpha_4(3\mathbf{u}_1 + 0\mathbf{u}_2 - \mathbf{u}_3 + 0\mathbf{u}_5) + \alpha_5\mathbf{u}_5 \\ &= (\alpha_1 + 3\alpha_4)\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + (\alpha_3 - \alpha_4)\mathbf{u}_3 + \alpha_5\mathbf{u}_5 \end{aligned}$$

Abbiamo quindi che

$$V = \langle \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5 \rangle = \langle \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_5 \rangle$$

Verifichiamo anche che $V = \langle \mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_5 \rangle$, ossia $\mathbf{Ax} = \mathbf{u}_2$, con $A = [\mathbf{u}_1\, \mathbf{u}_3\, \mathbf{u}_5]$:

$$\begin{bmatrix} 1 & 2 & 3 & -2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{cases} x_5 = 0 \\ x_3 = 0 \\ x_1 + 0 \cdot 2 + 0 \cdot 3 = -2 \end{cases} \qquad \begin{cases} x_5 = 0 \\ x_3 = 0 \\ x_1 = -2 \end{cases}$$

e quindi $\mathbf{u}_2$ è una combinazione lineare di $\mathbf{u}_1, \mathbf{u}_3, \mathbf{u}_5$.
Infine, come si nota da $\mathbf{A}$ di quest'ultimo caso non è più possibile eliminare altri vettori dall'insieme senza cambiare il sottospazio generato: per esempio $\mathbf{u}_5$ non è combinazione lineare di $\mathbf{u}_1$ e $\mathbf{u}_3$).

2.2.2 Indipendenza lineare

Definizione 2.2.3 (Insieme di vettori linearmente dipendente)**.** Un insieme di vettori $\mathscr{A} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ si dice linearmente dipendente se esistono scalari $\alpha_1, \alpha_2, \dots, \alpha_n$ *non tutti nulli* (basta anche uno solo $\neq 0$) tali che

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots + \alpha_n\mathbf{v}_n = \mathbf{0} \tag{2.13}$$

Osservazione 112. Per precisione di linguaggio, la dipendenza lineare non è una proprietà dei singoli vettori ma dell'insieme che essi formano.

Osservazione 113. Ogni insieme in cui compaia il vettore nullo è linearmente dipendente in quanto se $\mathbf{v}_1 = \mathbf{0}$ possiamo prendere $\alpha_1 = 1$ e $\alpha_i = 0$ per $i = 2, \ldots, n$, ottenendo $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$.

Proposizione 2.2.4. *L'insieme costituito da un solo vettore* $\mathbf{v}$ *è linearmente dipendente se e solo se* $\mathbf{v} = \mathbf{0}$

Dimostrazione. Supponiamo che $\{\mathbf{v}\}$ sia linearmente dipendente: esiste uno scalare non nullo t tale che $t\mathbf{v} = \mathbf{0}$. Allora per la legge di annullamento del prodotto per uno scalare $\mathbf{v}$ è il vettore nullo. Viceversa il vettore nullo costituisce un insieme linearmente dipendente perché $1\mathbf{0} = \mathbf{0}$ e lo scalare 1 è non nullo. □

Esempio 2.2.7. Se nell'insieme $\mathscr{A}$ ci sono due vettori uguali, esso è linearmente dipendente (basta scegliere coefficienti opposti per tali vettori e nulli per i restanti)

Esempio 2.2.8. Considerando ancora l'insieme di vettori in $\mathbb{C}^3$

$$\mathscr{A} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}; \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}; \mathbf{v}_3 = \begin{bmatrix} -1 \\ 4 \\ 0 \end{bmatrix}; \right\}$$

una combinazione lineare dei vettori di $\mathscr{A}$ sarà un vettore della forma

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \begin{bmatrix} \alpha_1 + 2\alpha_2 - \alpha_3 \\ -\alpha_1 + \alpha_2 + 4\alpha_3 \\ 0 \end{bmatrix}$$

ed è immediato verificare che, per $\alpha_1 = 3, \alpha_2 = -1, \alpha_3 = 1$ la combinazione lineare corrispondente è il vettore nullo. Pertanto $\mathscr{A}$ è linearmente dipendente.

Proposizione 2.2.5. *Se* $\mathscr{A}$ *è un sottoinsieme non vuoto dell'insieme di vettori* $\mathscr{B}$*, e* $\mathscr{A}$ *è linearmente dipendente, allora anche* $\mathscr{B}$ *è linearmente dipendente.*

Dimostrazione. Supponendo che

$$\begin{aligned} \mathscr{A} &= \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \\ \mathscr{B} &= \{\mathbf{v}_1, \ldots, \mathbf{v}_m, \mathbf{v}_{m+1}, \ldots, \mathbf{v}_n\} \end{aligned}$$

e che

$$\sum_{i=1}^{m} \alpha_i \mathbf{v}_i = \mathbf{0}$$

per α_i per $i = 1, \ldots, m$ non tutti nulli. Allora ponendo $\alpha_{m+1} = \alpha_n = 0$ e abbiamo

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \sum_{i=1}^{m} \alpha_i \mathbf{v}_i + \sum_{i=m+1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

da cui la tesi. □

Definizione 2.2.4 (Insieme linearmente indipendente)**.** Un insieme di vettori $\mathscr{A} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ è *linearmente indipendente* se non è linearmente dipendente, ovvero si ha

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots + \alpha_n \mathbf{v}_n = \mathbf{0}$$

se e solo $\alpha_1 = \ldots = \alpha_n = 0$

Osservazione 114. Per convenzione, anche l'insieme vuoto è linearmente indipendente.

Esempio 2.2.9. L'insieme $\mathscr{E}$ delle colonne della matrice identità $\mathbf{I}_n$ è un insieme indipendente di vettori di $\mathbb{C}^n$, infatti si ha

$$\alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_n\mathbf{e}_n = \mathbf{0}$$

se e solo se $\alpha_1 = \ldots = \alpha_n = 0$, come si può facilmente verificare (ad esempio per $n = 4$).

Esempio 2.2.10 (Verifica indipendenza vettori)**.** Verifichiamo che l'insieme di vettori in $\mathbb{C}^4$

$$\mathscr{A} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} ; \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 0 \end{bmatrix} ; \mathbf{v}_3 = \begin{bmatrix} 0 \\ 3 \\ 2 \\ 0 \end{bmatrix} \right\}$$

è linearmente indipendente. Siamo interessati ai valori $\alpha_1, \alpha_2, \alpha_3$ per i quali

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

il che equivale a impostare il sistema lineare

$$\begin{cases} \alpha_1 + 2\alpha_2 = 0 \\ -\alpha_1 + \alpha_2 + 3\alpha_3 = 0 \\ 2\alpha_1 + 2\alpha_2 + 2\alpha_3 = 0 \end{cases}$$

(si è tralasciata l'ultima equazione perché identità) ossia $\mathbf{A}\mathbf{x} = \mathbf{0}$ con

$$\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{v}_3] = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 3 \\ 2 & 2 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Il sistema dato ammette solo la soluzione $\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$ per $\mathscr{A}$ è linearmente indipendente.

Osservazione 115. Come si vedrà un insieme di vettori è indipendente se il nucleo della matrice che formano contiene solamente il vettore nullo.

Esempio 2.2.11. L'insieme

$$\mathscr{B} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} \right\}$$

è linearmente dipendente. Ponendo

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{0}$$

che si può riscrivere come sistema $\mathbf{A}\mathbf{x} = \mathbf{0}$ con

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 2 & -2 \\ 3 & 1 & -5 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}, \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Risolvendo il sistema si trova che una soluzione possibile è $[2, -1, 1]^T$ che corrisponde alla relazione

$$2\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3 = \mathbf{0}$$

Motivo per cui l'insieme $\mathscr{B}$ è linearmente dipendente.

Esempio 2.2.12. Si dica quale dei seguenti insiemi di $\mathbb{R}^3$ è linearmente indipendente:

$$S_1 = \left\{ \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \right\}, S_2 = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 0 \end{bmatrix} \right\}$$

Sono indipendenti se l'unica soluzione di $\mathbf{Ax} = \mathbf{0}$ (sistemi determinati) è $\mathbf{0}$. Partendo dalla matrice aumentata costruita per S_1:

$$\begin{bmatrix} 4 & 4 & 2 & 0 \\ 2 & 6 & -1 & 0 \\ 2 & 2 & 1 & 0 \end{bmatrix} \qquad r_3 = r_3 \cdot 2$$

$$\begin{bmatrix} 4 & 4 & 2 & 0 \\ 2 & 6 & -1 & 0 \\ 4 & 4 & 2 & 0 \end{bmatrix} \qquad r_3 = r_3 - r_1$$

$$\begin{bmatrix} 4 & 4 & 2 & 0 \\ 2 & 6 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad r_2 = 2 \cdot r_2 - r_1$$

$$\begin{bmatrix} 4 & 4 & 2 & 0 \\ 0 & 8 & -4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Quindi questo primo sistema è indeterminato ed S_1 è non linearmente indipendente. Per S_2 invece

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 3 & 0 \\ 1 & 2 & 0 & 0 \end{bmatrix} \qquad r_2 = r_2 - r_1,\ r_3 = r_3 - r_1$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad r_3 = r_3 - r_2$$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & -3 & 0 \end{bmatrix}$$

Da cui

$$\begin{cases} -3x_3 = 0 \\ x_2 + 3 \cdot 0 = 0 \\ x_1 + 0 \cdot 1 = 0 \end{cases} \quad \begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{cases}$$

Il sistema è determinato con $\mathbf{0}$ come soluzione quindi è S_2 ad essere linearmente indipendente.

Esempio 2.2.13. Sia V uno spazio vettoriale complesso ed $A = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ un sottoinsieme linearmente indipendente di vettori di V. Si dica quale dei seguenti insiemi di vettori di V è linearmente indipendente:

$$A_1 = \{\mathbf{v}_2 + \mathbf{v}_3; \mathbf{v}_1 + \mathbf{v}_3; \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3\}$$
$$A_2 = \{\mathbf{v}_1 - 2\mathbf{v}_3; \mathbf{v}_1 + \mathbf{v}_2; \mathbf{v}_2 + 2\mathbf{v}_3\}$$

Tentativo di soluzione: partendo dal primo, i tre vettori sono linearmente indipendenti se $\mathbf{A}\mathbf{x} = \mathbf{0}$ con

$$\mathbf{A} = \begin{bmatrix} \mathbf{v}_2 + \mathbf{v}_3 & \mathbf{v}_1 + \mathbf{v}_3 & \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

o, guardando al prodotto di blocchi colonna per riga, equivalentemente se

$$(\mathbf{v}_2 + \mathbf{v}_3)x_1 + (\mathbf{v}_1 + \mathbf{v}_3)x_2 + (\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3)x_3 = \mathbf{0}$$

che sviluppata porta a

$$\mathbf{v}_1(x_2 + x_3) + \mathbf{v}_2(x_1 + x_3) + \mathbf{v}_3(x_1 + x_2 + x_3) = \mathbf{0}$$

Dato che $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ è linearmente indipendente l'uguaglianza può essere verificata se e solo se

$$\begin{cases} x_2 + x_3 = 0 \\ x_1 + x_3 = 0 \\ x_1 + x_2 + x_3 = 0 \end{cases} \qquad \begin{cases} x_1 + 0 = 0 \\ x_2 = x_3 \\ 0 + x_3 = 0 \end{cases} \qquad \begin{cases} x_1 = 0 \\ x_2 = 0 \\ x_3 = 0 \end{cases} \to \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Effettuando gli stessi passi con A_2 si giunge a

$$\mathbf{v}_1(x_1 + x_2) + \mathbf{v}_2(x_2 + x_3) + \mathbf{v}_3(2x_3 - 2x_1) = 0$$

che conduce al sistema

$$\begin{bmatrix} 1 & 0 & -2 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Il sistema è indeterminato, quindi A_2 non è linearmente indipendente.

Osservazione 116. Si vedrà che un insieme di vettori è linearmente dipendente se e solo se il nucleo della matrice $\mathbf{A}$ che formano contiene anche un solo vettore $\mathbf{v} \neq \mathbf{0}$.

Proposizione 2.2.6. *Se $\mathscr{A}$ è un insieme linearmente indipendente, ogni suo sottoinsieme è ancora linearmente indipendente.*

Dimostrazione. Se un sottoinsieme di $\mathscr{A}$ fosse linearmente dipendente, anche $\mathscr{A}$ lo sarebbe per la proposizione 2.2.5. □

Proposizione 2.2.7. *Se $\mathscr{A}$ è un insieme linearmente indipendente, esso consiste di vettori a due a due distinti e non nulli.*

Dimostrazione. Infatti ve ne fossero anche solo due uguali il sistema non sarebbe più indipendente (basterebbe porli in combinazione con coefficienti opposti); e se anche un solo vettore fosse nullo nella combinazione gli si potrebbe dare qualsiasi coefficiente diverso da 0 facendo si che l'insieme cui appartiene sia linearmente dipendente. □

Proposizione 2.2.8. *Un insieme di vettori $\mathscr{A}$ è linearmente indipendente se e solo se nessuno dei vettori di $\mathscr{A}$ è combinazione lineare degli altri.*

Dimostrazione. Supponiamo che $\mathscr{A}$ sia linearmente indipendente e dimostriamo che nessuno dei vettori di $\mathscr{A}$ è combinazione lineare degli altri. Se è l'insieme vuoto evidentemente non ci sono vettori che possono essere combinazioni lineari di altri (poiché non ce ne sono). Se invece $\mathscr{A} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, supponiamo per assurdo che $\mathbf{v}_n$ sia combinazione degli altri, ossia:

$$\mathbf{v}_n = \sum_{i=1}^{n-1} \alpha_i \mathbf{v}_i$$

da cui, se poniamo $\alpha_n = -1$ si ottiene (portando $\mathbf{v}_n$ a destra):

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

ma essendo $\mathscr{A}$ linearmente indipendente si ha che l'uguaglianza di sopra è garantita solo se $\alpha_n = 0$, il che è assurdo perché avevamo ipotizzato $\alpha_n = -1$.
Viceversa nell'ipotesi che nessuno dei vettori di $\mathscr{A}$ sia combinazione lineare degli altri dimostriamo che $\mathscr{A}$ è linearmente indipendente. Se $\alpha_i \in \mathbb{C}$, $(i = 1, \ldots, n)$ e:

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

con $\alpha_n \neq 0$ di fatto si ha che

$$\mathbf{v}_n = \sum_{i=1}^{n-1} \beta_i \mathbf{v}_i$$

con $\beta_i = -\alpha_n^{-1}\alpha_i$. Ma questo non è possibile per ipotesi (ossia che un vettore sia combinazione lineare degli altri), e dunque deve essere $\alpha_n = 0$; allo stesso modo si dimostra che $\alpha_i = 0$ per $i = 1, \ldots, n-1$. □

Teorema 2.2.9. *Sia $\mathscr{A}$ un insieme linearmente indipendente nello spazio vettoriale V e sia $\mathbf{v} \in V$ tale che $\mathbf{v} \notin \langle \mathscr{A} \rangle$. Allora l'insieme $\mathscr{B} = \mathscr{A} \sqcup \{\mathbf{v}\}$ è linearmente indipendente.*

Dimostrazione. La cosa è ovvia se $\mathscr{A}$ è l'insieme vuoto: si ha che $\mathscr{B} = \{\mathbf{v}\}$, ma $\mathbf{v} \neq \mathbf{0}$ dato che $\mathbf{v} \notin \langle \emptyset \rangle$, ed un insieme di un solo elemento è linearmente indipendente se e solo l'elemento non è l'elemento nullo.
Supponiamo allora $\mathscr{A} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$; si ha indipendenza se nell'ipotesi che

$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n + \alpha \mathbf{v} = \mathbf{0}$$

si ha che $\alpha_1 = \ldots = \alpha_n = \alpha = 0$. Dobbiamo mostrare questo. Per ipotesi $\mathscr{A}$ è linearmente indipendente quindi $\alpha_1 = \ldots = \alpha_n = 0$. Venendo al coefficiente del vettore aggiuntivo, supponiamo per assurdo che sia $\neq 0$; si ha che

$$-\alpha \mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$

e ponendo $\beta_i = -\alpha^{-1}\alpha_i$, otteniamo che $\mathbf{v}$ si può riscrivere come combinazione lineare

$$\mathbf{v} = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$$

da cui seguirebbe $\mathbf{v} \in \langle \mathscr{A} \rangle$, contraddizione. Dunque $\alpha = 0$ e quindi come si voleva tutti gli scalari sono nulli. □

Proposizione 2.2.10. *Sia* $\mathbf{U}$ *una matrice* $n \times p$ *in forma ridotta e siano* $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ *le sue colonne dominanti. Allora l'insieme di vettori* $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_k\}$ *in* $\mathbb{C}^n$ *è linearmente indipendente*

Dimostrazione. Le colonne dominanti dopo la EG assumono la forma

$$\mathbf{v}_i = \begin{bmatrix} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow i \\ \\ \\ \\ \end{matrix}$$

con gli asterischi ad indicare coefficienti non nulli. Procediamo per induzione. Se c'è una sola colonna dominante, l'insieme è formato da un solo vettore non nullo e quindi è linearmente indipendente. Supponiamo allora che $k > 1$ e battezziamo come $\mathbf{U}'$ la matrice che ha come colonne quelle precedenti l'ultima colonna dominante di $\mathbf{U}$; si ha che $\mathbf{U}'$ è ancora in forma ridotta e ha esattamente come colonne dominanti le prime $k-1$ colonne dominanti di $\mathbf{U}$, ossia $\{\mathbf{v}_1, \ldots, \mathbf{v}_{k-1}\}$: supponiamo per ipotesi che l'insieme di queste colonne sia linearmente indipendente e vogliamo verificare che lo sia anche $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Ci basta allora verificare che $\mathbf{v}_k \notin \langle \mathbf{v}_1, \ldots, \mathbf{v}_{k-1} \rangle$: ma una combinazione lineare delle prime $k-1$ colonne dominanti ha tutti i coefficienti della k-esima riga in poi nulli, e quindi non può essere uguale a $\mathbf{v}_k$ che sulla k-esima riga ha 1. □

Proposizione 2.2.11. *In una matrice complessa* $\mathbf{U}$ *di dimensioni* $m \times n$ *in forma ridotta, l'insieme delle righe non nulle è linearmente indipendente in* $\mathbb{R}_n$.

Dimostrazione. Infatti se si traspone $\mathbf{U}$ si ottiene una matrice triangolare inferiore in cui, alla prima riga, vi è un coefficiente diverso da 0 nella prima colonna e coefficienti nulli nelle successive per cui si ha un caso speculare a quello della dimostrazione di proposizione 2.2.10. □

2.3 Basi e dimensioni

2.3.1 Basi

Osservazione 117. Dalla proposizione 2.2.8 deriva che la procedura della proposizione 2.2.3, usabile per eliminare da un insieme di generatori vettori che siano combinazione lineare degli altri, a un certo punto termina e precisamente quando l'insieme ottenuto è linearmente indipendente.

Definizione 2.3.1 (Base di V)**.** Un insieme $\mathscr{A}$ di vettori di V si dice una base di V se è un insieme di generatori ed è linearmente indipendente.

Osservazione 118. In pratica una base sta a uno spazio vettoriale come un sistema di riferimento sta allo spazio euclideo; fissata una base è possibile identificare gli elementi dello spazio vettoriale astratto mediante delle coordinate $\alpha_1, \dots, \alpha_n$.

Osservazione 119. Si noti che la definizione non esclude la possibilità che uno spazio vettoriale possa avere molteplici basi, differenti tra loro. Di fatto ogni spazio vettoriale di dimensione finita ha infinite basi, con l'eccezione dello spazio vettoriale nullo che ha come unica base l'insieme vuoto.

Definizione 2.3.2 (Componenti/coordinate di $\mathbf{v} \in V$)**.** Gli scalari $\alpha_1, \dots, \alpha_n$, dette *componenti* o *coordinate* del vettore $\mathbf{v}$ rispetto alla base $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ sono appunto univocamente determinati una volta che si è scelto quale vettore si voglia rappresentare e quale base impiegare.

Proposizione 2.3.1. *Se $\mathscr{B}$ è una base dello spazio vettoriale V, ogni elemento $\mathbf{v} \in V$ si può scrivere in modo unico come combinazione lineare dei vettori di $\mathscr{B}$.*

Dimostrazione. Infatti se $\mathscr{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ è la base e ipotizzando per assurdo che il vettore $\mathbf{v} \in V$ si possa scrivere con due set di coordinate α_i, β_i si ha:

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$$

Ma allora

$$\mathbf{0} = \mathbf{v} - \mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i - \sum_{i=1}^{n} \beta_i \mathbf{v}_i = \sum_{i=1}^{n} (\alpha_i - \beta_i) \mathbf{v}_i$$

e, per l'indipendenza lineare $(\alpha_i - \beta_i) = 0$ da cui $\alpha_i = \beta_i$. □

Teorema 2.3.2. *Sia V uno spazio vettoriale finitamente generato. Allora V ha una base.*

Dimostrazione. Se $\mathscr{A}$ è un insieme di generatori di V, applichiamo la procedura della proposizione 2.2.3 fino a che possibile ottenendo un insieme $\mathscr{B}$ che è ancora un insieme di generatori ed è anche linearmente indipendente. □

Esempio 2.3.1. Lo spazio vettoriale nullo $\{\mathbf{0}\}$ ha come base l'insieme vuoto: questo è un insieme di generatori di $\{\mathbf{0}\}$, per definizione, ed è linearmente indipendente

Esempio 2.3.2 (Base canonica di $\mathbb{C}^n$)**.** Lo spazio vettoriale $\mathbb{C}^n$ ha come base l'insieme $\mathscr{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ dei vettori coordinati; infatti ogni elemento di $\mathbf{v} = [\alpha_1 \, \ldots \, \alpha_n]^T \in \mathbb{C}^n$ può essere scritto in un solo modo come combinazione lineare dei vettori della base

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i \tag{2.14}$$

e quindi $\mathscr{E}$ è insieme di generatori. Inoltre

$$\sum_{i=1}^{n} \alpha_i \mathbf{e}_i = [\alpha_1 \, \ldots \, \alpha_n]^T$$

e perciò se $\sum_{i=1}^{n} \alpha_i \mathbf{e}_i = \mathbf{0}$ segue che $\alpha_1 = \ldots = \alpha_n = 0$, ossia che $\mathscr{E}$ è anche linearmente indipendente.

Esempio 2.3.3. Anche lo spazio vettoriale delle matrici ha delle proprie basi: la base canonica consiste delle matrici $\mathbf{E}_{ij}$ che hanno l'elemento di posto (i, j) pari a 1 e i rimanenti a 0. Ad esempio, nel caso delle matrici quadrate 2×2 la base canonica è composta da:

$$\mathbf{E}_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \mathbf{E}_{21} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \mathbf{E}_{12} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{E}_{12} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Esempio 2.3.4. Verifichiamo che i seguenti tre vettori di $\mathbb{C}^3$ formano una base di $\mathbb{C}^3$:

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ 5 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix}$$

Per verificare innanzitutto che $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ è insieme di generatori, si tratta di vedere se l'insieme genera $\mathbb{C}^3$; ossia, se preso un generico vettore $\mathbf{v} = [a \, b \, c]^T \in \mathbb{C}^3$, vi sono tre scalari $x_1, x_2, x_3 \in \mathbb{C}$ tali che

$$x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + x_3 \mathbf{v}_3 = \mathbf{v}$$

Ciò corrisponde al fatto che il sistema $\mathbf{A}\mathbf{x} = \mathbf{v}$ con $\mathbf{x} = [x_1 \, x_2 \, x_3]^T$ e $\mathbf{A} = [\mathbf{v}_1 \, \mathbf{v}_2 \, \mathbf{v}_3]$ abbia soluzioni.
Per verificare che $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ sia linearmente indipendente, occorre invece verificare se il sistema $\mathbf{A}\mathbf{x} = \mathbf{0}$ abbia come unica soluzione lo $\mathbf{0}$.
Il teorema 1.6.12 dice che ciascuno dei fatti precedenti è equivalente all'invertibilità della matrice $\mathbf{A}$.

Esempio 2.3.5. Si provi che l'insieme

$$S = \left\{ \mathbf{s}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \mathbf{s}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{s}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -2 \end{bmatrix}, \mathbf{s}_4 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{s}_5 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

è un insieme di generatori per $\mathbb{C}^4$ e si trovi una base di $\mathbb{C}^4$ contenuta in S.
Per verificare che l'insieme sia di generatori, dato un generico $v = [a \, b \, c \, d]$ con

$a, b, c, d \in \mathbb{C}$ bisogna verificare che $\mathbf{Ax} = \mathbf{v}$ ammetta soluzioni:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & a \\ 0 & 1 & 1 & 0 & 1 & b \\ 0 & 0 & 0 & 1 & -1 & c \\ 2 & 0 & -2 & 0 & 1 & d \end{bmatrix} \qquad r_4 = r_4 - 2r_1$$

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & a \\ 0 & 1 & 1 & 0 & 1 & b \\ 0 & 0 & 0 & 1 & -1 & c \\ 0 & -2 & -2 & -2 & -1 & d-2a \end{bmatrix} \qquad r_4 = r_4 + 2r_2$$

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & a \\ 0 & 1 & 1 & 0 & 1 & b \\ 0 & 0 & 0 & 1 & -1 & c \\ 0 & 0 & 0 & -2 & 1 & d-2a+2b \end{bmatrix} \qquad r_4 = r_4 + 2r_3$$

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & a \\ 0 & 1 & 1 & 0 & 1 & b \\ 0 & 0 & 0 & 1 & -1 & c \\ 0 & 0 & 0 & 0 & -1 & d-2a+2b+2c \end{bmatrix}$$

Il sistema ha infinite soluzioni pertanto S è un insieme di generatori. Le colonne 1,2,4 e 5 sono dominanti; la 3 è non dominante e può essere combinazione lineare delle altre. Verifichiamolo partendo già dalla forma ridotta (i passaggi precedenti sarebbero stati gli stessi):

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{bmatrix} \rightarrow \begin{cases} x_5 = 0 \\ x_4 - x_5 = 0 \\ x_2 + x_5 = 1 \\ x_1 + x_2 + x_4 + x_5 = 0 \end{cases} \qquad \begin{cases} x_5 = 0 \\ x_4 = 0 \\ x_2 = 1 \\ x_1 = -1 \end{cases}$$

Pertanto la terza colonna può esser scritta come

$$\begin{bmatrix} 0 \\ 1 \\ 0 \\ 2 \end{bmatrix} = -1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix} + 1 \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

e $\mathbf{s}_3 = \mathbf{s}_2 - \mathbf{s}_1$. La base è $\{\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_4, \mathbf{s}_5\}$.

Esempio 2.3.6. L'insieme $\{\mathbf{p}_1 = 1 - X; \mathbf{p}_2 = 1 + X; \mathbf{p}_3 = 1 - X^2\}$ è una base di $\mathscr{P}_3(\mathbb{C})$. Per verificarlo dimostriamo che è sia un insieme di generatori che un insieme linearmente indipendente.
Notiamo innanzitutto che un generico polinomio $\mathbf{b} = a_0 + a_1X + a_2X^2$ possa esser scritto come combinazione lineare di $\{\mathbf{p}_1; \mathbf{p}_2; \mathbf{p}_3\}$, ossia esistono $\alpha_1, \alpha_2, \alpha_3$ tali che

$$\alpha_1\mathbf{p}_1 + \alpha_2\mathbf{p}_2 + \alpha_3\mathbf{p}_3 = a_0 + a_1X + a_2X^2$$

Dato che in $\mathscr{P}_3(\mathbb{C})$ i vettori $\mathbf{p}_1$, $\mathbf{p}_2$ e $\mathbf{p}_3$ possono essere riscritti come:

$$\mathbf{p}_1 = 1 - 1X + 0X^2$$
$$\mathbf{p}_2 = 1 + 1X + 0X^2$$
$$\mathbf{p}_3 = 1 + 0X + -1X^2$$

Se ci si focalizza solo sui coefficienti del polinomio (tralasciando le X) si può determinare le incognite $\alpha_1, \alpha_2, \alpha_3$ della combinazione lineare risolvendo il sistema $\mathbf{Ax} = \mathbf{b}$ con $\mathbf{A} = [\mathbf{p}_1\, \mathbf{p}_2\, \mathbf{p}_3]$ e $\mathbf{x} = [\alpha_1\, \alpha_2\, \alpha_3]^T$:

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

ossia attraverso il sistema determiniamo $\alpha_1, \alpha_2, \alpha_3$ i coefficienti del polinomio risultante siano a_0, a_1, a_2; in altre parole il sistema

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = a_0 \\ -\alpha_1 + \alpha_2 = a_1 \\ -\alpha_3 = a_2 \end{cases}$$

la matrice aumentata, sviluppata in EG conduce a

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & a_0 \\ -1 & 1 & 0 & a_1 \\ 0 & 0 & -1 & a_2 \end{array}\right] \dots \left[\begin{array}{ccc|c} 1 & 1 & 1 & a_0 \\ 0 & 1 & 1/2 & (a_0 + a_1)/2 \\ 0 & 0 & 1 & -a_2 \end{array}\right]$$

da cui si vede che ogni sistema del genere, al variare di a_0, a_1, a_2 (ossia del polinomio di secondo grado da rappresentare) ha un'unica soluzione. Dunque la soluzione esiste e l'insieme $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ è di generatori.
Inoltre quando $a_0 = 0, a_1 = 0, a_2 = 0$, cioè il sistema è omogeneo, anch'esso ha un'unica soluzione, quella con $\alpha_1 = 0, \alpha_2 = 0, \alpha_3 = 0$; perciò l'insieme dato è anche linearmente indipendente, in accordo all'osservazione 115.

Esempio 2.3.7. Anche l'insieme $\left\{1; X; X^2\right\}$ è una base di $\mathscr{P}_3(\mathbb{C})$, come si verifica immediatamente: il sistema lineare costruito come il precedente ha come matrice dei coefficienti l'identità.

Esempio 2.3.8. Se $\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ è una base di V, per ogni scalare $\gamma \neq 0$ anche

$$\mathscr{B}' = \{\mathbf{v}_1' = \gamma\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

è una base di V. Dimostriamo che $\mathscr{B}'$ è un insieme di generatori: se $\mathbf{v} \in V$ dalla definizione di base sappiamo che esistono scalari $\alpha_1, \alpha_2, \dots, \alpha_n$ tali che:

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

e quindi

$$\mathbf{v} = \alpha_1\gamma^{-1}\mathbf{v}_1' + \alpha_2\mathbf{v}_2 + \dots \alpha_n\mathbf{v}_n$$

ma essendo anche $\alpha_1\gamma^{-1}$ uno scalare, si ha che $\mathbf{v} \in \langle\mathscr{B}'\rangle$.
Per la dimostrazione che $\mathscr{B}'$ è linearmente indipendente: dato che $\mathscr{B}$ è linearmente indipendente si ha che $\mathbf{v}_1 \neq \alpha_2\mathbf{v}_2 + \dots \alpha_n\mathbf{v}_n, \forall\alpha_i \in \mathbb{C}$. Questo implica che (moltiplicando per γ entrambi i lati) $\gamma\mathbf{v}_1 \neq \sum_{i=2}^{n} \gamma\alpha_i\mathbf{v}_i$ ossia $\mathbf{v}_1' \neq \sum_{i=2}^{n} \beta_i\mathbf{v}_i$ con $\beta_i = \gamma\alpha_i$; pertanto non essendo $\mathbf{v}_1'$ una combinazione lineare dei rimanenti $\mathbf{v}_2, \dots, \mathbf{v}_n$, l'insieme che formano è linearmente indipendente.

2.3.2 Dimensione

Osservazione 120. Negli esempi che abbiamo illustrato abbiamo visto che basi diverse dello stesso spazio vettoriale hanno in comune il numero di elementi: questo è un fatto fondamentale della teoria degli spazi vettoriali. Procediamo a passi nella dimostrazione ricordando due fatti: quando diciamo che un insieme finito di vettori ha n elementi, intendiamo contare anche gli elementi eventualmente ripetuti; infine, in un insieme linearmente indipendente non ci sono ripetizioni.

Teorema 2.3.3. *Sia V uno spazio vettoriale finitamente generato. Sia $\mathscr{A}$ un insieme di generatori di V (con $\operatorname{Card}(\mathscr{A}) = n$) e $\mathscr{B}$ un insieme linearmente indipendente di vettori di V (con $\operatorname{Card}(\mathscr{B}) = m$); allora $m \leq n$.*

Osservazione 121. In altre parole un insieme linearmente indipendente ha una cardinalità al più uguale (o minore) a quella di un insieme di generatori dello spazio.

Dimostrazione. Se $\mathscr{B}$ è l'insieme vuoto (linearmente indipendente per convenzione) non c'è niente da dimostrare (nel senso che qualsiasi insieme di generatori di V ha cardinalità ≥ 0). Pertanto assumiamo che $m > 0$: allora necessariamente $n \geq 1$. Infatti se fosse $n = 0$ sarebbe $\mathscr{A} = \emptyset$ e $\mathscr{A}$ genererebbe lo spazio nullo $\{\mathbf{0}\}$, nel quale non si potrebbero scegliere $m \geq 1$ vettori non nulli e distinti per formare $\mathscr{B}$ (si veda proposizione 2.2.7); pertanto $\mathscr{B}$ non sarebbe un insieme linearmente indipendente, differentemente dall'ipotesi.
Supponiamo perciò:

$$\mathscr{B} = \{\mathbf{w}_1, \dots, \mathbf{w}_m\} \quad \mathscr{A} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$$

Procediamo per induzione su k dimostrando la seguente affermazione: se $0 \leq k \leq m$, esiste un insieme di generatori $\mathscr{C}_k$ (di n elementi) formato da k vettori dell'insieme $\mathscr{B}$ e da $n - k$ vettori dell'insieme $\mathscr{A}$. Questo prova che $m \leq n$, altrimenti se fosse $m > n$ non potremmo costruire l'insieme $\mathscr{C}_m$ (credo poiché dovremmo scegliere m elementi da un insieme e $n - m < 0$ da un altro).
Per il passo base, nel caso $k = 0$ si prendono solo ed esclusivamente gli elementi di $\mathscr{A}$, ossia $\mathscr{C}_0 = \mathscr{A}$; ed essendo $\mathscr{A}$ un insieme di generatori, lo è anche $\mathscr{C}_0$.
Per ipotesi allora supponiamo che la cosa valga e di avere a disposizione l'insieme $\mathscr{C}_k$: e che $k + 1 \leq m$ (ossia $k \leq m - 1$). Non è restrittivo supporre che

$$\mathscr{C}_k = \left\{ \underbrace{\mathbf{w}_1, \dots, \mathbf{w}_k}_{k \text{ da } \mathscr{B}}, \underbrace{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n}_{n-k \text{ da } \mathscr{A}} \right\}$$

eventualmente riordinando gli insiemi da cui siamo partiti. Per ipotesi essendo $\mathscr{C}_k$ un insieme di generatori possiamo scrivere il $k+1$-esimo elemento di $\mathscr{B}$ come combinazione degli elementi di $\mathscr{C}_k$:

$$\mathbf{w}_{k+1} = \sum_{i=1}^{k} \alpha_i \mathbf{w}_i + \sum_{j=k+1}^{n} \beta_j \mathbf{v}_j \tag{2.15}$$

Non è possibile che $\beta_{k+1} = \ldots = \beta_n = 0$, perché in tal caso il $k + 1$-esimo elemento di $\mathscr{B}$

$$\mathbf{w}_{k+1} = \sum_{i=1}^{k} \alpha_i \mathbf{w}_i$$

sarebbe una combinazione dei primi k e l'insieme $\mathscr{B}$ non sarebbe linearmente indipendente. Dunque possiamo supporre che almeno uno sia non nullo, $\beta_{k+1} \neq 0$, eventualmente riordinando. Considerando ora l'insieme $\mathscr{C}_k \cup \{\mathbf{w}_{k+1}\}$:

$$\mathscr{C}'_k = \{\mathbf{w}_1, \ldots, \mathbf{w}_k, \mathbf{w}_{k+1}, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$$

poiché contiene un insieme di generatori ($\mathscr{C}_k$, cui si è aggiunto $\mathbf{w}_{k+1}$), anche $\mathscr{C}'_k$ è un insieme di generatori.
Ora che abbiamo introdotto un elemento da $\mathscr{B}$ possiamo eliminarne uno da $\mathscr{A}$ per tornare a un insieme di generatore di n elementi complessivamente (ma dove ne abbiamo preso uno in più da $\mathscr{B}$ e uno in meno da $\mathscr{A}$, e quindi ci spostiamo verso $\mathscr{C}_{k+1}$); rielaborando la 2.15:

$$\mathbf{w}_{k+1} = \sum_{i=1}^{k} \alpha_i \mathbf{w}_i + \sum_{j=k+1}^{n} \beta_j \mathbf{v}_j$$
$$\mathbf{w}_{k+1} = \sum_{i=1}^{k} \alpha_i \mathbf{w}_i + \sum_{j=k+2}^{n} \beta_j \mathbf{v}_j + \beta_{k+1}\mathbf{v}_{k+1}$$

da cui

$$\mathbf{v}_{k+1} = \sum_{i=1}^{k+1} \frac{\alpha_i}{\beta_{k+1}} \mathbf{w}_i + \sum_{j=k+2}^{n} \frac{\beta_j}{\beta_{k+1}} \mathbf{v}_j$$

(con $\alpha_{k+1} = -1$) e quindi l'insieme che si ottiene eliminando da $\mathscr{C}'_k$ il vettore $\mathbf{v}_{k+1}$ è ancora un insieme di generatori (perché l'elemento eliminato è una combinazione lineare degli elementi rimasti). Chiamiamo $\mathscr{C}_{k+1}$ questo insieme perché ha le proprietà richieste. □

Osservazione 122. La tecnica impiegata nella dimostrazione è importante di per se, quindi la isoliamo in un risultato a parte.

Teorema 2.3.4. *Siano $\mathscr{A}$ un insieme di generatori dello spazio V (con* $\operatorname{Card}(\mathscr{A}) =$ *n elementi) e $\mathscr{B}$ un insieme linearmente indipendente di vettori di V (di* $\operatorname{Card}(\mathscr{B}) =$ *m elementi). Allora esiste un insieme di generatori $\mathscr{C}$ formato da tutti i vettori di $\mathscr{B}$ e da $n - m$ vettori di $\mathscr{A}$.*

Osservazione 123. Siccome ogni spazio che consideriamo ha un insieme di generatori, possiamo usare questo per ottenere che *ogni insieme linearmente indipendente può essere completato a una base.*

Esempio 2.3.9 (Completamento ad una base di $\mathbb{C}^n$)**.** Per illustrare il risultato precedente applichiamo la tecnica della dimostrazione del teorema 2.3.3 agli insiemi $\mathscr{A} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4\}$ e $\mathscr{B} = \{\mathbf{u}, \mathbf{v}\}$ in $\mathbb{C}^4$ dove

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

Si esegue l'EG sulla matrice $[\mathbf{u}\,\mathbf{v}\,\mathbf{e}_1\,\mathbf{e}_2\,\mathbf{e}_3\,\mathbf{e}_4]$

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & -2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1/2 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

Le colonne dominanti sono le prime quattro e quindi l'insieme $\{\mathbf{u}, \mathbf{v}, \mathbf{e}_1, \mathbf{e}_2\}$ è linearmente indipendente, quindi una base di $\mathbb{C}^4$. Per maggiori dettagli si veda il Teorema 2.5.4

Esempio 2.3.10. Si trovi una base di $\mathbb{C}^4$ contenente i vettori

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

Faccio la riduzione della matrice $[\mathbf{v}_1\, \mathbf{v}_2\, \mathbf{I}_4]$ trovando le colonne corrispondenti a quelle dominanti

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad r_2 = r_2 - r_1$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad r_4 = r_4 - r_1$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 \end{bmatrix} \qquad r_4 = r_4 + r_3$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 \end{bmatrix} \qquad r_2 = r_2 + 2r_3$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 \end{bmatrix} \qquad \text{riordino}$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 \end{bmatrix} \qquad r_4 = r_4 - r_3$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 2 & 0 \\ 0 & 0 & 0 & -1 & -1 & 1 \end{bmatrix}$$

Le colonne dominanti sono le prime 4 pertanto la base è composta da $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1, \mathbf{e}_2\}$.

Teorema 2.3.5. *Sia V uno spazio vettoriale finitamente generato; siano $\mathscr{A}$ e $\mathscr{B}$ basi di V. Allora $\mathscr{A}$ e $\mathscr{B}$ hanno lo stesso numero di elementi.*

Dimostrazione. Se $\operatorname{Card}(\mathscr{A}) = n$ e $\operatorname{Card}(\mathscr{B}) = m$:

- siccome $\mathscr{A}$ è un insieme di generatori e $\mathscr{B}$ è linearmente indipendente, per il teorema 2.3.3 si ha che $m \leq n$;

- siccome $\mathscr{B}$ è un insieme di generatori e $\mathscr{A}$ è linearmente indipendente, per il teorema 2.3.3 si ha che $n \leq m$.

Pertanto deve essere $m = n$. □

Definizione 2.3.3 (Dimensione di uno spazio V)**.** Se V è uno spazio vettoriale finitamente generato, si chiama dimensione e si indica con $\dim V$ il numero di elementi di una sua base (ossia di tutte).

Esempio 2.3.11. Lo spazio $\mathbb{C}^n$ ha dimensione n, perché la base canonica $\mathscr{E}$ ha n elementi.

Esempio 2.3.12. Lo spazio $\{\mathbf{0}\}$ ha dimensione 0; una sua base è l'insieme vuoto.

Esempio 2.3.13. Sia $\mathbf{U}$ una matrice $m \times n$ in forma ridotta. Scriviamola come

$$\mathbf{U} = [\mathbf{u}_1\, \mathbf{u}_2\, \dots\, \mathbf{u}_n]$$

e consideriamo $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ che è un sottospazio di $\mathbb{C}^n$.
Per quanto visto nella proposizione 2.2.10, l'insieme delle colonne dominanti è linearmente indipendente. Ogni altra colonna è combinazione lineare delle colonne dominanti che la precedono (come suggerisce l'esempio 2.2.6) e pertanto le colonne dominanti formano una base di V.
Quindi se k è il numero delle colonne dominanti (che è uguale al numero di righe non nulle), $\dim V = k$.

Esempio 2.3.14. Sia V il sottospazio di $\mathbb{C}^4$ generato dall'insieme

$$S = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 4 \\ 7 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Si calcoli la dimensione di V e si trovi una base di V contenuta in S.
Occorre fare la riduzione di Gauss; $\dim V$ è il numero di colonne dominanti e la

base sono le corrispondenti colonne di $\mathbf{A}$:

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 4 & 0 \\ 2 & 4 & 1 & 7 & 1 \end{bmatrix} \qquad r_3 = r_3 - r_1$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ 2 & 4 & 1 & 7 & 1 \end{bmatrix} \qquad r_4 = r_3 - 2r_1$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 \end{bmatrix} \qquad r_4 = r_4 - r_3$$

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad r_4 = r_4 - r_3$$

Le colonne dominanti sono 2 e sono la prima e la terza, pertanto la base è composta dalle suddette e $\dim V = 2$.

Osservazione 124. Per calcolare la dimensione di uno spazio vettoriale dobbiamo contare gli elementi di una sua base; in certe situazioni non è facile determinarla perciò si può approssimarla accontentandosi di maggiorarla dal numero di elementi di un insieme di generatori (teorema 2.3.7 in seguito) o minorarla dal numero di elementi un insieme linearmente indipendente (teorema 2.3.8 in seguito).
Cominciamo però da un lemma intuitivo ma dalla dimostrazione non del tutto ovvia.

Lemma 2.3.6. *Se V è uno spazio vettoriale finitamente generato, con* $\dim V = n$*, e U è un sottospazio di V; allora U è finitamente generato e* $\dim U \leq \dim V$.

Dimostrazione. Procedendo per assurdo assumiamo che U non sia finitamente generato. Pertanto nessun sottoinsieme di vettori di U è un insieme di generatori; in particolare $U \neq \{\mathbf{0}\}$ perché altrimenti l'insieme vuoto sarebbe un insieme di generatori. Dunque esiste $\mathbf{u}_1 \in U, \mathbf{u}_1 \neq \mathbf{0}$ (ossia U ha un elemento che non è il vettore nullo).
L'insieme $\mathscr{A}_1 = \{\mathbf{u}_1\}$ non genera U (per ipotesi dell'assurdo), quindi esiste un vettore $\mathbf{u}_2 \in U$ tale che $\mathbf{u}_2 \notin \langle \mathscr{A}_1 \rangle$. Ma allora (si veda il teorema 2.2.9) l'insieme $\mathscr{A}_2 = \{\mathbf{u}_1, \mathbf{u}_2\}$ è linearmente indipendente.
Possiamo andare avanti per induzione allo stesso modo: se abbiamo ottenuto l'insieme linearmente indipendente $\mathscr{A}_k$ in U, questo non genera U e quindi, prendendo $\mathbf{u}_{k+1} \in U \setminus \langle \mathscr{A}_k \rangle$, l'insieme $\mathscr{A}_{k+1} = \mathscr{A}_k \sqcup \{\mathbf{u}_{k+1}\}$ è linearmente indipendente.
Dato $\dim V = n$ esiste una base di n elementi; ma abbiamo costruito con procedimento ricorsivo l'insieme $\mathscr{A}_{n+1}$ che è linearmente indipendente in U, quindi anche in V (poiché U è sottoinsieme di V). Questo contraddice il teorema 2.3.3:

l'insieme linearmente indipendente $\mathscr{A}_{n+1}$ ha cardinalità maggiore di una base di V (ossia di un insieme di generatori). Dunque U è finitamente generato e perciò ha una base che avrà m elementi, con $m \leq \dim V = n$, sempre per il teorema 2.3.3. □

Teorema 2.3.7. *Sia V uno spazio vettoriale finitamente generato. Allora $\dim V$ è il numero minimo di elementi di un insieme di generatori di V.*

Dimostrazione. Sia $\mathscr{A}$ un insieme di generatori con numero minimo n di elementi: ossia ogni altro insieme di generatori ha numero di elementi $\geq n$.
Ipotizziamo per assurdo che $\mathscr{A}$ non sia linearmente indipendente; possiamo ottenere un altro insieme di generatori togliendo un vettore che sia combinazione lineare dei rimanenti (proposizione 2.2.3). Ma dato che avevamo detto che tutti gli altri insiemi di generatori hanno un numero maggiore di elementi non dovremmo avere un insieme di generatori diminuendo il numero di elementi. Allora $\mathscr{A}$ è linearmente indipendente, è una base di V, e $\dim V = n$. □

Osservazione 125. In altre parole eliminiamo vettori dipendenti da un insieme di generatori sino a che tutti sono indipendenti: il numero di elementi diminuisce e alla fine coincide con $\dim V$, ossia il numero di elementi della base (dove sono indipendenti).

Teorema 2.3.8. *Sia V uno spazio vettoriale finitamente generato. Allora $\dim V$ è il numero massimo di elementi di un insieme linearmente indipendente in V.*

Dimostrazione. Il numero massimo di elementi esiste, perché un insieme linearmente indipendente ha al più tanti elementi quanti un insieme di generatori (teorema 2.3.3) e l'insieme V ha dei sottoinsiemi che fungono da generatori (al massimo si usa V stesso).
Ipotizziamo che $\mathscr{A}$ sia un insieme linearmente indipendente con un numero massimo di vettori (ossia ha lo stesso numero di un insieme di generatori di V, e un elemento aggiuntivo sarebbe una mera combinazione lineare dei vettori già presenti in $\mathscr{A}$), ma che $\langle \mathscr{A} \rangle \neq V$. Allora per il teorema 2.2.9 potremmo aggiungere ad $\mathscr{A}$ un vettore di $\mathbf{v} \in V$ tale che $\mathbf{v} \notin \langle \mathscr{A} \rangle$, ottenendo ancora un insieme linearmente indipendente. Ma è assurdo perché un insieme indipendente ha al più tanti elementi quanto un insieme di generatori e qui partendo da un insieme $\mathscr{A}$ col numero pari a quello di un insieme di generatori abbiamo aumentato di uno. Pertanto $\mathbf{v} \in \langle \mathscr{A} \rangle$ e aggiungerlo renderebbe $\mathscr{A}$ non più indipendente. □

Teorema 2.3.9. *Se $\dim V = n$ e $\mathscr{A}$ è un insieme linearmente indipendente di vettori di V con n elementi, allora $\mathscr{A}$ è una base di V.*

Dimostrazione. Se $\mathscr{A}$ non fosse una base di V, il teorema 2.3.4 ci permetterebbe di completarla a una base che avrebbe $m > n$ elementi. Ma abbiamo ipotizzato che la base abbia n elementi, assurdo. □

Proposizione 2.3.10. *Come conseguenza otteniamo una tecnica per dimostrare che un sottospazio U di V coincide con V*

$$U = V \iff \dim U = \dim V \tag{2.16}$$

Dimostrazione. Infatti una implicazione è ovvia: se $U = V$, allora di sicuro $\dim U = \dim V$. Se viceversa $\dim U = \dim V$ abbiamo che una base di U ha numero di elementi pari al numero massimo di componenti di un insieme linearmente indipendente in V (2.3.8). Pertanto è una base anche di V (2.3.9); da questo deriva che ogni vettore di V è combinazione lineare di una base di U e quindi appartiene a U, ossia $V \subseteq U$. E svolgendo la trafila a parti inverse si conclude che $U = V$. □

2.3.3 Altri argomenti

2.3.3.1 Introduzione allo spazio delle colonne

Osservazione 126. Introduciamo uno dei quattro cosiddetti sottospazi fondamentali di una matrice che saranno esaminati più in dettaglio nel seguito.

Definizione 2.3.4 (Spazio delle colonne di **A**)**.** Se $\mathbf{A}$ è una matrice $m \times n$, $\mathbf{A} = [\mathbf{v}_1 \dots \mathbf{v}_n]$ lo spazio delle colonne è definito da

$$\mathrm{C}\,(\mathbf{A}) = \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle \tag{2.17}$$

Osservazione 127. Si tratta di un sottospazio di $\mathbb{C}^m$, quindi $\dim \mathrm{C}\,(\mathbf{A}) \le m$ (lemma 2.3.6). Inoltre $\mathrm{C}\,(\mathbf{A})$ ha un insieme di generatori con n elementi (l'insieme delle colonne $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, banalmente), quindi $\dim \mathrm{C}\,(\mathbf{A}) \le n$ (teorema 2.3.3), perché non è detto che le colonne siano linearmente indipendenti.

Osservazione 128. Nel caso $\mathbf{A}$ sia in forma ridotta, sappiamo calcolare la dimensione di $\mathrm{C}\,(\mathbf{A})$: è precisamente il numero di colonne dominanti. Per il caso generale ($\mathbf{A}$ non in forma ridotta) avremo bisogno di un concetto più astratto (applicazione lineare) approfondito in seguito.

2.3.3.2 Indipendenza di sottospazi

Proposizione 2.3.11 (Dimensione della somma di sottospazi)**.** *Se U_1 e U_2 sono sottospazi di V allora la loro somma (definizione 2.1.7) ha dimensione:*

$$\dim(U_1 + U_2) = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2) \tag{2.18}$$

Dimostrazione. Sia $\mathscr{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ una base di $U_1 \cap U_2$ (per cui $\dim(U_1 \cap U_2) = k$). Possiamo estendere $\mathscr{B}$ a una base di U_1, e ipotizzando che lo spazio sia di dimensione m:

$$\mathscr{B}' = \{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_m\}$$

Tale base ha effettivamente $k + (m - (k + 1) + 1) = m$ elementi, pertanto $\dim U_1 = m$. Allo stesso modo estenderlo a una base di U_2:

$$\mathscr{B}'' = \{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{w}_{k+1}, \dots, \mathbf{w}_n\}$$

(similmente $\dim U_2 = n$). Consideriamo allora la:

$$\mathscr{C} = \{\mathbf{u}_1, \dots, \mathbf{u}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_m, \mathbf{w}_{k+1}, \dots, \mathbf{w}_n\}$$

che ha $k+(m-(k+1)+1)+(n-(k+1)+1) = m+n-k$ elementi, e verifichiamo che $\mathscr{C}$ è una base di $U_1 + U_2$ (la cui dimensione è appunto $\dim U_1 + \dim U_2 -$

$\dim(U_1 \cap U_2) = m + n - k$).
Partiamo verificando che $\mathscr{C}$ è un insieme di generatori. Se $\mathbf{x} \in U_1$ è un generico elemento di U_1, lo possiamo scrivere come combinazione lineare dei vettori di $\mathscr{B}'$:

$$\mathbf{x} = \alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k + \gamma_{k+1} \mathbf{v}_{k+1} + \ldots + \gamma_m \mathbf{v}_m$$

Se $\mathbf{y} \in U_2$ abbiamo, in maniera speculare per $\mathscr{B}''$:

$$\mathbf{y} = \beta_1 \mathbf{u}_1 + \ldots + \beta_k \mathbf{u}_k + \delta_{k+1} \mathbf{w}_{k+1} + \ldots + \delta_n \mathbf{w}_n$$

Allora un generico vettore di $U_1 + U_2$ può essere scritto come combinazione lineare degli elementi di $\mathscr{C}$:

$$\mathbf{x}+\mathbf{y} = (\alpha_1+\beta_1)\mathbf{u}_1+\ldots+(\alpha_k+\beta_k)\mathbf{u}_k+\gamma_{k+1}\mathbf{v}_{k+1}+\ldots+\gamma_m\mathbf{v}_m+\delta_{k+1}\mathbf{w}_{k+1}+\ldots+\delta_n\mathbf{w}_n \tag{2.19}$$

e quindi $\mathscr{C}$ è un insieme di generatori di $U_1 + U_2$.
Per verificare l'indipendenza lineare dei generatori, supponiamo ora che la combinazione lineare della base sia nulla deriviamo che i rispettivi coefficienti siano nulli. Se:

$$\mathbf{0} = \alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k + \gamma_{k+1} \mathbf{v}_{k+1} + \ldots + \gamma_m \mathbf{v}_m + \delta_{k+1} \mathbf{w}_{k+1} + \ldots + \delta_n \mathbf{w}_n \tag{2.20}$$

allora un generico elemento di U_1:

$$\begin{aligned} \mathbf{x} &= \alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k + \gamma_{k+1} \mathbf{v}_{k+1} + \ldots + \gamma_m \mathbf{v}_m \\ &= -(\delta_{k+1} \mathbf{w}_{k+1} + \ldots + \delta_n \mathbf{w}_n) \in U_1 \cap U_2 \end{aligned}$$

ossia dato che è anche combinazione lineare dei $\mathbf{w}$ appartiene anche a U_2 e quindi in definitiva fa parte dell'intersezione tra U_1 e U_2.
Perciò, dato che $\mathscr{B}$ è un'altra base di $U_1 \cap U_2$, $\mathbf{x} = \beta_1 \mathbf{u}_1 + \ldots + \beta_k \mathbf{u}_k$ e quindi da

$$\mathbf{x} = -(\delta_{k+1} \mathbf{w}_{k+1} + \ldots + \delta_n \mathbf{w}_n) = \beta_1 \mathbf{u}_1 + \ldots + \beta_k \mathbf{u}_k$$

si ha che

$$\delta_{k+1} \mathbf{w}_{k+1} + \ldots + \delta_n \mathbf{w}_n = -\beta_1 \mathbf{u}_1 - \ldots - \beta_k \mathbf{u}_k$$

ed andando a sostituire il secondo membro al posto di $\delta_{k+1} \mathbf{w}_{k+1} + \ldots + \delta_n \mathbf{w}_n$ nella 2.20 si ha

$$(\alpha_1 - \beta_1)\mathbf{u}_1 + \ldots + (\alpha_k - \beta_k)\mathbf{u}_k + \gamma_{k+1} \mathbf{v}_{k+1} + \ldots + \gamma_m \mathbf{v}_m = \mathbf{0}$$

da cui $\gamma_{k+1} = \ldots = \gamma_m = 0$ perché $\mathscr{B}'$ è linearmente indipendente (cioè ora l'eguaglianza dipende da vettori di una sola base, che essendo linearmente indipendente impone che per la soddisfazione dell'equazione debbano essere tutti nulli, compresi i γ).
Analogamente se si fa tutta la trafila per $\mathbf{y} \in U_2$ si conclude che $\delta_{k+1} = \ldots = \delta_n = 0$. Ma allora (se sono tutti nulli), si ha che

$$\alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k = \mathbf{0}$$

e $\alpha_1 = \ldots = \alpha_k = 0$ perché si sta sfruttando una base di $U_1 \cap U_2$ e l'uguaglianza è soddisfatta solo se tutti i coefficienti sono nulli. □

Osservazione 129. Nel caso della somma diretta dovrebbe essere che $\dim(U_1 + U_2) = \dim U_1 + \dim U_2$ dato che $\dim(U_1 \cap U_2) = \dim(\{\mathbf{0}\}) = 0$, essendo $\emptyset$ una base di $\{\mathbf{0}\}$.

Proposizione 2.3.12. *Siano U_1 e U_2 sottospazi di V. Le seguenti affermazioni sono equivalenti:*

1. $U_1 + U_2 = U_1 \oplus U_2$
2. *ogni elemento di $U_1 + U_2$ si scrive in modo unico come somma di un elemento di U_1 e un elemento di U_2*
3. *se $\mathbf{u}_1 \in U_1$, $\mathbf{u}_2 \in U_2$ e $\mathbf{u}_1 + \mathbf{u}_2 = \mathbf{0}$ allora $\mathbf{u}_1 = \mathbf{0}$ e $\mathbf{u}_2 = \mathbf{0}$*

Dimostrazione. Per

- l'implicazione $1 \to 2$: per assurdo ipotizziamo che l'elemento si possa scrivere in 2 modi, ossia $\mathbf{x} + \mathbf{y} = \mathbf{x}' + \mathbf{y}'$ con $\mathbf{x}, \mathbf{x}' \in U_1$ e $\mathbf{y}, \mathbf{y}' \in U_2$ e $\mathbf{x} \neq \mathbf{x}'$ e $\mathbf{y} \neq \mathbf{y}'$. Allora spostando a destra gli elementi di U_1 e a sinistra quelli di U_2 si ha
$$\mathbf{x} - \mathbf{x}' = \mathbf{y}' - \mathbf{y} \in U_1 \cap U_2$$
il primo vettore appartiene a U_1, il secondo a U_2, ma essendo uguali allora appartengono sia all'uno che all'altro insieme quindi all'intersezione. Per ipotesi $U_1 \cap U_2 = \{\mathbf{0}\}$, quindi se $\mathbf{x} - \mathbf{x}' = \mathbf{0} = \mathbf{y}' - \mathbf{y}$, allora $\mathbf{x} = \mathbf{x}'$ e $\mathbf{y} = \mathbf{y}'$ e arriviamo a contraddire l'ipotesi che $\mathbf{x} \neq \mathbf{x}'$ e $\mathbf{y} \neq \mathbf{y}'$;
- l'implicazione $2 \to 3$: è evidente perché $\mathbf{0} = \mathbf{u}_1 + \mathbf{u}_2 = \mathbf{0} + \mathbf{0}$ (due elementi opposti non appartengono a sottospazi diversi);
- l'implicazione $3 \to 1$: sia $\mathbf{v} \in U_1 \cap U_2$, e $\mathbf{0} = \mathbf{v} + (-\mathbf{v})$. Ma dato che $\mathbf{v} \in U_1$ e $-\mathbf{v} \in U_2$, per la 3 si ha che $\mathbf{v} = -\mathbf{v} = \mathbf{0}$.

□

Definizione 2.3.5 (Sottospazi indipendenti)**.** Diremo che i sottospazi $U_1, U_2, \ldots, U_k$ di V sono indipendenti se da $\mathbf{u}_i \in U_i$ $(i = 1, 2, \ldots, k)$ e $\mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_k = \mathbf{0}$ segue che $\mathbf{u}_1 = \mathbf{u}_2 = \ldots = \mathbf{u}_k = \mathbf{0}$.

Esempio 2.3.15. Se $\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ è un insieme linearmente indipendente, allora i sottospazi $\langle \mathbf{v}_1 \rangle, \ldots, \langle \mathbf{v}_k \rangle$ sono indipendenti.

Osservazione 130. La definizione di somma di due sottospazi si estende facilmente a un numero qualunque di essi.

Proposizione 2.3.13. *Se i sottospazi $U_1, U_2, \ldots, U_k$ di V sono indipendenti, allora*

$$\dim(U_1 + U_2 + \ldots + U_k) = \dim U_1 + \dim U_2 + \ldots + \dim U_k \tag{2.21}$$

In tal caso se $\mathscr{B}_1, \mathscr{B}_2, \ldots, \mathscr{B}_k$ sono base di $U_1, U_2, \ldots, U_k$ rispettivamente, allora $\mathscr{B}_1 \sqcup \mathscr{B}_2 \sqcup \ldots \sqcup \mathscr{B}_k$ è una base di $U_1 + U_2 + \ldots + U_k$.

Dimostrazione. Per esercizio □

2.4 Applicazioni lineari

2.4.1 Introduzione

Osservazione 131. Applicazione è praticamente sinonimo di funzione.

Definizione 2.4.1 (Applicazione lineare)**.** Siano V e W spazi vettoriali e sia $f : V \to W$ un'applicazione tra V e W. La chiameremo lineare se:

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2) \qquad \mathbf{v}_1, \mathbf{v}_2 \in V \tag{2.22}$$

$$f(\alpha\mathbf{v}) = \alpha f(\mathbf{v}) \qquad \mathbf{v} \in V, \alpha \in \mathbb{C} \tag{2.23}$$

Esempio 2.4.1. Siano $V = W = M_2(\mathbb{C})$; si dica quali delle seguenti applicazioni $f_i : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ sono applicazioni lineari:

$$f_1(\mathbf{A}) = \mathbf{A}\mathbf{A}^T, \quad f_2(\mathbf{A}) = \mathbf{A} - \mathbf{I}_2, \quad f_3(\mathbf{A}) = 2\mathbf{A}, \quad f_4(\mathbf{A}) = \mathbf{A}^2$$

Per f_1 si ha che $f_1(\alpha\mathbf{A}) = \alpha\mathbf{A}\alpha\mathbf{A}^T = \alpha^2\mathbf{A}\mathbf{A}^T = \alpha^2 f_1(\mathbf{A}) \neq \alpha f_1(\mathbf{A})$ quindi f_1 non è applicazione lineare.
Per f_2, $f_2(\alpha\mathbf{A}) = \alpha\mathbf{A} - \alpha\mathbf{I}_2 = \alpha(\mathbf{A} - \mathbf{I}_2) = \alpha f_2(\mathbf{A})$; tuttavia se $f_2(\mathbf{B}) = \mathbf{B} - \mathbf{I}_2$, si ha che $f_2(\mathbf{A}) + f_2(\mathbf{B}) = \mathbf{A} + \mathbf{B} - 2\mathbf{I}_2 \neq f_2(\mathbf{A} + \mathbf{B}) = \mathbf{A} + \mathbf{B} - \mathbf{I}_2$ quindi f_2 non è lineare.
Per f_3, $f_3(\alpha\mathbf{A}) = 2\alpha\mathbf{A} = \alpha 2\mathbf{A} = \alpha f_3(\mathbf{A})$; inoltre $f(\mathbf{A} + \mathbf{B}) = 2(\mathbf{A} + \mathbf{B}) = 2\mathbf{A} + 2\mathbf{B} = f_3(\mathbf{A}) + f_3(\mathbf{B})$, per cui f_3 è lineare.
Infine per f_4, $f_4(\alpha A) = (\alpha\mathbf{A})^2 = \alpha^2\mathbf{A}^2 = \alpha^2 f_4(\mathbf{A}) \neq \alpha f_4(\mathbf{A})$. Pertanto f_4 non è lineare.

Esempio 2.4.2. Se λ è uno scalare, l'applicazione $V \to V$ definita da $\mathbf{v} \to \lambda\mathbf{v}$ è lineare.

Esempio 2.4.3 (Restrizione di f al sottospazio U)**.** Se $f : V \to W$ è una applicazione lineare e U è un sottospazio di V, allora l'applicazione $f|U : U \to W$ definita tramite $\mathbf{u} \to f(\mathbf{u})$ è lineare.

Esempio 2.4.4 (Scelta delle componenti $i_1, i_2, \ldots, i_d$)**.** Scegliamo d indici per un vettore $\mathbf{v} = [\alpha_1\, \alpha_2\, \ldots\, \alpha_n]^T \in \mathbb{C}^n$ di n elementi, tali che $1 \leq i_1 < i_2 < \ldots < i_d \leq n$. Se poniamo

$$f(\mathbf{v}) = \begin{bmatrix} \alpha_{i_1} \\ \alpha_{i_2} \\ \ldots \\ \alpha_{i_d} \end{bmatrix}$$

l'applicazione $f : \mathbb{C}^n \to \mathbb{C}^d$ così ottenuta è lineare e consiste nella scelta delle componenti corrispondenti agli indici scelti.

Esempio 2.4.5 (Applicazione nulla)**.** L'applicazione $f : V \to W$ definita ponendo $f(\mathbf{v}) = \mathbf{0}$ per ogni $\mathbf{v} \in V$ è lineare.

Proposizione 2.4.1. *Se $f : V \to W$ è un'applicazione lineare, allora $f(\mathbf{0}) = \mathbf{0}$.*

Dimostrazione. Poniamo $\mathbf{w} = f(\mathbf{0})$. Siccome $\mathbf{0} = \mathbf{0} + \mathbf{0}$ (stiamo usando il vettore nullo di V) abbiamo per la 2.22 che

$$\mathbf{w} = f(\mathbf{0}) = f(\mathbf{0} + \mathbf{0}) = f(\mathbf{0}) + f(\mathbf{0}) = \mathbf{w} + \mathbf{w}$$

da cui $\mathbf{w} = \mathbf{0}$ (il vettore nullo di W). □

Proposizione 2.4.2 (Verifica linearità di un'applicazione)**.** *Siano V e W spazi vettoriali e sia $f : V \to W$ un'applicazione di dominio V e codominio W. Allora f è lineare se e solo se, per ogni $\mathbf{v}_1, \mathbf{v}_2 \in V$ e per $\alpha_1, \alpha_2 \in \mathbb{C}$ si ha:*

$$f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) \tag{2.24}$$

Dimostrazione. Se f è lineare, basta applicare le condizioni 2.22 e 2.23. Supponiamo viceversa che per f valga la condizione data e mostriamo invece che valgono 2.22 e 2.23. Se prendiamo $\alpha_1 = \alpha_2 = 1$ abbiamo

$$f(\mathbf{v}_1 + \mathbf{v}_2) = f(1\mathbf{v}_1 + 1\mathbf{v}_2) = 1f(\mathbf{v}_1) + 1f(v_2) = f(\mathbf{v}_1) + f(\mathbf{v}_2)$$

che è la 2.22. Se prendiamo $\mathbf{v}_1 = \mathbf{v}$, $\mathbf{v}_2 = \mathbf{0}$, $\alpha_1 = \alpha$ e $\alpha_2 = 0$ abbiamo

$$f(\alpha\mathbf{v}) = f(\alpha\mathbf{v} + 0\mathbf{0}) = \alpha f(\mathbf{v}) + 0f(\mathbf{0}) = \alpha f(\mathbf{v})$$

che è la 2.23 □

Corollario 2.4.3. *Sia $f : V \to W$ un'applicazione lineare tra spazi vettoriali; siano $\mathbf{v}_i \in V$ e α_i scalari $(i = 1, \ldots, n)$. Allora*

$$f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n) = \alpha_1 f(\mathbf{v}_1) + \ldots + \alpha_n f(\mathbf{v}_n) \tag{2.25}$$

Dimostrazione. Si dimostra per induzione su n. Per il passo base $f(\alpha_1\mathbf{v}_1) = \alpha_1 f(\mathbf{v}_1)$ per le proprietà della funzione. Ipotizzando la relazione vera per $n-1$ si ha che:

$$f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_{n-1}\mathbf{v}_{n-1}) = \alpha_1 f(\mathbf{v}_1) + \ldots + \alpha_{n-1} f(\mathbf{v}_{n-1})$$

Ora se aggiungiamo ad entrambi i membri $\alpha_n f(\mathbf{v}_n)$ si ha:

$$f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_{n-1}\mathbf{v}_{n-1}) + \alpha_n f(\mathbf{v}_n) = \alpha_1 f(\mathbf{v}_1) + \ldots + \alpha_{n-1} f(\mathbf{v}_{n-1}) + \alpha_n f(\mathbf{v}_n)$$

ma dato che al primo membro f che è lineare, per le proprietà della definizione si ha che:

$$\begin{aligned} f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_{n-1}\mathbf{v}_{n-1}) + \alpha_n f(\mathbf{v}_n) &= f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_{n-1}\mathbf{v}_{n-1}) + f(\alpha_n\mathbf{v}_n) \\ &= f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n) \end{aligned}$$

e si conclude con

$$f(\alpha_1\mathbf{v}_1 + \ldots + \alpha_n\mathbf{v}_n) = \alpha_1 f(\mathbf{v}_1) + \ldots + \alpha_n f(\mathbf{v}_n)$$

□

Esempio 2.4.6 (Applicazione indotta da matrice $\mathbf{A}$)**.** Sia $\mathbf{A}$ una matrice $m \times n$ (da considerare come parametro); dato (come input) $\mathbf{v} \in \mathbb{C}^n$, considerato che $\mathbf{A}\mathbf{v} \in \mathbb{C}^m$, possiamo definire un'applicazione

$$\begin{aligned} &f_{\mathbf{A}} : \mathbb{C}^n \to \mathbb{C}^m \\ &\mathbf{v} \to \mathbf{A}\mathbf{v} \end{aligned}$$

che dato un vettore $\mathbf{v}$ lo restituisce pre-moltiplicato per $\mathbf{A}$.
La funzione è lineare (ciò segue facilmente dalle proprietà del prodotto di matrici), in quanto

$$\begin{aligned} &\mathbf{A}(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{A}\mathbf{v}_1 + \mathbf{A}\mathbf{v}_2 \\ &\mathbf{A}(\lambda\mathbf{v}) = \lambda\mathbf{A}\mathbf{v} \end{aligned}$$

Osservazione 132. Più avanti si vedrà come ogni applicazione lineare tra spazi vettoriali di dimensione finita può essere rappresentata mediante una opportuna applicazione indotta (scegliendo una opportuna matrice $\mathbf{A}$).

Esempio 2.4.7 (Scambio delle componenti)**.** Dato un vettore $\mathbf{x} \in \mathbb{C}^2, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ la funzione indotta dalla matrice $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, del tipo $f_{\mathbf{A}} : \mathbb{R}^2 \to \mathbb{R}^2$ è l'applicazione lineare che inverte gli elementi del vettore fornito:

$$f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$

Esempio 2.4.8 (Matrice di rotazione di un angolo α)**.** Dato un punto $\mathbf{x} \in \mathbb{C}^2$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, la funzione indotta dalla matrice $\mathbf{B} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$ restituisce le coordinate di $\mathbf{x}$ in un sistema cartesiano ottenuto facendo ruotare quello di partenza dell'angolo α:

$$f_{\mathbf{B}}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1\cos\alpha - x_2\sin\alpha \\ x_1\sin\alpha + x_2\cos\alpha \end{bmatrix}$$

Proposizione 2.4.4 (Composizione di applicazioni lineari è lineare)**.** *Siano U, V, W spazi vettoriali e siano $f : U \to V$, $g : V \to W$ applicazioni lineari. Allora $g \circ f : U \to W$ è un'applicazione lineare.*

Dimostrazione. Basta eseguire le verifiche. Abbiamo:

$$\begin{aligned} g \circ f(\mathbf{u}_1 + \mathbf{u}_2) &= g(f(\mathbf{u}_1 + \mathbf{u}_2)) = g(f(\mathbf{u}_1) + f(\mathbf{u}_2)) \\ &= g(f(\mathbf{u}_1)) + g(f(\mathbf{u}_2)) = g \circ f(\mathbf{u}_1) + g \circ f(\mathbf{u}_2) \end{aligned}$$

e questo prova 2.22. Analogamente:

$$g \circ f(\alpha\mathbf{u}) = g(f(\alpha\mathbf{u})) = g(\alpha f(\mathbf{u})) = \alpha(g(f(\mathbf{u}))) = \alpha(g \circ f(\mathbf{u}))$$

prova la 2.23. □

Esempio 2.4.9 (Composizione e applicazioni indotte da matrici)**.** Consideriamo le due applicazioni indotte dalle matrici $\mathbf{A}$, di tipo $m \times n$, e $\mathbf{B}$ di tipo $n \times p$:

$$f_{\mathbf{A}} : \mathbb{C}^n \to \mathbb{C}^m \quad f_{\mathbf{B}} : \mathbb{C}^p \to \mathbb{C}^n$$

La composizione $f_{\mathbf{A}} \circ f_{\mathbf{B}}$ ha senso ed è lineare (prima applico $f_{\mathbf{B}}$ per avere un vettore in $\mathbb{C}^n$, poi $f_{\mathbf{A}}$ per l'output in $\mathbb{C}^m$). Se $\mathbf{v} \in \mathbb{C}^p$:

$$f_{\mathbf{A}} \circ f_{\mathbf{B}}(\mathbf{v}) = f_{\mathbf{A}}(f_{\mathbf{B}}(\mathbf{v})) = f_{\mathbf{A}}(\mathbf{B}\mathbf{v}) = \mathbf{A}(\mathbf{B}\mathbf{v}) = (\mathbf{A}\mathbf{B})\mathbf{v} = f_{\mathbf{A}\mathbf{B}}(\mathbf{v})$$

Osservazione 133. Pertanto, nel caso di applicazioni lineari indotte da matrici, la composizione corrisponde al prodotto tra matrici. Questa è la motivazione profonda della definizione di prodotto di matrici che si è adottato.

2.4.2 Spazio nullo, immagine, iniettività

Definizione 2.4.2 (Spazio nullo $\mathrm{N}(f)$ di un'applicazione lineare)**.** Se $f : V \to W$ è un'applicazione lineare tra spazi vettoriali. Si definisce spazio nullo di f il sottoinsieme del dominio V:

$$\mathrm{N}(f) = \{\mathbf{v} \in V : f(\mathbf{v}) = \mathbf{0}\} \tag{2.26}$$

costituito dai vettori che, attraverso la funzione, risultano nel vettore nullo.

Proposizione 2.4.5. *$\mathrm{N}(f)$ è un sottospazio di V.*

Dimostrazione. Seguiamo i passi di 2.1.8.
Abbiamo già mostrato che $\mathbf{0} \in \mathrm{N}(f)$, infatti $f(\mathbf{0}) = \mathbf{0}$; pertanto $\mathrm{N}(f) \neq \emptyset$.
Inoltre $\mathbf{v}_1, \mathbf{v}_2 \in \mathrm{N}(f)$ e siano α_1, α_2 scalari. Allora

$$f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) = \alpha_1\mathbf{0} + \alpha_2\mathbf{0} = \mathbf{0}$$

s e dunque dato che $\mathbf{0} \in \mathrm{N}(f)$, concludiamo che $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 \in \mathrm{N}(f)$. □

Definizione 2.4.3 (Immagine $\mathrm{Im}(f)$ di un'applicazione lineare)**.** Se $f : V \to W$ è un'applicazione lineare tra spazi vettoriali, allora definiamo immagine di f come il sottoinsieme del codominio W composto dagli elementi che sono immagine di almeno un elemento del dominio:

$$\mathrm{Im}(f) = \{f(\mathbf{v}) : \mathbf{v} \in V\} \tag{2.27}$$

Proposizione 2.4.6. $\mathrm{Im}(f)$*è un sottospazio di W.*

Dimostrazione. Analogamente (2.1.8), è ovvio che $\mathrm{Im}(f)$ non sia vuota (basta che V non sia vuoto, direi).
Siano $\mathbf{w}_1, \mathbf{w}_2 \in \mathrm{Im}(f)$ definiti come $\mathbf{w}_1 = f(\mathbf{v}_1)$ e $\mathbf{w}_2 = f(\mathbf{v}_2)$ per opportuni $\mathbf{v}_1, \mathbf{v}_2 \in V$; siano $\alpha_1, \alpha_2 \in \mathbb{C}$. Allora possiamo scrivere:

$$\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2 = \alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2) = f(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) \in \mathrm{Im}(f)$$

□

Osservazione 134. Notiamo che per dimostrare che $\mathbf{w} \in \mathrm{Im}(f)$ basta riuscire a scrivere $\mathbf{w} = f(\mathbf{v})$ per un $\mathbf{v} \in V$.

Esempio 2.4.10. Consideriamo l'applicazione lineare di scelta della seconda, terza e quinta componente in $\mathbb{C}^5$ ossia la $f : \mathbb{C}^5 \to \mathbb{C}^3$ definita da:

$$f\left(\begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\alpha_4\\\alpha_5\end{bmatrix}\right) = \begin{bmatrix}\alpha_2\\\alpha_3\\\alpha_5\end{bmatrix}$$

Lo spazio nullo $\mathrm{N}(f)$ è l'insieme dei vettori con nulli gli elementi scelti, ovvero tutte le combinazioni lineari di $\mathbf{e}_1, \mathbf{e}_4$:

$$\mathrm{N}(f) = \left\{\begin{bmatrix}\alpha_1 & 0 & 0 & \alpha_4 & 0\end{bmatrix}^T,\ \alpha_1, \alpha_4 \in \mathbb{C}\right\} = \langle\mathbf{e}_1; \mathbf{e}_4\rangle$$

L'immagine, come si verifica facilmente, è $\mathrm{Im}(f) = \mathbb{C}^3$

Esempio 2.4.11. Sia $f : \mathbb{R}^3 \to \mathbb{R}^2$ è l'applicazione indotta dalla matrice

$$\begin{bmatrix} 1 & 2 & -1 \\ -2 & -4 & 2 \end{bmatrix}$$

ossia

$$f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 - x_3 \\ -2x_1 - 4x_2 + 2x_3 \end{bmatrix}$$

Si nota che ogni elemento di $\mathrm{Im}\,(f)$ è un multiplo scalare di $\mathbf{w} = [1\ -2]^T$ poiché:

$$\begin{bmatrix} x_1 + 2x_2 - x_3 \\ -2x_1 - 4x_2 + 2x_3 \end{bmatrix} = \underbrace{(x_1 + 2x_2 - x_3)}_{\text{costante}} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

quindi $\mathrm{Im}\,(f) = \langle \mathbf{w} \rangle$ (inoltre $\mathbf{w} = f(\mathbf{e}_1)$).
Un vettore $\mathbf{v} = [\alpha_1\, \alpha_2\, \alpha_3]^T$ appartiene a $\mathrm{N}\,(f)$ se e solo se $\alpha_1 + 2\alpha_2 - \alpha_3 = 0$, poiché in tal caso $f(\mathbf{v}) = \mathbf{0}$.

Proposizione 2.4.7 (Applicazione lineare iniettiva)**.** *Un'applicazione lineare $f : V \to W$ è iniettiva se e solo se* $\mathrm{N}\,(f) = \{\mathbf{0}\}$ *(solo il vettore nullo, attraverso la funzione, risulta nel vettore nullo).*

Dimostrazione. Supponiamo che f sia iniettiva e dimostriamo che solo $\mathbf{0} \in \mathrm{N}\,(f)$. Sia $\mathbf{v} \in \mathrm{N}\,(f)$: allora $f(\mathbf{v}) = \mathbf{0} = f(\mathbf{0})$ e quindi $\mathbf{v} = \mathbf{0}$, cioè l'unico elemento di $\mathrm{N}\,(f)$ è il vettore nullo.
Viceversa supponiamo che $\mathrm{N}\,(f) = \{\mathbf{0}\}$ e mostriamo che f è iniettiva. Consideriamo due elementi $\mathbf{v}_1, \mathbf{v}_2$ che abbiano la stessa immagine: $f(\mathbf{v}_1) = f(\mathbf{v}_2)$. Allora

$$\mathbf{0} = f(\mathbf{v}_1) - f(\mathbf{v}_2) = f(\mathbf{v}_1 - \mathbf{v}_2)$$

cioè $\mathbf{v}_1 - \mathbf{v}_2 \in \mathrm{N}\,(f)$ (ossia la differenza appartiene allo spazio nullo). Per ipotesi allora, $\mathbf{v}_1 - \mathbf{v}_2 = \mathbf{0}$ (dato che $\mathrm{N}\,(f)$ ha solo $\mathbf{0}$), quindi $\mathbf{v}_1 = \mathbf{v}_2$ e f è iniettiva. □

Osservazione 135 (Applicazione su un insieme di vettori)*.* Data un'applicazione lineare $f : V \to W$ e un insieme finito $\mathscr{A} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di vettori in V consideriamo l'insieme $f\,[\mathscr{A}]$ ottenuto valutando f nei vettori di $\mathscr{A}$, ossia:

$$f\,[\mathscr{A}] = \{f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)\} \tag{2.28}$$

Osservazione 136*.* Per convenzione poniamo $f\,[\emptyset] = \emptyset$.

2.4.3 Biettività e isomorfismo

Osservazione 137 (Applicazione lineare suriettiva)*.* Ricordiamo che una generica applicazione $g : V \to W$ è suriettiva se $\mathrm{Im}\,(f) = W$.

Definizione 2.4.4 (Applicazione lineare biettiva)**.** Applicazione lineare iniettiva e suriettiva.

Definizione 2.4.5 (Isomorfismo)**.** Un'applicazione lineare $f : V \to W$ si dice un isomorfismo se è biettiva, cioè iniettiva e suriettiva.

Definizione 2.4.6 (Spazi isomorfi)**.** Diciamo che gli spazi V e W sono isomorfi se esiste un isomorfismo $f : V \to W$ che li lega.

Proposizione 2.4.8. *Se un'applicazione lineare $f : V \to W$ è biettiva, anche l'inversa $f^{-1} : W \to V$ è lineare.*

Dimostrazione. Supponiamo $\mathbf{w}_1, \mathbf{w}_2 \in W$ e siano α_1 e α_2 scalari. Allora

$$\mathbf{w}_1 = f(\mathbf{v}_1) \quad \text{e} \quad \mathbf{w}_2 = f(\mathbf{v}_2)$$

ed essendo f biettiva si ha $\mathbf{v}_1 = f^{-1}(\mathbf{w}_1)$ e $\mathbf{v}_2 = f^{-1}(\mathbf{w}_2)$. Allora, riportando una combinazione lineare di elementi del codominio nel dominio si ha:

$$\begin{aligned} f^{-1}(\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2) &= f^{-1}(\alpha_1 f(\mathbf{v}_1) + \alpha_2 f(\mathbf{v}_2)) \\ &\overset{(1)}{=} f^{-1}(f(\alpha_1\mathbf{v}_1 + \alpha_2 f\mathbf{v}_2)) \\ &= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 \\ &= \alpha_1 f^{-1}(\mathbf{w}_1) + \alpha_2 f^{-1}(\mathbf{w}_2) \end{aligned}$$

(dove (1) è giustificata dal fatto che f è lineare) e quindi si conclude che anche f^{-1} è lineare. □

2.4.4 Applicazioni lineari, indipendenza e dimensione

Proposizione 2.4.9. *Sia $f : V \to W$ un'applicazione lineare e sia $\mathscr{A}$ un insieme finito di vettori di V:*

1. *se $f[\mathscr{A}]$ è linearmente indipendente, allora $\mathscr{A}$ è linearmente indipendente;*
2. *se f è iniettiva e $\mathscr{A}$ è linearmente indipendente, allora $f[\mathscr{A}]$ è linearmente indipendente.*

Dimostrazione. Entrambi i casi sono ovvi se $\mathscr{A}$ è vuoto. Se non è vuoto, sia $\mathscr{A} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.
Per il primo punto supponiamo che $f[\mathscr{A}]$ sia linearmente indipendente e che una combinazione lineare degli elementi di $\mathscr{A}$ sia nulla:

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$$

Mostriamo che da questo deriva che $\alpha_i = 0$ (ossia $\mathscr{A}$ è linearmente indipendente). Abbiamo in W:

$$\mathbf{0} = f(\mathbf{0}) = f\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right) = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i)$$

e per l'ipotesi che $f[\mathscr{A}]$ è linearmente indipendente, dal fatto che $\sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i) = 0$ si ha che $\alpha_1 = \ldots = \alpha_n = 0$.
Per il secondo punto supponiamo che sia:

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i)$$

nell'ipotesi che f sia iniettiva e $\mathscr{A}$ linearmente vogliamo mostrare che $\alpha_i = 0$ per concludere che $f[\mathscr{A}]$ è linearmente indipendente. Allora, dato che si tratta di funzione lineare:

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i) = f\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right)$$

e quindi $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i \in \mathrm{N}(f)$ (per la definizione di spazio nullo). Per ipotesi allora $\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = \mathbf{0}$ (dato che f è iniettiva) e dunque (dato che $\mathscr{A}$ è linearmente indipendente) $\alpha_1 = \ldots = \alpha_n = 0$. □

Teorema 2.4.10 (Nullità + rango)**.** *Sia $f : V \to W$ un'applicazione lineare tra spazi vettoriali e supponiamo che V sia finitamente generato. Allora:*

1. $\mathrm{Im}(f)$ *è finitamente generato*
2. *si ha che:*
$$\dim V = \dim \mathrm{N}(f) + \dim \mathrm{Im}(f) \tag{2.29}$$

Osservazione 138. Diremo nullità di f $\dim \mathrm{N}(f)$, rango di f $\dim \mathrm{Im}(f)$.

Dimostrazione. Nel caso in cui $\mathscr{A}$ sia l'insieme vuoto il teorema è di facile verifica; sia allora $\mathscr{A} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un insieme di generatori di V.

Per il primo punto, poniamo $\mathbf{w}_i = f(\mathbf{v}_i)$ $(i = 1, \ldots, n)$ la trasformazione dei generatori attraverso f. Se un generico $\mathbf{w} \in \mathrm{Im}(f)$, allora per definizione esiste un $\mathbf{v} \in V$ tale che $\mathbf{w} = f(\mathbf{v})$. Ora tal $\mathbf{v}$ si può scrivere combinazione lineare dei vettori di $\mathscr{A}$:
$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$$
Per il corollario 2.4.3 abbiamo:
$$\mathbf{w} = f(\mathbf{v}) = f\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right) = \sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i) = \sum_{i=1}^{n} \alpha_i \mathbf{w}_i$$
e dunque $f[\mathscr{A}] = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ è un insieme di generatori di $\mathrm{Im}(f)$ che quindi è finitamente generato (dato che è la trasformazione tramite f di un insieme finito $\mathscr{A}$).

Per il secondo punto facciamo vedere che, se $\dim \mathrm{N}(f) = n$ e $\dim \mathrm{Im}(f) = k$, allora V ha una base con $n + k$ elementi, quindi $\dim V = n + k = \dim \mathrm{N}(f) + \dim \mathrm{Im}(f)$. Le notazioni del seguito non hanno nessuna relazione con le precedenti.
Se $\dim \mathrm{Im}(f) = 0$ allora $\mathrm{Im}(f) = \{\mathbf{0}\}$ e quindi $V = \mathrm{N}(f)$, da cui $\dim V = \dim \mathrm{N}(f)$; la relazione da dimostrare è, in questo caso, valida.
Sia $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ una base di $\mathrm{Im}(f)$ (pertanto $\dim \mathrm{Im}(f) = k$); possiamo allora scrivere $\mathbf{w}_i = f(\mathbf{v}_i)$ $(i = 1, \ldots, k)$ per opportuni vettori $\mathbf{v}_i \in V$ (pertanto $\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ può essere vista come una $f[\mathscr{B}]$ linearmente indipendente, poiché base, e dalla 2.4.9, secondo punto, deriva che anche $\mathscr{B}$ è linearmente indipendente) e poniamo $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Poniamo ora attenzione su una base $\mathscr{A}$ di $\mathrm{N}(f)$:

- nel caso in cui $\mathscr{A} = \emptyset$, cioè $\mathrm{N}(f) = \{\mathbf{0}\}$ $(\dim \mathrm{N}(f) = 0)$, pertanto f è iniettiva. Dato che iniettiva, cardinalità di dominio e immagine coincidono e idem sarà[2] per i loro generatori, quindi $\dim V = k$; **TODO**: rivedere qui

[2] Per 2.4.9 $\dim V \leq k$ (componenti della base del dominio sono $\leq$ a quelli dell'immagine) perché insiemi linearmente indipendenti vanno in insiemi linearmente indipendenti. Ora l'insieme $\mathscr{B}$ è linearmente indipendente, per 2.4.9, perciò $\dim \mathrm{Im}(f) = k \leq \dim V$. In definitiva $\dim V = \dim \mathrm{Im}(f) = k$ come si voleva dimostrare, avendo $\dim \mathrm{N}(f) = 0$

- se $\mathscr{A} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ si ha che $\dim \mathrm{N}(f) = n$; dimostriamo che $\mathscr{C} = \mathscr{A} \sqcup \mathscr{B}$ è una base di V (da cui $\dim V = n+k$, per la 2.3.11, dato che $\mathscr{A} \cap \mathscr{B} = \emptyset$):

  1. l'insieme $\mathscr{C}$ è un insieme di generatori di V: se $\mathbf{v} \in V$, siccome $f(\mathbf{v}) \in \mathrm{Im}(f)$ possiamo scrivere $f(\mathbf{v})$ come combinazione lineare della base dell'immagine:
  $$f(\mathbf{v}) = \sum_{j=1}^{k} \beta_j \mathbf{w}_j$$
  Consideriamo il vettore
  $$\mathbf{v}' = \sum_{j=1}^{k} \beta_j \mathbf{v}_j$$
  ossia una combinazione lineare con gli stessi coefficienti di prima ma che fa uso dei vettori della base $\mathscr{B}$ (ossia appartenenti al dominio). In generale non potremo aspettarci che $\mathbf{v}' = \mathbf{v}$. Tuttavia
  $$\begin{aligned} f(\mathbf{v} - \mathbf{v}') = f(\mathbf{v}) - f(\mathbf{v}') &= \sum_{j=1}^{k} \beta_j \mathbf{w}_j - f\left(\sum_{j=1}^{k} \beta_j \mathbf{v}_j\right) \\ &= \sum_{j=1}^{k} \beta_j \mathbf{w}_j - \sum_{j=1}^{k} \beta_j f(\mathbf{v}_j) = \sum_{j=1}^{k} \beta_j \mathbf{w}_j - \sum_{j=1}^{k} \beta_j \mathbf{w}_j \\ &= \mathbf{0} \end{aligned}$$
  e se $f(\mathbf{v} - \mathbf{v}') = \mathbf{0}$, $\mathbf{v} - \mathbf{v}' \in \mathrm{N}(f)$. Dunque possiamo scrivere come combinazione lineare degli elementi della base di $\mathrm{N}(f)$:
  $$\mathbf{v} - \mathbf{v}' = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i$$
  da cui
  $$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \mathbf{v}' = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \sum_{j=1}^{k} \beta_j \mathbf{v}_j$$
  e quindi $\mathbf{v} \in \langle \mathscr{C} \rangle$ poiché combinazione lineare degli $\mathbf{u}_i \in \mathscr{A}$ e $\mathbf{v}_j \in \mathscr{B}$.

  2. l'insieme $\mathscr{C}$ è linearmente indipendente: supponiamo di avere
  $$\mathbf{0} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \sum_{j=1}^{k} \beta_j \mathbf{v}_j \tag{2.30}$$
  Allora
  $$\mathbf{0} \overset{(1)}{=} f(\mathbf{0}) \overset{(2)}{=} f\left(\sum_{i=1}^{n} \alpha_i \mathbf{u}_i + \sum_{j=1}^{k} \beta_j \mathbf{v}_j\right) = \sum_{i=1}^{n} \alpha_i \underbrace{f(\mathbf{u}_i)}_{=0, \mathbf{u}_i \in \mathrm{N}(f)} + \sum_{j=1}^{k} \beta_j \underbrace{f(\mathbf{v}_j)}_{=\mathbf{w}_j}$$
  $$= \sum_{j=1}^{k} \beta_j \mathbf{w}_j$$

dove la (1) deriva dalla 2.4.1, la (2) per sostituzione di 2.30. Pertanto abbiamo che $\sum_{j=1}^{k} \beta_j \mathbf{w}_j = \mathbf{0}$. Per l'indipendenza lineare di $f[\mathscr{B}]$ possiamo concludere che $\beta_1 = \ldots = \beta_k = 0$. Ma allora, riprendendo la 2.30:

$$\mathbf{0} = \sum_{i=1}^{n} \alpha_i \mathbf{u}_i$$

e per l'indipendenza lineare di $\mathscr{A}$ (base di $\mathrm{N}(f)$), abbiamo che $\alpha_1 = \ldots = \alpha_n = 0$ (e quindi sia gli α_i che i β_j sono nulli).

□

Corollario 2.4.11. *Se V è uno spazio vettoriale finitamente generato e $f : V \to V$ è lineare, allora f è iniettiva se e solo se è anche suriettiva (ossia se $\mathrm{Im}(f) = W$).*

Osservazione 139. Occhio al tipo di funzione: $f : V \to V$. È una funzione di un insieme in se stesso

Dimostrazione. Nel caso di f iniettiva si ha che $\mathrm{N}(f) = \{\mathbf{0}\}$ da cui $\dim \mathrm{N}(f) = 0$ poiché $\emptyset$ è base. Il teorema nullità + rango dice allora che $\dim V = 0 + \dim \mathrm{Im}(f)$ cioè che $\dim V = \dim \mathrm{Im}(f)$. Per la 2.3.10 allora che $\mathrm{Im}(f) = V$, e pertanto la funzione è anche suriettiva (poiché V è anche il codominio).
Viceversa, se f è suriettiva abbiamo $\mathrm{Im}(f) = V$ e quindi $\dim \mathrm{Im}(f) = \dim V$. Per il teorema nullità + rango, $\dim V = \dim \mathrm{N}(f) + \dim \mathrm{Im}(f)$ e quindi $\mathrm{N}(f) = 0$ (ossia è anche iniettiva). □

Osservazione 140. Attenzione che quest'ultimo corollario vale se dominio e codominio di f coincidono o più in generale, quando hanno la stessa dimensione. Infatti sia $f : V \to W$ con $\dim V \neq \dim W$; seguendo gli stessi passi della dimostrazione del corollario 2.4.11 si arriva a

$$\dim V = 0 + \dim \mathrm{Im}(f) \neq \dim W$$

quindi $W \neq \mathrm{Im}(f)$ e la funzione non è suriettiva.

2.4.5 Altri argomenti misti

Proposizione 2.4.12 (Immagine di applicazione indotta)**.** *Sia $f_{\mathbf{A}} : \mathbb{C}^n \to \mathbb{C}^m$ applicazione indotta $\mathbf{v} \to \mathbf{A}\mathbf{v}$ con $\mathbf{A}$ matrice $m \times n$ e $\mathbf{v} = [\alpha_1 \ldots \alpha_n]^T \in \mathbb{C}^n$; si ha che*

$$\mathrm{Im}(f_{\mathbf{A}}) = \mathrm{C}(\mathbf{A}) \tag{2.31}$$

Dimostrazione. Se scriviamo $A = [\mathbf{a}_1\, \mathbf{a}_2 \ldots \mathbf{a}_n]$ e prendiamo $\mathbf{v} = [\alpha_1 \ldots \alpha_n]^T \in \mathbb{C}^n$, abbiamo

$$f_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{v} = \alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$$

e dunque $f_{\mathbf{A}} \in \mathrm{C}(\mathbf{A})$, ossia un elemento dell'immagine appartiene allo spazio delle colonne $\mathrm{Im}(f_{\mathbf{A}}) \subseteq \mathrm{C}(\mathbf{A})$.
Viceversa, se un vettore $\mathbf{w} \in \mathrm{C}(\mathbf{A})$, esso si può scrivere come combinazione lineare dell'insieme delle colonne di $\mathbf{A}$, $\mathbf{w} = \alpha_1 \mathbf{a}_1 + \ldots + \alpha_n \mathbf{a}_n$ per opportuni α_i, ed è evidente che $\mathbf{w} = \mathbf{A}\mathbf{v} = f_{\mathbf{A}}(\mathbf{v})$ dove $\mathbf{v} = [\alpha_1 \ldots \alpha_n]^T \in \mathbb{C}^n$. Quindi un elemento dello spazio delle colonne appartiene all'immagine di $f_{\mathbf{A}}$, da cui $\mathrm{C}(\mathbf{A}) \subseteq \mathrm{Im}(f_{\mathbf{A}})$.
In definitiva $\mathrm{C}(\mathbf{A}) = \mathrm{Im}(f_{\mathbf{A}})$. □

Osservazione 141 (Immagine dell'applicazione indotta: definizione alternativa). In base a quest'ultima considerazione, $\mathrm{C}(\mathbf{A})$ è l'insieme dei vettori $\mathbf{b} \in \mathbb{C}^m$ per i quali il sistema lineare $\mathbf{Ax} = \mathbf{b}$ ha soluzione (e la soluzione $\mathbf{x}$ è il vettore con i coefficienti della combinazione lineare).

Definizione 2.4.7 (Spazio nullo di $f_{\mathbf{A}}$)**.** Lo spazio nullo di $f_{\mathbf{A}}$ è l'insieme dei vettori $\mathbf{v} \in \mathbb{C}^n$ tali che $\mathbf{Av} = \mathbf{0}$. Si può chiamare anche spazio nullo di $\mathbf{A}$ e si indica con $\mathrm{N}(\mathbf{A})$

Osservazione 142. Avendo definito come rango di una applicazione lineare f la dimensione di $\mathrm{Im}(f)$ (osservazione 138) e visto che per una matrice $\mathbf{A}$ risulta $\mathrm{Im}(f_{\mathbf{A}}) = \mathrm{C}(\mathbf{A})$, possiamo dare una nuova definizione di rango, in relazione alle applicazioni indotte da matrice.

Definizione 2.4.8 (Rango di $\mathbf{A}$, $\mathrm{rk}\,\mathbf{A}$)**.** Consiste nella dimensione dello spazio delle colonne di una matrice $\mathbf{A}$:

$$\mathrm{rk}\,\mathbf{A} = \dim \mathrm{Im}(f_{\mathbf{A}}) = \dim \mathrm{C}(\mathbf{A}) \tag{2.32}$$

Osservazione 143. Si vedrà che questa definizione coincide con quella associata alle righe (numero righe indipendenti) data in precedenza.

Corollario 2.4.13. *In una matrice* $\mathbf{A}$ *di tipo* $m \times n$ *il numero di colonne* n *coincide con la somma di rango e nullità.*

Dimostrazione. Si tratta di una applicazione del teorema nullità + rango alla funzione $f_{\mathbf{A}}$ □

In maxima per ottenere:

- la nullità di una matrice si usa la funzione `nullity`
- per il rango `rank`
- per il numero di colonne si usa `length` (numero delle righe) sulla trasposta (`transpose`) della matrice

```
##
## A:matrix([1,2,3],[4,5,6])
##                                   [ 1  2  3 ]
##                                   [         ]
##                                   [ 4  5  6 ]
## length(transpose(A))
##                                        3
## nullity(A)
##                                        1
## rank(A)
##                                        2
```

Proposizione 2.4.14 (Rango di composizione di funzioni lineari)**.** *Date* $f : U \to V$ *e* $g : V \to W$ *due generiche funzioni lineari (non necessariamente* *indotte da matrici), il rango di $g \circ f : U \to W$ è maggiorato dai ranghi delle singole funzioni;*

$$\dim \mathrm{Im}\,(g \circ f) \leq \dim \mathrm{Im}\,(f) \tag{2.33}$$
$$\dim \mathrm{Im}\,(g \circ f) \leq \dim \mathrm{Im}\,(g) \tag{2.34}$$

Dimostrazione. Possiamo applicare il teorema nullità + rango a $g \circ f$:

$$\dim U = \dim \mathrm{N}\,(g \circ f) + \dim \mathrm{Im}\,(g \circ f) \tag{2.35}$$

Si ha che

- se $f(\mathbf{u}) = \mathbf{0}$ allora anche $g \circ f(\mathbf{u}) = g(\mathbf{0}) = \mathbf{0}$. Per cui $\mathrm{N}\,(f) \subseteq \mathrm{N}\,(g \circ f)$ e
$$\dim \mathrm{N}\,(f) \leq \dim \mathrm{N}\,(g \circ f) \tag{2.36}$$
- inoltre dal teorema nullità+rango applicato alla sola f, per definizione
$$\dim U = \dim \mathrm{N}\,(f) + \dim \mathrm{Im}\,(f) \tag{2.37}$$

pertanto riscriviamo la 2.35 come

$$\dim \mathrm{N}\,(g \circ f) = \dim U - \dim \mathrm{Im}\,(g \circ f)$$

e la 2.37 come
$$\dim \mathrm{N}\,(f) = \dim U - \dim \mathrm{Im}\,(f)$$
Data la 2.36 si ha quindi che:

$$\dim U - \dim \mathrm{Im}\,(f) \leq \dim U - \dim \mathrm{Im}\,(g \circ f)$$

da cui sottraendo $\dim U$ ad entrambi i membri si conclude:

$$\dim \mathrm{Im}\,(g \circ f) \leq \dim \mathrm{Im}\,(f) \tag{2.38}$$

e quindi il rango di $g \circ f$ non supera il rango di f.
Specularmente avviene per g , facendo gli stessi passaggi, ma al posto di 2.37 si ha
$$\dim V = \dim \mathrm{N}\,(g) + \dim \mathrm{Im}\,(g)$$
da cui
$$\dim \mathrm{Im}\,(g \circ f) \leq \dim \mathrm{Im}\,(g) \tag{2.39}$$
(si poteva concludere allo stesso modo notando che $\mathrm{Im}\,(g \circ f) \subseteq \mathrm{Im}\,(g)$) e quindi il rango di $g \circ f$ non supera nemmeno il rango di g. □

Osservazione 144. Nel caso in cui le applicazioni lineari siano indotte da matrici, sappiamo (esempio 2.4.9) che $f_{\mathbf{B}} \circ f_{\mathbf{A}} = f_{\mathbf{BA}}$, quindi ne ricaviamo

$$\mathrm{rk}\,\mathbf{BA} \leq \mathrm{rk}\,\mathbf{A}$$
$$\mathrm{rk}\,\mathbf{BA} \leq \mathrm{rk}\,\mathbf{B}$$

Teorema 2.4.15. *Se $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ è una base dello spazio vettoriale V e $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ è un insieme di vettori nello spazio vettoriale W, allora esiste una e una sola applicazione lineare $f : V \to W$ tale che $f(\mathbf{v}_i) = \mathbf{w}_i$, $(i = 1, \ldots, n)$.*

Dimostrazione. Diamo solamente una idea. Se un'applicazione lineare f del tipo voluto esiste, essa è unica. Infatti, supponiamo che anche $g : V \to W$ abbia la proprietà richiesta, cioè che $f(\mathbf{v}_i) = \mathbf{w}_i = g(\mathbf{v}_i)$, $(i = 1, 2, \ldots, n)$. Se $\mathbf{v} \in V$, possiamo scrivere $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{v}_i$ in funzione della base, e dunque

$$g(\mathbf{v}) = \sum_{i=1}^n \alpha_i g(\mathbf{v}_i) = \sum_{i=1}^n \alpha_i \mathbf{w}_i = \sum_{i=1}^n \alpha_i f(\mathbf{v}_i) = f(\mathbf{v})$$

ossia coincide con f ed è unica.
Come si definisce allora f? Proprio sfruttando l'ultimo e il terzultimo membro della precedente:

$$f(\mathbf{v}) = \sum_{i=1}^n \alpha_i \mathbf{w}_i$$

e l'aspetto complicato della verifica è appunto controllare che questa posizione definisce un'applicazione lineare. Si tratta di un noioso controllo dei dettagli e lo omettiamo. □

2.5 I quattro sottospazi fondamentali di una matrice

Data una qualunque matrice $\mathbf{A}$ di dimensioni $m \times n$ si possono associare a essa quattro spazi vettoriali detti *sottospazi fondamentali*:

$\mathrm{C}(\mathbf{A})$ Spazio delle *colonne* di $\mathbf{A}$, definito come il sottospazio di $\mathbb{C}^m$ generato dai vettori colonna di $\mathbf{A}$

$\mathrm{N}(\mathbf{A})$ spazio *nullo* di $\mathbf{A}$ sottospazio di $\mathbb{C}^n$ formato dalle soluzioni del sistema omogeneo $\mathbf{Ax} = \mathbf{0}$

$\mathrm{C}\left(\mathbf{A}^H\right)$ Spazio delle *righe*, sottospazio di $\mathbb{C}^n$, coincidente con lo spazio delle colonne della H-trasposta di A

$\mathrm{N}\left(\mathbf{A}^H\right)$ spazio *nullo sinistro* di $\mathbf{A}$, coincidente con lo spazio nullo dell'H-trasposta, sottospazio di $\mathbb{C}^m$

Osservazione 145. Abbiamo legato lo spazio delle colonne di una matrice $\mathbf{A}$ alla soluzione di sistemi lineari (abbiamo detto nell'osservazione 141 che coincide con l'insieme dei vettori $\mathbf{b}$ per i quali il sistema $\mathbf{Ax} = \mathbf{b}$ ha soluzione).
Vediamo ora come il rango (dimensione dello spazio colonne) caratterizzi l'esistenza di inverse di $\mathbf{A}$ (e quindi sia anch'esso legato all'esistenza di soluzioni per il sistema). Diamo per buono quanto vedremo più avanti (ossia coincidenza delle definizioni di rango tra righe e colonne date).

Proposizione 2.5.1. *Sia* $\mathbf{A}$ *una matrice* $m \times n$*:*

- $\mathbf{A}$ *ha una inversa destra se e solo se* $\operatorname{rk}\mathbf{A} = m \leq n$ *(ossia se è orizzontale e inoltre il rango coincide con il numero di righe);*

- $\mathbf{A}$ *ha una inversa sinistra se e solo se* $\operatorname{rk}\mathbf{A}^H = n \leq m$ *(ossia se verticale e il rango dell'H-trasposta coincide con il numero di colonne).*

Dimostrazione. Rispettivamente:

- se $\mathbf{A}$ ha inversa destra $\mathbf{R}$ i sistemi $\mathbf{Ax} = \mathbf{e}_i$ $(i = 1, 2, \ldots, m)$ sono risolubili (con soluzione $\mathbf{x} = \mathbf{Re}_i$ dato che $\mathbf{ARe}_i = \mathbf{e}_i$) e dunque l'equazione è soddisfatta. Pertanto i vettori $\mathbf{e}_i$ appartengono allo spazio delle colonne $\mathrm{C}(\mathbf{A})$ (insieme dei vettori per i quali $\mathbf{Ax} = \mathbf{b}$ ha soluzione. Ossia $\mathrm{C}(\mathbf{A})$ contiene $\mathscr{E}$, una base di $\mathbb{C}^m$: questo avviene se e solo se $\mathrm{C}(\mathbf{A}) = \mathbb{C}^m$. Passando alle dimensioni di questi spazi:

TODO: non chiaro perché $\iff \mathrm{C}(\mathbf{A}) = \mathbb{C}^m$

$$\dim \mathrm{C}(\mathbf{A}) = \dim \mathbb{C}^m$$
$$\operatorname{rk} \mathbf{A} = m$$

  Dal fatto che il rango coincida con il numero di colonne dominanti (che possono essere $\leq n$) si ha che deve essere $\operatorname{rk} \mathbf{A} \leq n$ e quindi complessivamente che deve essere $\operatorname{rk} \mathbf{A} = m \leq n$;

- basta ricordare che $\mathbf{A}$ ha una inversa sinistra se e solo se $\mathbf{A}^H$ ha una inversa destra, visto che:

$$\mathbf{LA} = \mathbf{I} \iff (\mathbf{LA})^H = \mathbf{I}^H \iff \mathbf{A}^H \mathbf{L}^H = \mathbf{I}$$

□

Osservazione 146. Si vedrà nel teorema 2.5.9 che $\operatorname{rk} \mathbf{A} = \operatorname{rk} \mathbf{A}^H$: per cui una matrice non quadrata non può avere inversa destra *e* inversa sinistra (per cui le sole matrici quadrate sono invertibili tout court).

Corollario 2.5.2. *Una matrice quadrata* $\mathbf{A}$ *di ordine* n *è invertibile se e solo se* $\operatorname{rk} \mathbf{A} = n$

Osservazione 147. Rimane il problema di calcolare il rango di una matrice. Era stato definito come il numero di colonne dominanti (in una forma ridotta); quello che manca è la dimostrazione che questo numero è lo stesso, qualunque sia il procedimento di eliminazione seguito. Sappiamo comunque che per ogni matrice $\mathbf{A}$ esistono una matrice in forma ridotta $\mathbf{U}$ e una matrice invertibile $\mathbf{F}$ (una matrice di permutazione opportuna) tali che $\mathbf{FA} = \mathbf{U}$ (si vedano il lemma 1.6.2, punto 2 e il teorema 1.8.8) .

Osservazione 148 (Setup). Affrontiamo il problema più in generale:

- consideriamo due matrici $\mathbf{A}$ e $\mathbf{B}$ di dimensioni $m \times n$ tali che esista $\mathbf{F}$ invertibile di ordine m per la quale $\mathbf{FA} = \mathbf{B}$ (es una matrice di permutazione);
- scriviamo le colonne delle due matrici $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$ e $\mathbf{B} = [\mathbf{b}_1 \ldots \mathbf{b}_n]$;
- gli insiemi di vettori $\mathscr{A} = \{\mathbf{a}_1 \ldots \mathbf{a}_n\}$ e $\mathscr{B} = \{\mathbf{b}_1 \ldots \mathbf{b}_n\}$ sono insiemi di generatori di $\mathrm{C}(\mathbf{A})$ e $\mathrm{C}(\mathbf{B})$ rispettivamente: da essi perciò possiamo estrarre sottoinsiemi che sono basi (non tutte le colonne magari sono indipendenti);
- ricordiamo che si ha che $\mathbf{b}_i = \mathbf{Fa}_i$;

- vi è un link tra i due insiemi in quanto la pre-moltiplicazione per $\mathbf{F}$ di $\mathbf{a}_i$ da $\mathbf{b}_i$. In altre parole dato un sottoinsieme di uno dei due è chiaro quale sia il sottoinsieme corrispondente nell'altro. Per esempio il sottoinsieme corrispondente a $\{\mathbf{a}_1, \mathbf{a}_3\}$ è $\{\mathbf{b}_1, \mathbf{b}_3\}$.

Fino al teorema 2.5.4 ci poniamo in questa ipotesi.

Proposizione 2.5.3. *Nell'ipotesi* $\mathbf{FA} = \mathbf{B}$, *un vettore* $\mathbf{v} \in \mathbb{C}^m$ *è combinazione lineare dell'insieme delle colonne di* $\mathbf{A}$ *(*$\mathbf{v} \in \mathrm{C}(\mathbf{A})$*) se e solo se* $\mathbf{Fv}$ *è combinazione lineare dell'insieme delle colonne di* $\mathbf{B}$ *con gli stessi coefficienti.*

Dimostrazione. Se il vettore appartiene a $\mathrm{C}(\mathbf{A})$ può essere scritto come combinazione lineare delle colonne di $\mathbf{A}$: $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{a}_i$, dove abbiamo usato coefficienti α_i opportuni. Allora

$$\mathbf{Fv} = \mathbf{F} \sum_{i=1}^{n} \alpha_i \mathbf{a}_i = \sum_{i=1}^{n} \alpha_i \mathbf{F} \mathbf{a}_i = \sum_{i=1}^{n} \alpha_i \mathbf{b}_i$$

dove si conclude facendo uso degli stessi coefficienti α_i, ma delle colonne di $\mathbf{B}$. Viceversa, se $\mathbf{Fv} = \sum_{i=1}^{n} \alpha_i \mathbf{b}_i$ abbiamo (dato che $\mathbf{F}$ è invertibile per ipotesi) anche:

$$\mathbf{v} \overset{(1)}{=} \mathbf{F}^{-1}(\mathbf{Fv}) = \mathbf{F}^{-1}\left(\sum_{i=1}^{n} \alpha_i \mathbf{b}_i\right) = \sum_{i=1}^{n} \alpha_i \mathbf{F}^{-1} \mathbf{b}_i \overset{(2)}{=} \sum_{i=1}^{n} \alpha_i \mathbf{a}_i$$

dove (1) è una identità che sfrutta il fatto che $\mathbf{F}$ sia invertibile, e per (2) $\mathbf{F}^{-1}\mathbf{b} = \mathbf{a}_i$ perché da $\mathbf{b}_i = \mathbf{F}\mathbf{a}_i$ pre-moltiplicando per $\mathbf{F}^{-1}$ entrambi i membri si ottiene proprio $\mathbf{F}^{-1}\mathbf{b} = \mathbf{a}_i$. Quindi $\mathbf{v}$ è stata scritta come combinazione lineare delle colonne di $\mathbf{A}$ e appartiene dunque a $\mathrm{C}(\mathbf{A})$. □

Osservazione 149 (Applicazione della matrice $\mathbf{F}$ ad un subset delle colonne di $\mathbf{A}$). Scegliendo k indici corrispondenti ad opportune colonne di $\mathbf{A}$ e $\mathbf{B}$, con $1 \leq i_1 < \ldots < i_k \leq n$ siano $\mathbf{A}' = [\mathbf{a}_{i_1}\, \mathbf{a}_{i_2} \ldots \mathbf{a}_{i_k}]$ e $\mathbf{B}' = [\mathbf{b}_{i_1}\, \mathbf{b}_{i_2} \ldots \mathbf{b}_{i_k}]$ le matrici con le colonne selezionate. Si ha allo stesso modo che:

$$\mathbf{B}' = \mathbf{FA}'$$

dato che la colonna i-esima di $\mathbf{B}$ è data dalla pre-moltiplicazione della i-esima di $\mathbf{A}$ per $\mathbf{F}$. Pertanto possiamo applicare a questa situazione la proposizione 2.5.3.

Osservazione 150 (Caratteristiche dell'applicazione indotta da $\mathbf{F}$). La pre-moltiplicazione per $\mathbf{F}$ definisce un'applicazione f:

- lineare (2.4.6) del tipo $f : \mathrm{C}(\mathbf{A}) \to \mathrm{C}(\mathbf{B})$: ossia trasforma una colonna della prima matrice, appartenente comunque allo spazio colonne, in una della seconda, facente parte dello spazio colonne di questo;

- biettiva: infatti, nell'ipotesi che $\mathbf{FA} = \mathbf{B}$ con $\mathbf{F}$ invertibile, se pre-moltiplico entrambi i membri per $\mathbf{F}'$ si ha $\mathbf{A} = \mathbf{F}'\mathbf{B}$, quindi l'applicazione indotta da $\mathbf{F}'$ trasforma $\mathbf{B}$ in $\mathbf{A}$.
  Inoltre se f è definita come $\mathbf{A} \to \mathbf{FA}$ e g come $\mathbf{A} \to \mathbf{F}'\mathbf{A}$ si ha che $f \circ g$ è definita come $\mathbf{FF}'\mathbf{A} = \mathbf{A}$ ossia è l'identità, per cui g è inversa di f.

Osservazione 151 (Caratteristiche di $\mathbf{A}$ e $\mathbf{B}$, poste in relazione da f). Si ha che:

- $\mathrm{N}(\mathbf{A}) = \mathrm{N}(\mathbf{B})$. Infatti considerando un vettore $\mathbf{v} \in \mathrm{N}(\mathbf{A})$, ossia se $\mathbf{Av} = \mathbf{0}$, si ha $\mathbf{Bv} = \mathbf{FAv} = \mathbf{F0} = \mathbf{0}$, e quindi $\mathrm{N}(\mathbf{A}) \subseteq \mathrm{N}(\mathbf{B})$; analogamente il viceversa, usando che $\mathbf{A} = \mathbf{F}^{-1}\mathbf{B}$.

- $\operatorname{rk}\mathbf{A} = \operatorname{rk}\mathbf{B}$. Infatti, applicando il teorema nullità + rango alle funzioni di pre-moltiplicazione f : e g per A e B (entrambi $m \times n$, quindi sia f che g sono del tipo $\mathbb{C}^n \to \mathbb{C}^m$)

$$\begin{aligned} \dim \mathbb{C}^n &= \dim \mathrm{N}(f) + \dim \mathrm{Im}(f) \\ \dim \mathbb{C}^n &= \dim \mathrm{N}(g) + \dim \mathrm{Im}(g) \end{aligned}$$

  che si evolvono in

$$\begin{aligned} n &= \dim \mathrm{N}(\mathbf{A}) + \operatorname{rk}\mathbf{A} \\ n &= \dim \mathrm{N}(\mathbf{B}) + \operatorname{rk}\mathbf{B} \end{aligned}$$

  Dato poi che $\dim \mathrm{N}(\mathbf{A}) = \dim \mathrm{N}(\mathbf{B})$ allora si conclude che $\operatorname{rk}\mathbf{A} = \operatorname{rk}\mathbf{B}$.

Come trovare sottoinsiemi di colonne linearmente indipendenti di $\mathbf{A}$ una volta noti sottoinsiemi linearmente indipendenti di $\mathbf{B}$?

Teorema 2.5.4. *Sia* $\mathbf{B} = \mathbf{FA}$ *con* $\mathbf{F}$ *matrice invertibile. Un sottoinsieme di colonne di* $\mathbf{A}$*:*

1. *è linearmente indipendente se e solo se il corrispondente sottoinsieme di colonne di* $\mathbf{B}$ *lo è;*

2. *è una base di* $\mathrm{C}(\mathbf{A})$ *se e solo se il sottoinsieme corrispondente di colonne di* $\mathbf{B}$ *lo è (è una base di* $\mathrm{C}(\mathbf{B})$*)*

Dimostrazione. Rispettivamente:

1. se $\mathscr{A}$ è un sottoinsieme di colonne di $\mathbf{A}$ allora il sottoinsieme di colonne di $\mathbf{B}$ corrispondente è precisamente $f[\mathscr{A}]$ (per l'osservazione 149, poiché f trasforma colonne di $\mathbf{A}$ in colonne di $\mathbf{B}$), dove f è l'applicazione lineare $f : \mathrm{C}(\mathbf{A}) \to \mathrm{C}(\mathbf{B})$ indotta dalla moltiplicazione per $\mathbf{F}$.
   Siccome f è iniettiva (essendo biettiva), supponendo che $\mathscr{A}$ sia linearmente indipendente, abbiamo che $f[\mathscr{A}]$ è linearmente indipendente per la proposizione 2.4.9.
   Il viceversa si ottiene considerando l'applicazione lineare $g = f^{-1}$.

2. sia $\mathscr{A}$ una base di $\mathrm{C}(\mathbf{A})$ e abbia $k = \operatorname{rk}\mathbf{A}$ elementi. Il sottoinsieme corrispondente $f[\mathscr{B}]$ è linearmente indipendente (sempre per 2.4.9) e di ugual numerosità k. Ma dato che per le considerazioni dell'osservazione 151 $\operatorname{rk}\mathbf{A} = \operatorname{rk}\mathbf{B}$, si ha che $k = \operatorname{rk}\mathbf{B}$. Quindi essendo linearmente indipendente ed avendo lo stesso numero di elementi di una base di $\mathrm{C}(\mathbf{B})$ (definizione 2.4.8), esso è una base di $\mathrm{C}(\mathbf{B})$ per il teorema 2.3.9.
   Anche qui il viceversa si ottiene considerando l'applicazione inversa.

□

Osservazione 152 (Procedura per determinare una base di $\mathrm{C}(\mathbf{A})$). Occorre:

1. eseguire l'EG su $\mathbf{A}$ ottenendo una matrice in forma ridotta $\mathbf{U}$ e una matrice di permutazione invertibile $\mathbf{F}$ tali che $\mathbf{U} = \mathbf{FA}$
2. le colonne dominanti di $\mathbf{U}$ formano una base di $\mathrm{C}(\mathbf{U})$ (esempio 2.3.13)
3. le colonne di $\mathbf{A}$ corrispondenti alle colonne dominanti di $\mathbf{U}$ formano una base di $\mathrm{C}(\mathbf{A})$ (teorema 2.5.4)

Osservazione 153. È bene ricordare che, in generale, gli spazi delle colonne di $\mathbf{A}$ e $\mathbf{U}$ sono diversi; hanno però la stessa dimensione.

Proposizione 2.5.5. *Se* $\mathbf{B} = \mathbf{AG}$ *con* $\mathbf{G}$ *matrice invertibile; si ha che* $\mathrm{C}(\mathbf{B}) = \mathrm{C}(\mathbf{A})$.

Osservazione 154. Ossia a differenza di prima, post-moltiplico per la matrice di "trasformazione" invece di pre-moltiplicare.

Dimostrazione. Infatti, fissando $\mathbf{G} = [\gamma_{ij}]$ e scrivendo le colonne delle matrici $\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_n]$ $\mathbf{B} = [\mathbf{b}_1 \ldots \mathbf{b}_n]$:

- per il prodotto righe per colonne a blocchi si ha:

$$\mathbf{b}_i = \sum_{j=1}^{n} \mathbf{a}_j \gamma_{ji}, \quad (i = 1, \ldots, n)$$

  ossia la i-esima colonna di $\mathbf{AG}$ coincide con il prodotto di $\mathbf{A}$ per la i-esima colonna di $\mathbf{G}$. Pertanto ogni colonna di $\mathbf{B}$ è combinazione lineare dell'insieme delle colonne di $\mathbf{A}$, da cui $\mathrm{C}(\mathbf{B}) \subseteq \mathrm{C}(\mathbf{A})$.

- l'inclusione inversa discende dall'uguaglianza $\mathbf{A} = \mathbf{BG}^{-1}$, ottenuta post-moltiplicando $\mathbf{B} = \mathbf{AG}$ per $\mathbf{G}^{-1}$; anche se la matrice per la quale si post-moltiplica è cambiata ($\mathbf{G}^{-1}$ invece che $\mathbf{G}$) rimane il fatto che $\mathbf{A}$ è combinazione lineare delle colonne di $\mathbf{B}$, per cui $\mathrm{C}(\mathbf{A}) \subseteq \mathrm{C}(\mathbf{B})$.

Pertanto si conclude che $\mathrm{C}(\mathbf{A}) = \mathrm{C}(\mathbf{B})$. □

Teorema 2.5.6. *Sia* $\mathbf{B} = \mathbf{FA}$ *con* $\mathbf{F}$ *matrice invertibile. Allora* $\mathrm{C}\left(\mathbf{B}^H\right) = \mathrm{C}\left(\mathbf{A}^H\right)$.

Osservazione 155. Cioè gli spazi delle righe di $\mathbf{A}$ e $\mathbf{B}$ coincidono. In particolare ciò accade quando $\mathbf{B} = \mathbf{U}$, forma ridotta di $\mathbf{A}$ (ossia si ha $\mathrm{C}\left(\mathbf{A}^H\right) = \mathrm{C}\left(\mathbf{U}^H\right)$) ed $\mathbf{F}$ è la matrice di permutazione dell'EG.

Dimostrazione. L'uguaglianza $\mathbf{B} = \mathbf{FA}$ è equivalente (H-trasponendo entrambi i membri) a $\mathbf{B}^H = \mathbf{A}^H \mathbf{F}^H$. Da questo, dato che $\mathbf{F}^H$ rimane invertibile per le proprietà 1.6.13, applicando la 2.5.5, si ha che $\mathrm{C}\left(\mathbf{B}^H\right) = \mathrm{C}\left(\mathbf{A}^H\right)$. □

Osservazione 156. Spieghiamo perché $\mathrm{C}\left(\mathbf{A}^H\right)$ si chiama spazio delle righe di $\mathbf{A}$: ciò è dovuto al fatto che noi privilegiamo gli spazi $\mathbb{C}^n$ dei vettori colonna rispetto agli spazi $\mathbb{C}_n$ dei vettori riga, quindi pensiamo alle righe di $\mathbf{A}$ come alle colonne di $\mathbf{A}^T$. Questo esaurisce la spiegazione per le matrici reali, in quanto per queste $\mathbf{A}^T = \mathbf{A}^H$.
Nel caso di matrici complesse il modo di esprimere l'ortogonalità tra vettori di $\mathbb{C}^n$ (che si vedrà in seguito) forza a considerare $\mathbf{A}^H$ invece di $\mathbf{A}^T$. Questo

non crea alcun problema ai fini di considerazioni su dipendenza e indipendenza lineare e insiemi di generatori, quindi di dimensione. Infatti un insieme di righe di $\mathbf{A}$ è linearmente indipendente se e solo se le stesse righe di $\overline{\mathbf{A}}$ lo sono (basta coniugare tutto), quindi se e solo se le stesse colonne di $(\overline{\mathbf{A}})^T = \mathbf{A}^H$ lo sono.

Proposizione 2.5.7. *Sia* $\mathbf{U}$ *una forma ridotta della matrice* $\mathbf{A}$*. Una base di* $\mathrm{C}\left(\mathbf{U}^H\right)$ *(e quindi di* $\mathrm{C}\left(\mathbf{A}^H\right)$ *per l'osservazione 155) è data dalle colonne non nulle di* $\mathbf{U}^H$*.*

Dimostrazione. Basta provare che le colonne non nulle di $\mathbf{U}^H$ sono linearmente indipendenti, ovvero che le righe non nulle di $\mathbf{U}$ sono linearmente indipendenti. Ciò è speculare/analogo a quanto visto in 2.2.10 per le colonne (l'idea è che ci sarà un pivot di una riga che non permetterà che una data riga sia combinazione lineare delle rimanenti). □

Corollario 2.5.8. *Il numero di righe non nulle di una forma ridotta* $\mathbf{U}$ *della matrice* $\mathbf{A}$ *coincide con* $\dim \mathrm{C}\left(\mathbf{A}^H\right) = \mathrm{rk}\, \mathbf{A}^H$*.*

Dimostrazione. Dato che $\mathrm{C}\left(\mathbf{U}^H\right) = \mathrm{C}\left(\mathbf{A}^H\right)$ allora $\dim \mathrm{C}\left(\mathbf{U}^H\right) = \dim \mathrm{C}\left(\mathbf{A}^H\right) = \mathrm{rk}\, \mathbf{A}^H$ (ultima uguaglianza data dalla definizione 2.4.8). Per la 2.5.7 il $\dim \mathrm{C}\left(\mathbf{U}^H\right)$ coincide con il numero di colonne non nulle di $\mathbf{U}^H$ (ossia con il numero di righe non nulle di una forma ridotta $\mathbf{U}$). Quindi si conclude che il numero di righe non nulle di una forma ridotta $\mathbf{U}$ coincide con il rango di $\mathbf{A}^H$ □

Teorema 2.5.9. *Per ogni matrice* $\mathbf{A}$*, si ha* $\mathrm{rk}\, \mathbf{A} = \mathrm{rk}\, \mathbf{A}^H$*.*

Osservazione 157. Ossia in generale il numero di elementi di una base delle righe corrisponde a quello di una base delle colonne (H come forma generalizzata di trasposizione).

Dimostrazione. Sia $\mathbf{U}$ una forma ridotta di $\mathbf{A}$, $\mathbf{U} = \mathbf{FA}$ con $\mathbf{F}$ invertibile. Allora

$$\mathrm{rk}\, \mathbf{A} = \dim \mathrm{C}\left(\mathbf{A}\right) = \dim \mathrm{C}\left(\mathbf{U}\right) \overset{(1)}{=} \dim \mathrm{C}\left(\mathbf{U}^H\right) = \dim \mathrm{C}\left(\mathbf{A}^H\right) = \mathrm{rk}\, \mathbf{A}^H$$

dove (1) dipende dal fatto che il numero di colonne dominanti di $\mathbf{U}$ coincide con il numero delle sue righe non nulle. □

Definizione 2.5.1 (Spazio nullo sinistro)**.** Se $\mathbf{A}$ ha dimensioni $m \times n$ lo spazio nullo sinistro è dato dai vettori $\mathbf{w} \in \mathbb{C}^m$ tali che $\mathbf{A}^H\mathbf{w} = 0$ o equivalentemente (H-trasponendo entrambi i membri) $\mathbf{w}^H\mathbf{A} = 0$ (quest'ultima ne giustifica il nome).

Osservazione 158. Ci interessa valutare la dimensione dello spazio nullo sinistro.

Proposizione 2.5.10. *La dimensione dello spazio nullo di una matrice* $\mathbf{A}$ *di dimensioni* $m \times n$ *è* $n - \mathrm{rk}\, \mathbf{A}$*. La dimensione dello spazio nullo sinistro di* $\mathbf{A}$ *è* $m - \mathrm{rk}\, \mathbf{A}$*.*

Dimostrazione. Per definizione $\mathrm{N}\left(\mathbf{A}\right) = \mathrm{N}\left(f_{\mathbf{A}}\right)$ dove $f_{\mathbf{A}} : \mathbb{C}^n \to \mathbb{C}^m$. Dall'applicazione pedissequa del teorema nullità + rango si ha:

$$\begin{aligned} \dim \mathbb{C}^n &= \dim \mathrm{N}\left(f_{\mathbf{A}}\right) + \dim \mathrm{C}\left(f_{\mathbf{A}}\right) \\ n &= \dim \mathrm{N}\left(\mathbf{A}\right) + \mathrm{rk}\, \mathbf{A} \\ \dim \mathrm{N}\left(\mathbf{A}\right) &= n - \mathrm{rk}\, \mathbf{A} \end{aligned}$$

Analogamente $\mathrm{N}\left(\mathbf{A}^H\right) = \mathrm{N}\left(f_{\mathbf{A}^H}\right)$ dove $f_{\mathbf{A}^H} : \mathbb{C}^m \to \mathbb{C}^n$. E quindi

$$\begin{aligned}
\dim \mathbb{C}^m &= \dim \mathrm{N}\left(f_{\mathbf{A}^H}\right) + \dim \mathrm{C}\left(f_{\mathbf{A}^H}\right) \\
m &= \dim \mathrm{N}\left(\mathbf{A}^H\right) + \operatorname{rk} \mathbf{A}^H \\
m &= \dim \mathrm{N}\left(\mathbf{A}^H\right) + \operatorname{rk} \mathbf{A} \\
\dim \mathrm{N}\left(\mathbf{A}^H\right) &= m - \operatorname{rk} \mathbf{A}
\end{aligned}$$

□

Osservazione 159. Da quanto visto finora si può ricavare un importante risultato che lega tra di loro, a due a due, i quattro sottospazi fondamentali di una matrice; esso potrà essere anche migliorato in seguito quando avremo a disposizione la nozione di ortogonalità.

Teorema 2.5.11. *Data una matrice* $\mathbf{A}$ *di dimensioni* $m \times n$ *si ha*

$$\mathbb{C}^m = \mathrm{C}\left(\mathbf{A}\right) \oplus \mathrm{N}\left(\mathbf{A}^H\right) \tag{2.40}$$

$$\mathbb{C}^n = \mathrm{C}\left(\mathbf{A}^H\right) \oplus \mathrm{N}\left(\mathbf{A}\right) \tag{2.41}$$

Dimostrazione. È sufficiente verificare la prima uguaglianza; la seconda si ricava dalla prima sostituendo $\mathbf{A}$ con $\mathbf{A}^H$ (e notando che si stanno sommando effettivamente due vettori di $\mathbb{C}^n$).
Cominciamo dimostrando che $\mathrm{C}\left(\mathbf{A}\right) \cap \mathrm{N}\left(\mathbf{A}^H\right) = \{\mathbf{0}\}$, come presupposto da $\oplus$: se $\mathbf{y} \in \mathrm{C}\left(\mathbf{A}\right) \cap \mathrm{N}\left(\mathbf{A}^H\right)$, allora si ha che:

$$\begin{aligned}
\mathbf{y} &= \mathbf{A}\mathbf{v} \quad \text{per un vettore } \mathbf{v} \in \mathbb{C}^n \\
\mathbf{A}^H\mathbf{y} &= \mathbf{0}
\end{aligned}$$

Sostituendo la prima nella seconda si arriva a:

$$\mathbf{A}^H\mathbf{A}\mathbf{v} = \mathbf{0}$$

il che comporta che $\mathbf{y} = \mathbf{A}\mathbf{v} = \mathbf{0}$, come visto nella dimostrazione del teorema 1.6.10 (punto $3 \to 4$). Quindi abbiamo mostrato che l'elemento in comune è il vettore nullo.
Dalle preposizioni 2.3.11 e 2.5.10 ricaviamo:

$$\begin{aligned}
\dim\left[\mathrm{C}\left(\mathbf{A}\right) \oplus \mathrm{N}\left(\mathbf{A}^H\right)\right] &= \dim \mathrm{C}\left(\mathbf{A}\right) + \dim \mathrm{N}\left(\mathbf{A}^H\right) - \dim\{\mathbf{0}\} \\
&= \dim \mathrm{C}\left(\mathbf{A}\right) + \dim \mathrm{N}\left(\mathbf{A}^H\right) \\
&= \operatorname{rk} \mathbf{A} + (m - \operatorname{rk} \mathbf{A}) \\
&= m
\end{aligned}$$

e dato che $U = V \iff \dim U = \dim V$ (proposizione 2.3.10) possiamo concludere che $\mathbb{C}^m = \mathrm{C}\left(\mathbf{A}\right) \oplus \mathrm{N}\left(\mathbf{A}^H\right)$. □

Osservazione 160. Procediamo infine per il calcolo di una base di $\mathrm{N}\left(\mathbf{A}\right)$ di una matrice $\mathbf{A}$ $m \times n$.

Proposizione 2.5.12. *Lo spazio nullo di* $\mathbf{A}$ *coincide con quello di una sua forma ridotta* $\mathbf{U}$*, ossia* $\mathrm{N}\left(\mathbf{A}\right) = \mathrm{N}\left(\mathbf{U}\right)$ *.*

Dimostrazione. Prendendo un vettore $\mathbf{v}$ tale che $\mathbf{v} \in \mathrm{N}(\mathbf{A})$ si ha che $\mathbf{Av} = \mathbf{0}$. Ma dato che $\mathbf{U} = \mathbf{FA}$, con $\mathbf{F}$ invertibile, se post moltiplico per $\mathbf{v}$ si ottiene che $\mathbf{Uv} = \mathbf{FAv} = \mathbf{F0} = \mathbf{0}$. Quindi $\mathbf{Uv} = \mathbf{0}$ e quindi $\mathbf{v} \in \mathrm{N}(\mathbf{U})$. □

Osservazione 161 (Procedura per il calcolo di una base di $\mathrm{N}(\mathbf{A})$). Per la 2.5.12, i vettori dello spazio nullo di $\mathbf{A}$ sono le soluzioni del sistema omogeneo $\mathbf{Ux} = \mathbf{0}$: una soluzione di questo si ottiene dando valori arbitrari alle incognite non dominanti, ottenendo in corrispondenza i valori delle incognite dominanti.
Tra le diverse soluzioni ci basterà trovare $n - \mathrm{rk}\,\mathbf{A}$ (da 2.5.10) vettori linearmente indipendenti:

- ci saranno $d = n - \mathrm{rk}\,\mathbf{A}$ colonne non dominanti, se se le incognite ad esse corrispondenti sono $x_{i_1}, x_{i_2}, \dots, x_{i_d}$, con possiamo dare a ciascuna di esse, a turno, il valore 1 e alle altre il valore 0 (per comodità);
- adottando questa regola otterremo d vettori $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d$. Questo insieme di vettori è linearmente indipendente: infatti a essi corrispondono, nell'applicazione che sceglie le componenti $i_1, i_2, \dots, i_d$, i vettori della base canonica di $\mathbb{C}^d$ che formano un insieme linearmente indipendente.

Si noti che è particolarmente conveniente usare, nel calcolo della base dello spazio nullo, la forma ridotta ottenuta con l'eliminazione in avanti e all'indietro.

Esempio 2.5.1. Considerando la seguente matrice 3×4

$$\mathbf{A} = \begin{bmatrix} 1 & i & 6 & 8i \\ 2i & -2 & 0 & 8 \\ 1 & i & 6 & 8i \end{bmatrix}$$

Vogliamo determinare i quattro sottospazi fondamentali con relative basi.
Effettuiamo l'eliminazione di Gauss:

$$\begin{bmatrix} 1 & i & 6 & 8i \\ 2i & -2 & 0 & 8 \\ 1 & i & 6 & 8i \end{bmatrix} \begin{bmatrix} 1 & i & 6 & 8i \\ -2 & -2i & 0 & 8i \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & i & 6 & 8i \\ 0 & 0 & 12 & 24i \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & i & 6 & 8i \\ 0 & 0 & 1 & 2i \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

produce dunque la forma ridotta di $\mathbf{A}$:

$$\mathbf{U} = \begin{bmatrix} 1 & i & 6 & 8i \\ 0 & 0 & 1 & 2i \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Da quanto descritto nell'osservazione 152 si ricava che $\mathrm{C}(\mathbf{A})$ ha come base le colonne di $\mathbf{A}$ corrispondenti alle colonne dominanti di $\mathbf{U}$, quindi la prima e terza colonna di $\mathbf{A}$:

$$\mathrm{C}(\mathbf{A}) = \left\langle \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix} ; \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix} \right\rangle$$

Seguendo la procedura illustrata in osservazione 161, partendo dal sistema $\mathbf{Ux} = \mathbf{0}$ corrispondente alla matrice aumentata (e relativo sistema equivalente)

$$\begin{bmatrix} 1 & i & 6 & 8i & 0 \\ 0 & 0 & 1 & 2i & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{cases} x_1 + i \cdot x_2 + 6 \cdot x_3 + 8_i \cdot x_4 = 0 \\ x_3 + 2i \cdot x_4 = 0 \end{cases}$$

Ora:

- se $x_2 = 0$, $x_4 = 1$ si ha che

$$\begin{cases} x_3 + 2i = 0 \\ x_1 + 6 \cdot x_3 + 8i = 0 \end{cases} \qquad \begin{cases} x_3 = -2i \\ x_1 = -6 \cdot (-2i) - 8i = 4i \end{cases}$$

Quindi una prima soluzione è $\begin{bmatrix} 4i & 0 & -2i & 1 \end{bmatrix}^T$;

- se viceversa $x_1 = 1, x_4 = 0$ si ha

$$\begin{cases} x_3 = 0 \\ x_1 + i = 0 \end{cases} \qquad \begin{cases} x_3 = 0 \\ x_1 = i \end{cases}$$

per cui questa soluzione è $\begin{bmatrix} -i & 1 & 0 & 0 \end{bmatrix}^T$

TODO: Perché a loro viene per il primo vettore $\begin{bmatrix} -2i \\ 0 \\ -i \\ 1 \end{bmatrix}$?

Pertanto

$$\mathrm{N}(\mathbf{A}) = \mathrm{N}(\mathbf{U}) = \left\langle \begin{bmatrix} 4i \\ 0 \\ -2i \\ 1 \end{bmatrix}; \begin{bmatrix} -i \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\rangle$$

La proposizione 2.5.7 mostra che i due vettori della di base di $\mathrm{C}\left(\mathbf{A}^H\right) = \mathrm{C}\left(\mathbf{U}^H\right)$ sono i vettori H-trasposti delle due righe non nulle di $\mathbf{U}$, pertanto se:

$$\mathbf{U}^H = \begin{bmatrix} 1 & 0 & 0 \\ -i & 0 & 0 \\ 6 & 1 & 0 \\ -8 & -2i & 0 \end{bmatrix}$$

si ha che

$$\mathrm{C}\left(\mathbf{A}^H\right) = \mathrm{C}\left(\mathbf{U}^H\right) = \left\langle \begin{bmatrix} 1 \\ -i \\ 6 \\ -8i \end{bmatrix}; \begin{bmatrix} 0 \\ 0 \\ 1 \\ -2i \end{bmatrix} \right\rangle$$

Infine un altro facile calcolo mostra che:

$$\mathrm{N}\left(\mathbf{A}^H\right) = \left\langle \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\rangle$$

Il lettore verifichi che quanto previsto dal teorema 2.5.11 si verifica.

Esempio 2.5.2. Sia

$$\mathbf{A}_\alpha = \begin{bmatrix} 1 & -\alpha^2 & 1 & 1 \\ 1 & 3 & 2 & \alpha+1 \\ 3 & -3\alpha^2 & 3 & \alpha+1 \end{bmatrix}, \quad \alpha \in \mathbb{C}$$

Determinare per ogni $\alpha \in \mathbb{C}$ il rango di $\mathbf{A}_\alpha$; inoltre per ogni $\alpha \in \mathbb{C}$ trovare una base di $\mathrm{C}(\mathbf{A}_\alpha), \mathrm{N}(\mathbf{A}_\alpha)$ e $\mathrm{C}\left(\mathbf{A}_\alpha^H\right)$.

Procediamo alla riduzione di $\mathbf{A}_\alpha$:

$$\begin{bmatrix} 1 & -\alpha^2 & 1 & 1 \\ 1 & 3 & 2 & \alpha+1 \\ 3 & -3\alpha^2 & 3 & \alpha+1 \end{bmatrix} \qquad r_3 = r_3 - 3r_1$$

$$\begin{bmatrix} 1 & -\alpha^2 & 1 & 1 \\ 1 & 3 & 2 & \alpha+1 \\ 0 & 0 & 0 & \alpha-2 \end{bmatrix} \qquad r_2 = r_2 - r_1$$

$$\begin{bmatrix} 1 & -\alpha^2 & 1 & 1 \\ 0 & 3+\alpha^2 & 1 & \alpha \\ 0 & 0 & 0 & \alpha-2 \end{bmatrix}$$

Si ha che:

- la prima colonna è sempre dominante
- l'altra colonna dominante è la seconda o la terza, a sconda del valore assunto da α (se $\alpha = \pm i\sqrt{3}$ è dominante la terza, altrimenti la seconda)
- l'ultima colonna è dominante solo se $\alpha \neq 2$

Pertanto:

- se $\alpha \neq 2$, $\operatorname{rk} \mathbf{A}_\alpha = 3$
- se $\alpha = 2$, $\operatorname{rk} \mathbf{A}_\alpha = 2$

Per la base di $\mathrm{C}(\mathbf{A}_\alpha)$:

- se $\alpha = 2$, le prime due colonne, ossia

$$\mathrm{C}(\mathbf{A}_\alpha) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -\alpha^2 \\ 3 \\ -3\alpha^2 \end{bmatrix} \right\}$$

- se $\alpha = \pm i\sqrt{3}$ colonne 1,3 e 4

$$\mathrm{C}(\mathbf{A}_\alpha) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ \alpha+1 \\ \alpha+1 \end{bmatrix} \right\}$$

- altrimenti colonne 1, 2 e 4:

$$\mathrm{C}(\mathbf{A}_\alpha) = \left\{ \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -\alpha^2 \\ 3 \\ -3\alpha^2 \end{bmatrix}, \begin{bmatrix} 1 \\ \alpha+1 \\ \alpha+1 \end{bmatrix} \right\}$$

Ai fini della base di $\mathrm{N}(\mathbf{A}_\alpha)$:

$$\begin{bmatrix} 1 & -\alpha^2 & 1 & 1 & 0 \\ 0 & 3+\alpha^2 & 1 & \alpha & 0 \\ 0 & 0 & 0 & \alpha-2 & 0 \end{bmatrix} \begin{cases} x_1 - \alpha^2 x_2 + x_3 + x_4 = 0 \\ (3+\alpha^2)x_2 + \alpha = 0 \\ (\alpha-2)x_4 = 0 \end{cases}$$

pertanto:

- se $\alpha = 2$ sono dominanti la prima e seconda colonna e vi sono da scegliere i coefficienti di x_3 e x_4:
  - se $x_3 = 0, x_4 = 1$

$$\begin{cases} x_1 - \alpha^2 x_2 + 1 = 0 \\ (3+\alpha^2)x_2 + \alpha = 0 \\ \alpha = 2 \end{cases} \quad \begin{cases} x_2 = -\frac{\alpha}{3+\alpha^2} \\ x_1 = -1 - \frac{\alpha}{3+\alpha^2} \end{cases} \quad \begin{cases} x_1 = -\frac{15}{7} \\ x_2 = -\frac{2}{7} \end{cases}$$

  - se $x_3 = 1, x_4 = 0$:

$$\begin{cases} x_1 - \alpha^2 x_2 + 1 = 0 \\ (3+\alpha^2)x_2 + 1 = 0 \end{cases} \quad \begin{cases} x_2 = -\frac{1}{3+\alpha^2} \\ x_1 + \frac{\alpha^2}{3+\alpha^2} + 1 = 0 \end{cases} \quad \begin{cases} x_1 = -\frac{3+2\alpha^2}{3+\alpha^2} = -\frac{11}{7} \\ x_2 = -\frac{1}{3+\alpha^2} = -\frac{1}{7} \end{cases}$$

  Pertanto se $\alpha = 2$:

$$\mathrm{N}(\mathbf{A}_\alpha) = \left\langle \begin{bmatrix} -\frac{15}{7} \\ -\frac{2}{7} \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{11}{7} \\ -\frac{1}{7} \\ 1 \\ 0 \end{bmatrix} \right\rangle$$

- se $\alpha = \pm i\sqrt{3}$ è non dominante la seconda; ponendo $x_2 = 1$:

$$\begin{cases} x_1 - \alpha^2 + x_3 + x_4 = 0 \\ (3+\alpha^2) + x^3 + \alpha x^4 = 0 \\ (\alpha - 2)x_4 = 0 \end{cases} \quad \begin{cases} x_4 = 0 \\ x_3 = -3 - \alpha^2 \\ x_1 = \alpha^2 - 3 - \alpha^2 \end{cases} \quad \begin{cases} x_1 = -3 \\ x_3 = -3 - \alpha^2 \\ x_4 = 0 \end{cases} \quad \to \mathrm{N}(\mathbf{A}_\alpha) = \left\langle \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\rangle$$

- altrimenti è non dominante la terza, ponendo $x_3 = 1$:

$$\begin{cases} x_1 - \alpha^2 x_2 + 1 + x_4 = 0 \\ (3+\alpha^2)x_2 + 1 + \alpha x_4 = 0 \\ (\alpha - 2)x_4 = 0 \end{cases} \quad \begin{cases} x_4 = 0 \\ x_2 = -\frac{1}{3+\alpha^2} \\ x_1 = \frac{-\alpha^2}{3+\alpha^2} - 1 = \frac{-2\alpha^2-3}{3+\alpha^2} \end{cases} \quad \to \mathrm{N}(\mathbf{A}_\alpha) = \left\langle \begin{bmatrix} \frac{-2\alpha^2-3}{3+\alpha^2} \\ -\frac{1}{3+\alpha^2} \\ 1 \\ 0 \end{bmatrix} \right\rangle$$

Per $\mathrm{C}\left(\mathbf{A}_\alpha^H\right)$ si ha che $\mathrm{C}\left(\mathbf{A}_\alpha^H\right) = \mathrm{C}\left(\mathbf{U}_\alpha^H\right)$ e

$$\mathbf{U}_\alpha^H = \begin{bmatrix} 1 & 0 & 0 \\ -\alpha^2 & 3+\alpha^2 & 0 \\ 1 & 1 & 0 \\ 1 & \alpha & \alpha - 2 \end{bmatrix}$$

La seconda colonna è nulla se $\alpha = 0 \wedge \alpha = \pm i\sqrt{3}$, quindi mai; la terza colonna è nullase $\alpha = 2$. Pertanto

- se $\alpha = 2$

$$\mathrm{C}\left(\mathbf{A}_\alpha^H\right) = \left\langle \begin{bmatrix} 1 \\ -4 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \\ 1 \\ 2 \end{bmatrix} \right\rangle$$

- se $\alpha \neq 2$

$$\mathrm{C}\left(\mathbf{A}_\alpha^H\right) = \left\langle \begin{bmatrix} 1 \\ -\alpha^2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3+\alpha^2 \\ 1 \\ \alpha \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \alpha-2 \end{bmatrix} \right\rangle$$

In maxima si può trovare una base dello spazio nullo mediante `nullspace`:

```
##
## A:matrix([1,2,3],[4,5,6],[7,8,9])
## nullity(A)
##                                       1
## nullspace(A)
##                                     [ - 3 ]
##                                     [     ]
##                                span([  6  ])
##                                     [     ]
##                                     [ - 3 ]
```

2.5.1 Il teorema di Rouché-Capelli

Osservazione 162. I concetti di rango e di spazio delle colonne di una matrice permettono di dare una veste precisa alle questioni sulla risolubilità di un sistema lineare. Il teorema che enunceremo è in realtà già stato dimostrato nel capitolo sulle matrici in forma leggermente diversa.

Osservazione 163 (Risolubilità e indipendenza di vettori colonna). Quello che sappiamo sino ad ora:

- un sistema lineare di m equazioni in n incognite si può scrivere nella forma $\mathbf{Ax} = \mathbf{b}$;
- il sistema è risolubile se e solo se $\mathbf{b} \in \mathrm{C}(\mathbf{A})$ ossia se $\mathbf{b}$ è combinazione lineare dell'insieme delle prime n colonne della matrice aumentata $[\mathbf{A}|\mathbf{b}]$;
- dal teorema 2.5.4 (primo punto, alla luce del fatto che $\mathbf{U} = \mathbf{FA}$ per opportuna matrice di trasformazione $\mathbf{F}$) supponendo che $[\mathbf{U}|\mathbf{c}]$ sia una forma ridotta della matrice aumentata il sistema è risolubile se e solo se $\mathbf{c}$ è non dominante, e ciò accade se e solo se l'ultima colonna di $[\mathbf{A}|\mathbf{b}]$ è combinazione lineare dell'insieme delle colonne precedenti (se $\mathbf{c}$ fosse dominante $\mathbf{b}$ non sarebbe combinazione lineare delle colonne di $\mathbf{A}$, per cui $\mathbf{b} \notin \mathrm{C}(\mathbf{A})$, da cui l'irrisolvibilità del sistema).

Osservazione 164 (Risolubilità e rango). Un altro modo di vedere la cosa usa il concetto di rango. Chiaramente si ha che $\mathrm{C}(\mathbf{A}) \subseteq \mathrm{C}([\mathbf{A}|\mathbf{b}])$; dunque (verosimilmente per lemma 2.3.6 e per la proposizione 2.3.10) $\operatorname{rk}\mathbf{A} \leq \operatorname{rk}[\mathbf{A}|\mathbf{b}]$.
Di quanto può aumentare il rango della matrice aumentata? Se il rango di $\mathbf{A}$ è k, il rango di $[\mathbf{A}|\mathbf{b}]$ è al più $k+1$ (al massimo se la colonna aggiunta non è combinazione lineare delle rimanenti), quindi $\dim \mathrm{C}([\mathbf{A}|\mathbf{b}]) \leq k+1$.

Osservazione 165. Il sistema omogeneo associato a quello dato (ossia $\mathbf{Ax} = \mathbf{0}$) fornisce un'altra informazione utile.

Proposizione 2.5.13 (Somma di soluzioni del sistema e dell'omogeneo associato)**.** *Sia* $\mathbf{v}$ *soluzione di* $\mathbf{Ax} = \mathbf{b}$ *e* $\mathbf{u}$ *una soluzione del sistema omogeneo* $\mathbf{Ax} = \mathbf{0}$ *(ossia* $\mathbf{u} \in \mathrm{N}(\mathbf{A})$*). Abbiamo allora che la somma* $\mathbf{u}+\mathbf{v}$ *delle 2 soluzioni è ancora soluzione del sistema dato.*

Dimostrazione. Infatti:

$$\mathbf{A}(\mathbf{u} + \mathbf{v}) = \mathbf{Au} + \mathbf{Av} = \mathbf{0} + \mathbf{b} = \mathbf{b}$$

□

Proposizione 2.5.14 (Differenza di soluzioni del sistema)**.** *Siano* $\mathbf{v_1}, \mathbf{v_2}$ *soluzioni del sistema* $\mathbf{Ax} = \mathbf{b}$*; la differenza* $\mathbf{v}_1 - \mathbf{v}_2$ *delle due soluzioni è soluzione del sistema associato* $\mathbf{Ax} = \mathbf{0}$*, cioè* $\mathbf{v_1} - \mathbf{v}_2 \in \mathrm{N}(\mathbf{A})$.

Dimostrazione. Infatti:

$$\mathbf{A}(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{Av}_1 - \mathbf{Av}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

□

Osservazione 166. Dalla 2.5.13 ricaviamo che una volta nota *una soluzione particolare* del sistema, le altre si ottengono sommando a questa un elemento a turno dello spazio nullo.
Pertanto, per essere in grado di determinare tutte le soluzione basta trovarne una e avere a disposizione una base dello spazio nullo (la cui dimensione è precisamente $n - \operatorname{rk} \mathbf{A}$).
È a questo che ci si riferisce quando si dice che le soluzioni di un sistema lineare, se esistono, dipendono da un numero di parametri pari al numero delle incognite (n) meno il numero di incognite dominanti ($\operatorname{rk} \mathbf{A}$): questi parametri sono i coefficienti che possiamo assegnare agli elementi di una base dello spazio nullo per costruire una loro combinazione lineare.

Teorema 2.5.15 (Rouché-Capelli)**.** *Sia* $\mathbf{Ax} = \mathbf{b}$ *un sistema lineare di* m *equazioni in* n *incognite. Il sistema è risolubile se e solo se*

$$\operatorname{rk} \mathbf{A} = \operatorname{rk}[\mathbf{A}|\mathbf{b}]$$

(ossia l'ultima colonna non è dominante) e in tal caso le soluzioni del sistema dipendono da $n - \operatorname{rk} \mathbf{A}$ *parametri.*
In particolare il sistema ammette una e una sola soluzione se e solo se

$$\operatorname{rk} \mathbf{A} = \operatorname{rk}[\mathbf{A}|\mathbf{b}] = n$$

(ossia la soluzione dipende da 0 parametri).

Dimostrazione. Fatta a pezzi in precedenza. □

2.6 Coordinate e matrici associate alle applicazioni lineari

Osservazione 167. In questa sezione:

- supponiamo gli spazi vettoriali sia non nulli (dimensione > 0) che finitamente generati
- facciamo uso di *basi ordinate*, ossia basi in cui l'ordine in cui si considerano gli elementi è importante.

Definizione 2.6.1 (Base ordinata)**.** Base definita sia indicandone gli elementi che fissandone l'ordine in cui sono considerati.

2.6.1 Applicazione delle coordinate rispetto a una base

Definizione 2.6.2 (Applicazione delle coordinate rispetto alla base ordinata $\mathscr{B}$)**.** Applicazione $C_{\mathscr{B}} : V \to \mathbb{C}^n$ definita ponendo:

$$\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \text{se e solo se} \quad \mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i \tag{2.42}$$

ossia che dato un vettore restituisce le sue coordinate, da applicare alla base $\mathscr{B}$, necessarie per ottenerlo.
$\mathrm{C}_{\mathscr{B}}(\mathbf{v})$ si chiama *vettore delle coordinate di* $\mathbf{v}$ *rispetto alla base ordinata* $\mathscr{B}$.

Osservazione 168. È proprio il fatto di aver fissato l'ordine dei vettori di $\mathscr{B}$ a permetterci di dare questa definizione, ossia che ci permette di trasformare un vettore appartenente a V in uno del "più comune" $\mathbb{C}^n$.

Proposizione 2.6.1. *L'applicazione delle coordinate $C_{\mathscr{B}} : V \to \mathbb{C}^n$ è lineare e biettiva.*

Dimostrazione. Sia $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$:

- per la linearità: siano $\mathbf{u}, \mathbf{v} \in V$ e poniamo:

$$\mathrm{C}_{\mathscr{B}}(\mathbf{u}) = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}$$

Questo significa che

$$\mathbf{u} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i, \quad \mathbf{v} = \sum_{i=1}^{n} \beta_i \mathbf{v}_i$$

da cui

$$\mathbf{u} + \mathbf{v} = \sum_{i=1}^{n} (\alpha_i + \beta_i) \mathbf{v}_i$$

e quindi:

$$\mathrm{C}_{\mathscr{B}}(\mathbf{u}+\mathbf{v}) = \begin{bmatrix} \alpha_1+\beta_1 \\ \vdots \\ \alpha_n+\beta_n \end{bmatrix} = \mathrm{C}_{\mathscr{B}}(\mathbf{u}) + \mathrm{C}_{\mathscr{B}}(\mathbf{v})$$

Inoltre si ha che $\mathrm{C}_{\mathscr{B}}(\alpha\mathbf{u}) = \alpha\,\mathrm{C}_{\mathscr{B}}(\mathbf{u})$ per ogni $\mathbf{u} \in V$ e ogni scalare α. Infatti

$$\alpha\mathbf{u} = \alpha\sum_{i=1}^{n}\alpha_i\mathbf{v}_i = \sum_{i=1}^{n}\alpha\alpha_i\mathbf{v}_i$$

ed applicando $C_{\mathscr{B}}$ al primo e ultimo membro si ottiene:

$$\mathrm{C}_{\mathscr{B}}(\alpha\mathbf{u}) = \begin{bmatrix} \alpha\alpha_1 \\ \vdots \\ \alpha\alpha_n \end{bmatrix} = \alpha\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha\,\mathrm{C}_{\mathscr{B}}(\mathbf{u})$$

- per la biettività: $C_{\mathscr{B}}$ è chiaramente suriettiva (ad ogni vettore di coordinate di $\mathbb{C}^n$ può essere associato un vettore di V), per definizione:

$$\begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \mathrm{C}_{\mathscr{B}}\left(\sum_{i=1}^{n}\alpha_i\mathbf{v}_i\right)$$

  Inoltre ai fini dell'iniettività, per il teorema nullità + rango la dimensione dello spazio nullo di $C_{\mathscr{B}}$ è zero, in quanto:

$$\dim V = \dim \mathrm{Im}\,(C_{\mathscr{B}}) + \dim \mathrm{N}\,(C_{\mathscr{B}})$$

  $\dim V = n$ poiché $\mathscr{B}$ è composto da n elementi, $\dim \mathrm{Im}\,(C_{\mathscr{B}}) = n$ poiché i vettori di $\mathbb{C}^n$ hanno una base di n elementi. E $\dim \mathrm{N}\,(C_{\mathscr{B}}) = 0$ per differenza; ergo $\mathrm{N}\,(C_{\mathscr{B}}) = \{\mathbf{0}\}$ (e appunto la funzione è iniettiva).

□

Osservazione 169. L'applicazione delle coordinate rispetto a una base, proprio perché è lineare e biettiva, è lo strumento che rende possibile trattare ogni questione su spazi vettoriali attraverso matrici. Ad esempio:

- un insieme di vettori $\mathscr{A}$ in V è linearmente indipendente se e solo se l'insieme $C_{\mathscr{B}}[\mathscr{A}]$ è linearmente indipendente, e quest'ultimo è un insieme di vettori in $\mathbb{C}^n$ e dunque si può trattare con i metodi matriciali.
- un vettore $\mathbf{v} \in V$ è combinazione lineare dei vettori di $\mathscr{A}$ se e solo se $\mathrm{C}_{\mathscr{B}}(\mathbf{v})$ è combinazione lineare dei vettori di $C_{\mathscr{B}}[\mathscr{A}]$ (TODO esercizio).

TODO: fixme

Esempio 2.6.1. Il vettore delle coordinate di un vettore della base $\mathscr{B}$ è il corrispondente vettore della base canonica di $\mathbb{C}^n$.
Infatti nelle notazioni precedenti abbiamo che

$$\mathrm{C}_{\mathscr{B}}(\mathbf{v}_i) = \mathbf{e}_i \quad (i = 1, \ldots, n)$$

infatti possiamo scrivere

$$\mathbf{v}_i = 0\mathbf{v}_1 + \ldots + 1\mathbf{v}_i + \ldots + 0\mathbf{v}_n$$

e dunque

$$\mathrm{C}_{\mathscr{B}}(\mathbf{v}_i) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\text{-esimo elemento} = \mathbf{e}_i$$

Esempio 2.6.2. Se lo spazio vettoriale V è $\mathbb{C}^n$ e si adotta la base canonica $\mathscr{E}$, si ha che $\mathrm{C}_{\mathscr{E}}(\mathbf{v}) = \mathbf{v}$, ossia è l'applicazione *identità*.
Infatti, secondo la definizione, $\mathrm{C}_{\mathscr{E}}(\mathbf{v})$ è la colonna formata dai coefficienti α_i che servono a scrivere $\mathbf{v}$ come combinazione lineare degli elementi di $\mathscr{E}$. Ossia dato che:

$$\mathbf{v} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \ldots + \alpha_n\mathbf{e}_n \quad \text{allora} \quad \mathrm{C}_{\mathscr{E}}(\mathbf{v}) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}$$

da cui appunto $\mathrm{C}_{\mathscr{E}}(\mathbf{v}) = \mathbf{v}$.

Esempio 2.6.3. L'insieme $\mathscr{B} = \{2, 1-X, 1+X^2\}$ è una base di $\mathscr{P}_3(\mathbb{C})$ (verificarlo per esercizio). Vogliamo calcolare $C_{\mathscr{B}}(4-2X+3X^2)$. Per fare questo occorre determinare i coefficienti α, β, γ per i quali moltiplicare gli elementi della base; **TODO**: fixme

$$4 - 2X + 3X^2 = \alpha \cdot 2 + \beta(1-X) + \gamma(1+X^2) \tag{2.43}$$

Ora considerando che

$$\begin{aligned} 2 &= 2 + 0X + 0X^2 \\ 1 - X &= 1 - 1X + 0X^2 \\ 1 + X^2 &= 1 + 0X + 1X^2 \end{aligned}$$

conduce alla seguente matrice dei vettori della base

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

traducendo la 2.43 a sistema lineare in forma matriciale (sottendendo gli X^0, X^1, X^2) si ha:

$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$$

Risolvendo il sistema

$$\begin{bmatrix} 2 & 1 & 1 & 4 \\ 0 & -1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/2 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix} \to \begin{cases} \gamma = 3 \\ \beta = 2 \\ \alpha + \frac{1}{2}\beta + \frac{1}{2}\gamma = 2 \end{cases} \qquad \begin{cases} \alpha = -\frac{1}{2} \\ \beta = 2 \\ \gamma = 3 \end{cases}$$

Per cui si ottiene

$$C_{\mathscr{B}}(4 - 2X + 3X^2) = \begin{bmatrix} -1/2 & 2 & 3 \end{bmatrix}^T$$

2.6.2 Cambiamento di base

Osservazione 170. È ovvio che se $\mathscr{B}$ e $\mathscr{D}$ sono basi ordinate distinte di V, le applicazioni delle coordinate saranno diverse. Come calcolare le coordinate di un vettore rispetto a una base conoscendo quelle rispetto a un'altra base?

Teorema 2.6.2 (Matrice del cambiamento di base)**.** *Siano $\mathscr{B}$ e $\mathscr{D} = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ basi dello spazio vettoriale V. Allora esiste una e una sola matrice $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$ tale che, per ogni $\mathbf{v} \in V$:*

$$\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\,\mathrm{C}_{\mathscr{D}}(\mathbf{v})$$

La matrice $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$ si chiama matrice del cambiamento di base*: è invertibile e la sua inversa è $\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}}$.*

Dimostrazione. Assumiamo che $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$, di ordine $n = \dim V$, sia tale per cui:

$$\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\,\mathrm{C}_{\mathscr{D}}(\mathbf{v}) \qquad \forall \mathbf{v} \in V$$

Come costruire tale matrice? Dato che $\mathrm{C}_{\mathscr{D}}(\mathbf{w}_i) = \mathbf{e}_i$ (coordinate di un elemento della base) allora sotto la nostra ipotesi deve essere che le coordinate di un elemento $\mathbf{w}_i \in \mathscr{D}$, in termini della base $\mathscr{B}$ siano:

$$\mathrm{C}_{\mathscr{B}}(\mathbf{w}_i) = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\,\mathrm{C}_{\mathscr{D}}(\mathbf{w}_i) = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\mathbf{e}_i = i\text{-esima colonna di } \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$$

Ossia la i-esima colonna di $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$ corrisponde alle coordinate dell'i-esimo elemento della base $\mathscr{D}$ secondo la base $\mathscr{B}$. E quindi possiamo scrivere $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$ nel modo seguente:

$$\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}} = \begin{bmatrix}\mathrm{C}_{\mathscr{B}}(\mathbf{w}_1) & \mathrm{C}_{\mathscr{B}}(\mathbf{w}_2) & \dots & \mathrm{C}_{\mathscr{B}}(\mathbf{w}_n)\end{bmatrix}$$

Abbiamo provato l'*unicità* (la matrice è definita univocamente nella maniera appena vista); è facile concludere anche sull'effettiva *esistenza* (un vettore $\mathbf{w}_i \in V$ deve poter essere scritto come combinazione lineare di elementi della base $\mathscr{B}$ per la definizione stessa di base).
Verifichiamo che la matrice così ottenuta effettivamente *funziona* come voluto in termini di cambio delle coordinate. Sia $\mathbf{v} \in V$ e poniamo

$$\mathrm{C}_{\mathscr{D}}(\mathbf{v}) = \begin{bmatrix}\beta_1 & \beta_2 & \dots & \beta_n\end{bmatrix}^T \quad \text{per cui} \quad \mathbf{v} = \sum_{i=1}^{n} \beta_i \mathbf{w}_i$$

allora lo stesso vettore $\mathbf{v} \in V$, in termini di coordinate su $\mathscr{B}$, sarà definito come (applicando $C_{\mathscr{B}}$) ad entrambi i membri dell'ultima):

$$\begin{aligned}
\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathrm{C}_{\mathscr{B}}\left(\sum_{i=1}^{n}\beta_i\mathbf{w}_i\right) &= \sum_{i=1}^{n}\beta_i\,\mathrm{C}_{\mathscr{B}}(\mathbf{w}_i) = \sum_{i=1}^{n}\beta_i\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\mathbf{e}_i \\
&= \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\left(\sum_{i=1}^{n}\beta_i\mathbf{e}_i\right) = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\begin{bmatrix}\beta_1\\ \beta_2\\ \dots\\ \beta_n\end{bmatrix} \\
&= \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\,\mathrm{C}_{\mathscr{D}}(\mathbf{v})
\end{aligned}$$

e quindi $\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\,\mathrm{C}_{\mathscr{D}}(\mathbf{v})$ come si voleva dimostrare.
Alcuni punti rilevanti per la discussione a seguire:

1. ciò che abbiamo dimostrato vale per due basi qualunque. In particolare la matrice $\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}}$ (a basi invertite rispetto a $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$) esiste ed è unica;

2. per verificare che una generica $\mathbf{M} = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$, "basta" mostrare che, per ogni $\mathbf{v} \in V$ si ha $\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathbf{M}\,\mathrm{C}_{\mathscr{D}}(\mathbf{v})$ (ossia verificare la formula "banalmente");

3. non abbiamo in nessun punto usato che $\mathscr{B}$ e $\mathscr{D}$ siano diverse; quindi la matrice $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{B}}$ esiste ed è unica. Come sarà definita? Siccome:

$$\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathbf{I}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v})$$

per il punto 2 (ma anche per l'unicità della matrice $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{B}}$) si ha che $\mathbf{I} = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{B}}$.

Rimane ora da mostrare che $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$ è *invertibile*. Proviamo a eseguire il prodotto $(\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}})\,\mathrm{C}_{\mathscr{B}}(\mathbf{v})$:

$$\begin{aligned}(\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}})\,\mathrm{C}_{\mathscr{B}}(\mathbf{v}) &= \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}(\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v}))\\ &= \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\,\mathrm{C}_{\mathscr{D}}(\mathbf{v})\\ &= \mathrm{C}_{\mathscr{B}}(\mathbf{v})\end{aligned}$$

Pertanto

$$\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}} = \mathbf{I} = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{B}}$$

ossia $\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}}$ è inversa di $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}$. Viceversa, analogamente (a parti invertite per le basi):

$$\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{B}}\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}} = \mathbf{I}$$

□

Osservazione 171. Un caso particolare di queste matrici si ha quando $V = \mathbb{C}^n$ e una delle basi è la base canonica $\mathscr{E}$.
Vogliamo calcolare $\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}$ con $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un'altra base di $\mathbb{C}^n$: secondo la definizione dobbiamo calcolare $\mathrm{C}_{\mathscr{E}}(\mathbf{v}_i) = \mathbf{v}_i$ (esempio 2.6.2). Dunque:

$$\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}} = \begin{bmatrix}\mathbf{v}_1 & \mathbf{v}_2 & \ldots & \mathbf{v}_n\end{bmatrix}$$

Esempio 2.6.4. Verificare che, date due basi $\mathscr{B}, \mathscr{D}$ di $\mathbb{C}^n$, ed $\mathscr{E}$ base canonica, si ha:

$$\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}} = \mathbf{M}^{-1}_{\mathscr{E}\leftarrow\mathscr{B}}\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{D}}$$

TODO: fixme

Esempio 2.6.5. Più in generale, dimostrare che se $\mathscr{B}, \mathscr{D}, \mathscr{F}$ sono basi dello spazio vettoriale V si ha:

$$\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{F}} = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{D}}\mathbf{M}_{\mathscr{D}\leftarrow\mathscr{F}}$$

TODO: fixme

Esempio 2.6.6. Siano

$$\mathscr{B} = \left\{ \begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\\1\end{bmatrix} \right\}$$

$$\mathscr{B}' = \left\{ \begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\-1\end{bmatrix}, \begin{bmatrix}0\\2\\-1\\1\end{bmatrix}, \begin{bmatrix}0\\0\\-1\\0\end{bmatrix} \right\}$$

basi ordinate di $\mathbb{R}^4$. Calcolare la matrice $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{B}'}$.
La i-esima colonna di $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{B}'}$ corrisponde a $\mathrm{C}_{\mathscr{B}}(\mathbf{w}_i)$ con $\mathbf{w}_i$ elemento di $\mathscr{B}'$. Quindi bisogna risolvere tanti sistemi quanti sono gli elementi di $\mathscr{B}'$, o più efficientemente farne 1 e portarsi a dietro tutto (ossia applicare l'algoritmo di calcolo dell'inversa con la matrice di vettori di $\mathscr{B}'$ al posto di $\mathbf{I}$ sulla parte destra della matrice). Si ha:

$$\begin{bmatrix}1&1&1&1&1&0&0&0\\0&1&0&0&0&0&2&0\\0&0&1&1&1&1&-1&-1\\1&1&0&1&0&-1&1&0\end{bmatrix} \qquad r_4 = r_4 - r_1$$

$$\begin{bmatrix}1&1&1&1&1&0&0&0\\0&1&0&0&0&0&2&0\\0&0&1&1&1&1&-1&-1\\0&0&-1&0&-1&-1&1&0\end{bmatrix} \qquad r_4 = r_4 + r_3$$

$$\begin{bmatrix}1&1&1&1&1&0&0&0\\0&1&0&0&0&0&2&0\\0&0&1&1&1&1&-1&-1\\0&0&0&1&0&0&0&-1\end{bmatrix} \qquad r_3 = r_3 - r_4$$

$$\begin{bmatrix}1&1&1&1&1&0&0&0\\0&1&0&0&0&0&2&0\\0&0&1&0&1&1&-1&0\\0&0&0&1&0&0&0&-1\end{bmatrix} \qquad r_1 = r_1 - r_4$$

$$\begin{bmatrix}1&1&1&0&1&0&0&1\\0&1&0&0&0&0&2&0\\0&0&1&0&1&1&-1&0\\0&0&0&1&0&0&0&-1\end{bmatrix} \qquad r_1 = r_1 - r_3$$

$$\begin{bmatrix}1&1&0&0&0&-1&1&1\\0&1&0&0&0&0&2&0\\0&0&1&0&1&1&-1&0\\0&0&0&1&0&0&0&-1\end{bmatrix} \qquad r_1 = r_1 - r_2$$

$$\begin{bmatrix}1&0&0&0&0&-1&-1&1\\0&1&0&0&0&0&2&0\\0&0&1&0&1&1&-1&0\\0&0&0&1&0&0&0&-1\end{bmatrix}$$

Per cui la matrice richiesta è

$$\begin{bmatrix} 0 & -1 & -1 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Esempio 2.6.7. Si verifichi che

$$\mathscr{F} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}, \quad \mathscr{G} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

sono basi di $\mathbb{C}^3$ e si trovino le due matrici del cambiamento di base.
Facciamo la verifica per $\mathscr{F}$; effettuando la riduzione (i passaggi effettuati sono suggeriti dall'ultima colonna):

$$\begin{bmatrix} 1 & 1 & 0 & a \\ 1 & 0 & 1 & b \\ 0 & 1 & 1 & c \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & a \\ 0 & -1 & 1 & b-a \\ 0 & 1 & 1 & c \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & a \\ 0 & -1 & 1 & b-a \\ 0 & 0 & 2 & c+b-a \end{bmatrix}$$

Pertanto il sistema è determinato e nel caso fossero stati $a = b = c = 0$ la soluzione sarebbe stata $\mathbf{0}$. Allo stesso modo avviene per l'altra base:

$$\begin{bmatrix} 1 & 1 & 1 & a \\ 1 & 1 & 0 & b \\ 1 & 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & a \\ 0 & 1 & 0 & b-c \\ 1 & 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & a \\ 0 & 1 & 0 & b-c \\ 1 & 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & a-b+c \\ 0 & 1 & 0 & b-c \\ 1 & 0 & 0 & c \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & a-b \\ 0 & 1 & 0 & b-c \\ 1 & 0 & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & c \\ 0 & 1 & 0 & b-c \\ 0 & 0 & 1 & a-b \end{bmatrix}$$

Ora invece di portare avanti entrambi le matrici formate dagli elementi di $\mathscr{F}$ e $\mathscr{G}$ proseguiamo su quelle sviluppate nella verifica portando a identità la pare sinistra e sfruttando l'ultima colonna (le lettere) per il calcolo della matrice inversa. Partiamo dalla seconda che è già in tale forma: calcolando allora le tre righe applicando le espressioni con le lettere (dove a sarebbe la prima riga della matrice formata dagli elementi di $\mathscr{F}$, b la seconda e così via) si ottiene che

$$\mathbf{B}_{\mathscr{G} \leftarrow \mathscr{F}} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

Portiamo ad identità la parte sinistra della matrice cui siamo giunti per la prima:

$$\begin{bmatrix} 1 & 1 & 0 & a \\ 0 & -1 & 1 & b-a \\ 0 & 0 & 2 & c+b-a \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & a \\ 0 & 1 & -1 & a-b \\ 0 & 0 & 1 & (c+b-a)/2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & a \\ 0 & 1 & 0 & a-b+(c+b-a)/2 \\ 0 & 0 & 1 & (c+b-a)/2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & a-a+b-(c+b-a)/2 \\ 0 & 1 & 0 & a-b+(c+b-a)/2 \\ 0 & 0 & 1 & (c+b-a)/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & (2b-c-b+a)/2 \\ 0 & 1 & 0 & (2a-2b+c+b-a)/2 \\ 0 & 0 & 1 & (c+b-a)/2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & (b-c+a)/2 \\ 0 & 1 & 0 & (a-b+c)/2 \\ 0 & 0 & 1 & (c+b-a)/2 \end{bmatrix}$$

Ora applicando l'ultima colonna alle righe della matrice formata dai vettori della prima base si ha che

$$\mathbf{B}_{\mathscr{F} \leftarrow \mathscr{G}} = \begin{bmatrix} 1/2 & 1 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & -1/2 \end{bmatrix}$$

Osservazione 172. Come detto le coordinate permettono di trasferire questioni sugli spazi vettoriali a questioni su matrici. Il caso principale è quello delle applicazioni lineari; ora assoceremo a un'applicazione lineare una matrice.

2.6.3 Matrici associate ad applicazione lineare

Teorema 2.6.3 (Matrice associata a f)**.** *Siano V e W spazi vettoriali con $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ ($\dim V = n$) e $\mathscr{D}$ ($\dim W = m$) rispettive basi. Se $f : V \to W$ è un'applicazione lineare tra gli spazi, esiste una e una sola matrice $\mathbf{A}$ di tipo $m \times n$ tale che:*

$$\mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v})\right) = \mathbf{A}\,\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}\right) \qquad \forall \mathbf{v} \in V$$

La matrice $\mathbf{A}$ si chiama matrice associata *a f rispetto alla base $\mathscr{B}$ sul dominio e alla base $\mathscr{D}$ sul codominio.*

Osservazione 173. Ossia dove la pre-moltiplicazione traduce le coordinate in $\mathscr{B}$ di un vettore a quelle in $\mathscr{D}$ della sua immagine secondo f.

Dimostrazione. Per la dimostrazione sfruttiamo una tecnica analoga a quella adottata per le matrici del cambiamento di base. Nell'ipotesi che una matrice siffatta esista (ricordando che $\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}_i\right) = \mathbf{e}_i$) per un elemento $\mathbf{v}_i \in \mathscr{B}$ deve essere:

$$\mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v}_i)\right) = \mathbf{A}\,\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}_i\right) = \mathbf{A}\mathbf{e}_i = \ i\text{-esima colonna di } \mathbf{A}$$

Ossia la i-esima colonna di $\mathbf{A}$ corrisponde alle coordinate secondo $\mathscr{D}$ di f applicato all'i-esimo elemento della base $\mathscr{B}$. Pertanto complessivamente si ha:

$$\mathbf{A} = \begin{bmatrix} \mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v}_1)\right) & \mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v}_2)\right) & \ldots & \mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v}_n)\right) \end{bmatrix}$$

E quindi l'unicità è mostrata; l'esistenza deriva ancora dall'esistenza delle colonne. Facciamo vedere in seguito che questa matrice funziona come desiderato. Sia $\mathbf{v} \in V$ e scriviamolo in funzione degli elementi della base $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i$, ossia si ha che $\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}\right) = \begin{bmatrix} \alpha_1 & \ldots & \alpha_n \end{bmatrix}^T$. Allora, ricordando che sia f che $C_{\mathscr{D}}$ sono funzioni lineari:

$$\begin{aligned}
\mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v})\right) = \mathrm{C}_{\mathscr{D}}\left(f\left(\sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right)\right) &= \mathrm{C}_{\mathscr{D}}\left(\sum_{i=1}^{n} \alpha_i f(\mathbf{v}_i)\right) = \sum_{i=1}^{n} \alpha_i\,\mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v}_i)\right) \\
&= \sum_{i=1}^{n} \alpha_i \mathbf{A}\mathbf{e}_i = \mathbf{A}\left(\sum_{i=1}^{n} \alpha_i \mathbf{e}_i\right) \\
&= \mathbf{A}\,\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}\right)
\end{aligned}$$

come si voleva dimostrare. □

Esempio 2.6.8 (Matrice associata ad una applicazione $f : \mathbb{C}^n \to \mathbb{C}^m$)**.** Considerando una applicazione $f : \mathbb{C}^n \to \mathbb{C}^m$ possiamo calcolare la matrice associata a f considerando le basi canoniche $\mathscr{E}_n$ e $\mathscr{E}_m$. Per definizione sarà la matrice $\mathbf{A}$ di dimensioni $m \times n$ tale che:

$$\mathrm{C}_{\mathscr{E}_m}\left(f(\mathbf{v})\right) = \mathbf{A}\,\mathrm{C}_{\mathscr{E}_n}\left(\mathbf{v}\right), \qquad \forall \mathbf{v} \in \mathbb{C}^n$$

ma dato che, per le proprietà delle basi $\mathrm{C}_{\mathscr{E}_m}\left(f(\mathbf{v})\right) = f(\mathbf{v})$ e $\mathrm{C}_{\mathscr{E}_n}\left(\mathbf{v}\right) = \mathbf{v}$, si ha:

$$f(\mathbf{v}) = \mathbf{A}\mathbf{v}$$

Dunque qualsiasi applicazione lineare $\mathbb{C}^n \to \mathbb{C}^m$ è rappresentabile mediante pre-moltiplicazione per una opportuna matrice associata.

Esempio 2.6.9. Si consideri l'applicazione $f : \mathbb{R}^3 \to \mathbb{R}^3$ definita da $f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x+y \\ y+z \end{bmatrix}$:

- verificare che f è una trasformazione lineare
- trovare la matrice associata ad f rispetto alle basi canoniche
- siano $\mathscr{F} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ e $\mathscr{G} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$. Dopo aver verificato che $\mathscr{F}$ è una base per $\mathbb{R}^3$ e $\mathscr{G}$ è una base per $\mathbb{R}^2$ trovare la matrice associata ad f rispetto a $\mathscr{F}$ e $\mathscr{G}$.

Per il primo punto basta vedere che

$$f\left(\alpha \begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = f\left(\begin{bmatrix} \alpha x \\ \alpha y \\ \alpha z \end{bmatrix}\right) = \begin{bmatrix} \alpha x + \alpha y \\ \alpha y + \alpha z \end{bmatrix} = \alpha \begin{bmatrix} x+y \\ y+z \end{bmatrix} = \alpha f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right)$$

e che

$$\begin{aligned} f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) &= f\left(\begin{bmatrix} x+a \\ y+b \\ z+c \end{bmatrix}\right) = \begin{bmatrix} x+a+y+b \\ y+b+z+c \end{bmatrix} = \begin{bmatrix} x+y \\ y+z \end{bmatrix} + \begin{bmatrix} a+b \\ b+c \end{bmatrix} \\ &= f\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) + f\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) \end{aligned}$$

Per il secondo punto se A è la matrice associata a f per le basi canoniche si ha che $\mathrm{C}_{\mathscr{E}_2}(f(\mathbf{v})) = A \cdot \mathrm{C}_{\mathscr{E}_3}(\mathbf{v})$, e si vede velocemente che

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x+y \\ y+z \end{bmatrix}$$

per cui $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.
Per il terzo punto, verifichiamo che costituiscano basi: ciò avviene se i sistemi generati $\mathbf{Ax} = \mathbf{b}$ e $\mathbf{Ax} = \mathbf{0}$ hanno rispettivamente soluzioni e sono determinati con $\mathbf{0}$ soluzione. Per quanto riguarda $\mathscr{F}$:

$$\begin{bmatrix} 1 & 1 & 1 & a \\ 1 & 0 & 1 & b \\ 0 & 0 & 1 & c \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & a \\ 0 & -1 & 0 & b-a \\ 0 & 0 & 1 & c \end{bmatrix} \to \text{sistema determinato, ha soluzioni}$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \to \text{sistema determinato, unica soluzione } \mathbf{0}$$

Per $\mathscr{G}$ invece:

$$\begin{bmatrix} 1 & 1 & a \\ 1 & 0 & b \end{bmatrix} \begin{bmatrix} 1 & 1 & a \\ 0 & -1 & b-a \end{bmatrix} \to \text{sistema determinato, ha soluzioni}$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix} \to \text{sistema determinato, unica soluzione } \mathbf{0}$$

Per la matrice associata ad f rispetto ad $\mathscr{F}$ e $\mathscr{G}$ applichiamo f agli i-esimi elementi della base di $\mathbb{C}^3$:

$$f\left(\begin{bmatrix}1\\1\\0\end{bmatrix}\right)=\begin{bmatrix}2\\1\end{bmatrix},\quad f\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right)=\begin{bmatrix}1\\0\end{bmatrix},\quad f\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right)=\begin{bmatrix}2\\2\end{bmatrix}$$

e ora troviamo le coordinate secondo la base di $\mathbb{C}^2$ dei vettori ottenuti; lo facciamo risolvendo i tre sistemi in un colpo solo se affianchiamo le due matrici e portiamo all'identità la prima

$$\begin{bmatrix}1&1&2&1&2\\1&0&1&0&2\end{bmatrix} \qquad r_1=r_1-r_2$$

$$\begin{bmatrix}0&1&1&1&0\\1&0&1&0&2\end{bmatrix} \qquad \text{ora swappo}$$

$$\begin{bmatrix}1&0&1&0&2\\0&1&1&1&0\end{bmatrix}$$

Per cui la matrice richiesta è $\begin{bmatrix}1&0&2\\1&1&0\end{bmatrix}$

Esempio 2.6.10. Sia $f:\mathbb{C}^2\to\mathbb{C}^3$ la trasformazione lineare tale che

$$f\left(\begin{bmatrix}1\\1\end{bmatrix}\right)=\begin{bmatrix}1\\0\\0\end{bmatrix},\quad f\left(\begin{bmatrix}1\\-1\end{bmatrix}\right)=\begin{bmatrix}0\\0\\1\end{bmatrix}$$

- si trovi la matrice associata ad f rispetto alla base $\mathscr{F}=\left\{\begin{bmatrix}1\\1\end{bmatrix},\begin{bmatrix}1\\-1\end{bmatrix}\right\}$ sul dominio e alla base canonica sul codominio;

- si trovi la matrice associata ad f rispetto alla base canonica sul dominio e sul codominio

Per il primo punto siamo interessati ad $\mathbf{A}$ tale che $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v}))=\mathbf{A}\,\mathrm{C}_{\mathscr{F}}(\mathbf{v})$; se $\mathbf{w}_i\in\mathscr{F}$, dato che $\mathrm{C}_{\mathscr{F}}(\mathbf{w}_i)=\mathbf{e}_i$ si ha che $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{w}_i))=\mathbf{A}\mathbf{e}_i$ ossia la i-esima colonna di $\mathbf{A}$ sono le coordinate $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{w}_i))$, ma dato che $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{w}_i))=f(\mathbf{w}_i)$ ci limitiamo ad applicare la funzione agli elementi di $\mathscr{F}$ e a costruire la relativa matrice

$$\mathbf{A}=\begin{bmatrix}1&0\\0&0\\0&1\end{bmatrix}$$

Per il secondo punto vogliamo $\mathbf{A}$ tale che $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v}))=\mathbf{A}\,\mathrm{C}_{\mathscr{E}_2}(\mathbf{v})$ conoscendo $\mathbf{A}'=\begin{bmatrix}1&0\\0&0\\0&1\end{bmatrix}$ tale che $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v}))=\mathbf{A}'\,\mathrm{C}_{\mathscr{F}}(\mathbf{v})$; sappiamo che $\mathrm{C}_{\mathscr{F}}(f(\mathbf{v}))=\mathbf{M}_{\mathscr{F}\leftarrow\mathscr{E}}\,\mathrm{C}_{\mathscr{E}_2}(\mathbf{v})$ per cui $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v}))=\mathbf{A}'\mathbf{M}_{\mathscr{F}\leftarrow\mathscr{E}}\,\mathrm{C}_{\mathscr{E}_2}(\mathbf{v})$ e si ha che $\mathbf{A}=\mathbf{A}'\mathbf{M}_{\mathscr{F}\leftarrow\mathscr{E}}$.

Al fine del calcolo di $\mathbf{A}$ pertanto abbiamo bisogni di determinare $\mathbf{M}_{\mathscr{F}\leftarrow\mathscr{E}}$:

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{bmatrix} \qquad r_2 = r_2 - r_1$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & -2 & -1 & 1 \end{bmatrix} \qquad r_2 = r_2 \cdot (-1/2)$$

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1/2 & -1/2 \end{bmatrix} \qquad r_1 = r_1 - r_2$$

$$\begin{bmatrix} 1 & 0 & 1/2 & 1/2 \\ 0 & 1 & 1/2 & -1/2 \end{bmatrix}$$

per cui $\mathbf{M}_{\mathscr{F}\leftarrow\mathscr{E}} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$ pertanto

$$\mathbf{A} = \mathbf{A}'\mathbf{M}_{\mathscr{F}\leftarrow\mathscr{E}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 0 \\ 1/2 & -1/2 \end{bmatrix}$$

Esempio 2.6.11. Siano

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}, \quad \mathbf{A} = \begin{bmatrix} 2 & 4 & 4 & -2 \\ -1 & -2 & -1 & -1 \\ -1 & -2 & -3 & 1 \\ 1 & 2 & 5 & -7 \end{bmatrix}$$

con $\mathscr{B}$ una base di $\mathbb{C}^4$ e $\mathbf{A}$ la matrice associata ad $f : \mathbb{C}^4 \to \mathbb{C}^4$ rispetto a $\mathscr{B}$ su dominio e codominio. Si trovi la matrice associata ad f rispetto alla base canonica su dominio e codominio.
Conosciamo $\mathbf{A}$ tale che $\mathrm{C}_{\mathscr{B}}(f(\mathbf{v})) = \mathbf{A} \cdot \mathrm{C}_{\mathscr{B}}(\mathbf{v})$; vogliamo determinare $\mathbf{B}$ tale che $\mathrm{C}_{\mathscr{E}_4}(f(\mathbf{v})) = \mathbf{B} \cdot \mathrm{C}_{\mathscr{E}_4}(\mathbf{v})$; sappiamo però che

$$\begin{aligned} \mathrm{C}_{\mathscr{B}}(f(\mathbf{v})) &= \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}} \cdot \mathrm{C}_{\mathscr{E}_4}(f(\mathbf{v})) \\ \mathrm{C}_{\mathscr{B}}(\mathbf{v}) &= \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}} \cdot \mathrm{C}_{\mathscr{E}_4}(\mathbf{v}) \end{aligned}$$

per cui

$$\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}}\mathrm{C}_{\mathscr{E}_4}(f(\mathbf{v})) = \mathbf{A}\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}} \cdot \mathrm{C}_{\mathscr{E}_4}(\mathbf{v})$$

ma dato che $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}}$ è invertibile con inversa $\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}$ pre-moltiplico per l'inversa ottenendo che

$$\mathrm{C}_{\mathscr{E}_4}(f(\mathbf{v})) = \mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}\mathbf{A}\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}}\mathrm{C}_{\mathscr{E}_4}(\mathbf{v})$$

per cui $\mathbf{B} = \mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}\mathbf{A}\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}}$. Occorre allora determinare $\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}$ e $\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}}$. Per la prima si ha immediatamente che

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$$

per cui

$$\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Per la seconda

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad r_1 = r_1 - r_3$$

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad r_2 = r_2 - r_1$$

$$\begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \text{swappo}$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & -1 & 0 \end{bmatrix}$$

per cui

$$\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix}$$

Concludendo

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 4 & -2 \\ -1 & -2 & -1 & -1 \\ -1 & -2 & -3 & 1 \\ 1 & 2 & 5 & -7 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} -18 & 9 & 21 & 6 \\ -8 & 2 & 8 & 0 \\ -6 & 4 & 8 & 4 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

Osservazione 174. Vediamo un primo uso della matrice associata: il rango di un'applicazione lineare si può calcolare tramite il rango della sua matrice associata.

Proposizione 2.6.4. *Siano V e W spazi vettoriali, con basi $\mathscr{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ e $\mathscr{D}$ rispettivamente, e sia $f : V \to W$ un'applicazione lineare tra gli spazi. Se $\mathbf{A}$ è la matrice associata a f rispetto alle basi considerate, allora:*

$$\dim \mathrm{Im}\,(f) = \mathrm{rk}\,\mathbf{A}$$

Di conseguenza $\dim \mathrm{N}\,(f) = \dim V - \mathrm{rk}\,\mathbf{A}$.

Dimostrazione. Sia $k = \operatorname{rk}\mathbf{A}$; allora sappiamo trovare un insieme di k colonne di $\mathbf{A}$ linearmente indipendenti (colonne corrispondenti alle dominanti nella forma ridotta); per non complicare le notazioni supponiamo che l'insieme sia quello delle prime k colonne $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Allora l'insieme:

$$\{\mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_1)), \mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_2)), \ldots \mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_k))\} \tag{2.44}$$

è linearmente indipendente poiché:

- dato che $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ sono linearmente indipendenti (fanno parte di una base) lo sono anche $\{\mathrm{C}_{\mathscr{B}}(\mathbf{v}_1), \ldots, \mathrm{C}_{\mathscr{B}}(\mathbf{v}_k)\}$ in quanto sono ottenuti tramite una biiezione (iniettiva) alla luce della della 2.4.9, punto 2
- ma dato che $\mathrm{C}_{\mathscr{B}}(\mathbf{v}_i) = \mathbf{e}_i$, si ha che $\mathbf{A}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v}_i) = \mathbf{A}\mathbf{e}_i$ che corrisponde all'i-esima colonna di $\mathbf{A}$. Quindi anche l'insieme delle prime k colonne di $\mathbf{A}$, scritto come:
$$\{\mathbf{A}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v}_1), \ldots, \mathbf{A}\,\mathrm{C}_{B}(\mathbf{v}_k)\}$$
è linearmente indipendente;
- ma notando che $\mathbf{A}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v}_i) = \mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_i))$ si conclude che
$$\{\mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_1)), \mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_2)), \ldots \mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_k))\}$$
è linearmente indipendente

Ma allora anche l'insieme:

$$\{f(\mathbf{v}_1), f(\mathbf{v}_2), \ldots, f(\mathbf{v}_k)\}$$

è linearmente indipendente (sempre per biezione con l'insieme della 2.44. Siccome è un insieme di vettori in $\operatorname{Im}(f)$ ne segue che $\dim \operatorname{Im}(f) \leq k = \operatorname{rk}\mathbf{A}$ per il teorema 2.3.7 (k numero di generatori indipendenti) e $\dim \operatorname{Im}(f) \leq \operatorname{rk}\mathbf{A}$.
La disuguaglianza inversa ($\dim \operatorname{Im}(f) \geq \operatorname{rk}\mathbf{A}$, dalla quale conclude $\dim \operatorname{Im}(f) = \operatorname{rk}\mathbf{A}$) si dimostra in modo analogo. Siccome $f[\mathscr{B}]$ è un insieme di generatori di $\operatorname{Im}(f)$ (per quanto visto nella dimostrazione di teorema 2.4.10), da essa possiamo estrarre una base (togliendo vettori non indipendenti). Possiamo allora riordinare i vettori di $\mathscr{B}$ in modo che la base estratta sia $\{f(\mathbf{v}_1), \ldots, f(\mathbf{v}_k)\}$ e ne deduciamo che $\{\mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_1)), \ldots, \mathrm{C}_{\mathscr{D}}(f(\mathbf{v}_k))\}$ è un insieme linearmente indipendente in $\mathbb{C}^n$ (poiché generato da funzione biettiva). Questi vettori sono k' colonne di $\mathbf{A}$, quindi $\dim \operatorname{Im}(f) = k' \leq \operatorname{rk}\mathbf{A}$ (la prima uguaglianza è dovuta al fatto che k' è il numero di elementi di una base di $\operatorname{Im}(f)$, la seconda disuguaglianza dal fatto che vi possono essere altre colonne indipendenti di $\mathbf{A}$ oltre alle k') pertanto $\dim \operatorname{Im}(f) \leq \operatorname{rk}\mathbf{A}$ come si voleva. □

TODO: ma qui non si voleva dimostrare che $\dim \operatorname{Im}(f) \geq \operatorname{rk}\mathbf{A}$?

Esempio 2.6.12. Sia $f : M_2(\mathbb{R}) \to \mathbb{R}^2$ l'applicazione lineare (verificarlo nel caso) definita da $f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a+b-c \\ a+b \end{bmatrix}$:

- calcolare la matrice A associata ad f rispetto alle basi, ordinatamente di $M_2(\mathbb{R})$ e $\mathbb{R}^2$:

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\}, \qquad \mathscr{D} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

- determinare $\mathrm{N}(f)$ e trovarne una base
- calcolare $\dim \mathrm{Im}(f)$ e provare che $\dim \mathrm{Im}(f) = \mathrm{rk}\, \mathbf{A}$, dove A è la matrice determinata al primo punto.

Per il primo punto applico innanzitutto f agli elementi della base del dominio

$$f\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad f\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad f\left(\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad f\left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Ora trovo le coordinate di questi elementi secondo la base di R^2 risolvendo i seguenti sistemi multipli

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 & 1 \\ -1 & 1 & 2 & 0 & 1 & 2 \end{bmatrix} \qquad r_2 = r_2 + r_1$$

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 4 & 0 & 2 & 3 \end{bmatrix}$$

Pertanto la matrice richiesta è $\begin{bmatrix} 2 & 0 & 1 & 1 \\ 4 & 0 & 2 & 3 \end{bmatrix}$.
Per il secondo punto, lo spazio nullo è definito come

$$\mathrm{N}(f) = \left\{ x \in M_2(\mathbb{R}) : \begin{bmatrix} a+b-c \\ a+b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

ossia

$$\begin{cases} a+b-c = 0 \\ a+b = 0 \end{cases} \quad \begin{cases} c = 0 \\ a = -b \end{cases} \quad \begin{cases} a = -k \\ b = k \\ c = 0 \\ d = \text{qualsiasi} \end{cases}$$

Tute le matrici con questa caratteristica appartengono allo spazio nullo. Per rappresentarle

$$k \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} + d \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

per cui lo spazio nullo è di dimensione 2 ed una base è l'insieme delle due matrici utilizzate per generare la combinazione lineare.
Infine si ha che $\dim f = 4$ (le matrici 2×2 hanno una base a 4 elementi) e applicando il teorema nullità + resto si ha che $\dim \mathrm{Im}(f) = 2$; e di fatti se si effettua la riduzione $\begin{bmatrix} 2 & 0 & 1 & 1 \\ 4 & 0 & 2 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ le colonne dominanti (o le righe non nulle) sono 2 per cui il rango è 2.

Esempio 2.6.13. Sia $f_\alpha : \mathbb{C}^3 \to \mathbb{C}^3$ la trasformazione lineare che ha $\mathbf{A}_\alpha = \begin{bmatrix} 1 & 1 & \alpha+1 \\ 0 & \alpha & 2\alpha \\ 1 & 1 & \alpha+1 \end{bmatrix}$ $(\alpha \in \mathbb{C})$ matrice associata rispetto alla base canonica sul dominio e alla base $\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ sul codominio:

- si trovi la matrice $\mathbf{A}'_\alpha$ di f_α rispetto alla base canonica su dominio e codominio

- al variare di $\alpha \in \mathbb{C}$ si trovi il rango di f_α e la dimensione di $\mathrm{N}(f_\alpha)$
- trovare per quali $\alpha \in \mathbb{C}$ il vettore $\mathbf{v} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \in \mathrm{Im}(f_\alpha)$

Per il primo punto abbiamo che $\mathrm{C}_{\mathscr{B}}(f(\mathbf{v})) = A_\alpha\, \mathrm{C}_{\mathscr{E}_3}(\mathbf{v})$ e vogliamo trovare $\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v})) = \mathbf{A}'_\alpha\, \mathrm{C}_{\mathscr{E}_3}(\mathbf{v})$. Sapendo che $\mathrm{C}_{\mathscr{B}}(f(\mathbf{v})) = \mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}}\, \mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v}))$ si ha che

$$\mathbf{M}_{\mathscr{B}\leftarrow\mathscr{E}}\, \mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v})) = \mathbf{A}_\alpha\, \mathrm{C}_{\mathscr{E}_3}(\mathbf{v})$$

per cui pre moltiplicando per $\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}$ si ha

$$\mathrm{C}_{\mathscr{E}_3}(f(\mathbf{v})) = \mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}\mathbf{A}_\alpha\, \mathrm{C}_{\mathscr{E}_3}(\mathbf{v})$$

dalla quale si conclude che $\mathbf{A}'_\alpha = \mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}\mathbf{A}_\alpha$. Troviamo $\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}$

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

per cui $\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ e concludendo

$$\mathbf{A}'_\alpha = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & \alpha+1 \\ 0 & \alpha & 2\alpha \\ 1 & 1 & \alpha+1 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 2\alpha+2 \\ 1 & \alpha+1 & 3\alpha+1 \\ 1 & 1 & \alpha+1 \end{bmatrix}$$

Per il secondo punto, dato che $\mathbf{A}'_\alpha$ è una combinazione lineare delle righe di $\mathbf{A}_\alpha$ il rango deve essere lo stesso e possiamo usarle entrambe; usando $\mathbf{A}_\alpha$

$$\begin{bmatrix} 1 & 1 & \alpha+1 \\ 0 & \alpha & 2\alpha \\ 1 & 1 & \alpha+1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \alpha+1 \\ 0 & \alpha & 2\alpha \\ 0 & 0 & 0 \end{bmatrix}$$

si ha che:

- se $\alpha = 0$ il rango $= 1$ (solo prima colonna dominante) per cui applicando il thm nullità $+$ rango $\dim \mathrm{N}(f) = 2$
- se $\alpha \neq 0$, $\mathrm{rk}\, f = 2$ (prima e seconda colonna dominante) da cui $\dim \mathrm{N}(f) = 1$

Infine per il terzo punto direi andando ad istinto credo di dover utilizzare $\mathbf{A}'_\alpha$ a sto giro; si ha:

$$\begin{bmatrix} 2 & 2 & 2\alpha+2 & 2 \\ 1 & \alpha+1 & 3\alpha+1 & 0 \\ 1 & 1 & \alpha+1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & 2\alpha+2 & 2 \\ 1 & \alpha+1 & 3\alpha+1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 2 & 2\alpha+2 & 2 \\ 0 & 2\alpha+2-2 & 6\alpha+2-2\alpha-2 & -2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & \alpha+1 & 1 \\ 0 & 2\alpha & 4\alpha & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 & \alpha+1 & 1 \\ 0 & \alpha & 2\alpha & -1 \end{bmatrix}$$

Nell'ultima vi sono soluzioni se $\alpha \neq 0$

Osservazione 175. Talvolta è data la matrice associata a un'applicazione lineare rispetto a basi che non sono quelle che ci interessano; quale sarà la relazione fra due matrici associate?

Teorema 2.6.5. *Sia $f : V \to W$ un'applicazione lineare e:*

- $\mathscr{B}$ e $\mathscr{B}'$ *due basi di V;*
- $\mathscr{D}$ e $\mathscr{D}'$ *due basi di W;*
- $\mathbf{A}$ *la matrice associata a f rispetto alle basi $\mathscr{B}$ e $\mathscr{D}$;*
- $\mathbf{A}'$ *la matrice associata a f rispetto alle basi $\mathscr{B}'$ e $\mathscr{D}'$.*

Allora $\mathbf{A}$ e $\mathbf{A}'$ sono legate dalla relazione:

$$\mathbf{A} = \mathbf{M}_{\mathscr{D} \leftarrow \mathscr{D}'} \mathbf{A}' \mathbf{M}_{\mathscr{B} \leftarrow \mathscr{B}'}^{-1} \tag{2.45}$$

Dimostrazione. Assumiamo di conoscere $\mathbf{A}'$ e cerchiamo di calcolare $\mathbf{A}$, usando ovviamente le matrici dei cambiamenti di base. Ci serve una matrice $\mathbf{A}$ con la proprietà che, per ogni $\mathbf{v} \in V$:

$$\mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v})\right) = \mathbf{A}\,\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}\right)$$

Abbiamo a disposizione le seguenti uguaglianze, valide per ogni $\mathbf{v} \in V$ e $\mathbf{w} \in W$:

1. $\mathrm{C}_{\mathscr{D}'}\left(f(\mathbf{v})\right) = \mathbf{A}'\,\mathrm{C}_{\mathscr{B}'}\left(\mathbf{v}\right)$
2. $\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}\right) = \mathbf{M}_{\mathscr{B} \leftarrow \mathscr{B}'}\,\mathrm{C}_{\mathscr{B}'}\left(\mathbf{v}\right)$
3. $\mathrm{C}_{\mathscr{D}}\left(\mathbf{w}\right) = \mathbf{M}_{\mathscr{D} \leftarrow \mathscr{D}'}\,\mathrm{C}_{\mathscr{D}'}\left(\mathbf{w}\right)$

Abbiamo allora, indicando sopra il segno di uguaglianza quale relazione stiamo impiegando

$$\begin{aligned}\mathrm{C}_{\mathscr{D}}\left(f(\mathbf{v})\right) &\overset{(3)}{=} \mathbf{M}_{\mathscr{D} \leftarrow \mathscr{D}'}\,\mathrm{C}_{\mathscr{D}'}\left(f(\mathbf{v})\right) \overset{(1)}{=} \mathbf{M}_{\mathscr{D} \leftarrow \mathscr{D}'}\mathbf{A}'\,\mathrm{C}_{\mathscr{B}'}\left(\mathbf{v}\right) \\ &\overset{(2)}{=} \mathbf{M}_{\mathscr{D} \leftarrow \mathscr{D}'}\mathbf{A}'\mathbf{M}_{\mathscr{B} \leftarrow \mathscr{B}'}^{-1}\,\mathrm{C}_{\mathscr{B}}\left(\mathbf{v}\right)\end{aligned}$$

ed essendo la matrice associata $\mathbf{A}$ unica si ha che:

$$\mathbf{A} = \mathbf{M}_{\mathscr{D} \leftarrow \mathscr{D}'} \mathbf{A}' \mathbf{M}_{\mathscr{B} \leftarrow \mathscr{B}'}^{-1}$$

□

Esempio 2.6.14. Sia

$$B = \begin{bmatrix} 1 & 2 & 0 \\ 2 & -1 & 1 \end{bmatrix}$$

la matrice associata all'applicazione lineare $f : \mathbb{C}^3 \to \mathbb{C}^2$ rispetto alla base

$$\mathscr{B} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

sul dominio $\mathbb{C}^3$ e alla base

$$\mathscr{D} = \left\{ \mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

sul codominio $\mathbb{C}^2$. Desideriamo conoscere la matrice $\mathbf{A}$ associata a f rispetto alle basi canoniche su dominio e codominio. Scriviamo dunque le formule necessarie dove poniamo $\mathbf{P} = \mathbf{M}_{\mathscr{E}_3 \leftarrow \mathscr{B}}$ e $\mathbf{S} = \mathbf{M}_{\mathscr{E}_2 \leftarrow \mathscr{D}}$ per brevità

1. $\mathrm{C}_{\mathscr{E}_2}(f(\mathbf{v})) = \mathbf{A}\,\mathrm{C}_{\mathscr{E}_3}(\mathbf{v})$
2. $\mathrm{C}_{\mathscr{D}}(f(\mathbf{v})) = \mathbf{B}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v})$
3. $\mathrm{C}_{\mathscr{E}_3}(\mathbf{v}) = \mathbf{P}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v})$, dalla quale si ha la $3'$: $\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathbf{P}^{-1}\,\mathrm{C}_{\mathscr{E}_3}(\mathbf{v})$
4. $\mathrm{C}_{\mathscr{E}_2}(\mathbf{w}) = \mathbf{S}\,\mathrm{C}_{\mathscr{D}}(\mathbf{w})$

Abbiamo indicato con $\mathscr{E}_2$ e $\mathscr{E}_3$ le basi canoniche di $\mathbb{C}^2$ e $\mathbb{C}^3$ rispettivamente; secondo quanto appreso la matrice $\mathbf{P} = \mathbf{M}_{\mathscr{E}_3 \leftarrow \mathscr{B}}$ si scrive come

$$\mathbf{P} = \begin{bmatrix} \mathrm{C}_{\mathscr{E}_3}(\mathbf{v}_1) & \mathrm{C}_{\mathscr{E}_3}(\mathbf{v}_2) & \mathrm{C}_{\mathscr{E}_3}(\mathbf{v}_3) \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$$

mentre $\mathbf{S} = \mathbf{M}_{\mathscr{E}_2 \leftarrow \mathscr{D}}$ si scrive come

$$\mathbf{S} = \begin{bmatrix} \mathrm{C}_{\mathscr{E}_2}(\mathbf{w}_1) & \mathrm{C}_{\mathscr{E}_2}(\mathbf{w}_2) \end{bmatrix} = \begin{bmatrix} \mathbf{w}_1 & \mathbf{w}_2 \end{bmatrix}$$

Vogliamo dunque esprimere una relazione tipo la (1) attraverso le altre tre relazioni. Sia $\mathbf{v} \in \mathbb{C}^3$; allora

$$\begin{aligned} \mathrm{C}_{\mathscr{E}_2}(f(\mathbf{v})) &\overset{(4)}{=} \mathbf{S}\,\mathrm{C}_{\mathscr{D}}(f(\mathbf{v})) \\ &\overset{(2)}{=} \mathbf{SB}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v}) \\ &\overset{(3')}{=} \mathbf{SBP}^{-1}\,\mathrm{C}_{\mathscr{E}_3}(\mathbf{v}) \end{aligned}$$

e dunque $\mathbf{A} = \mathbf{SBP}^{-1}$. Calcoliamo l'inversa di $\mathbf{P}$:

$$\left[\begin{array}{ccc|ccc} 1 & 1 & -1 & 1 & 0 & 0 \\ 1 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \to \ldots \to \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 1 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & 1/2 & 1/2 \end{array}\right]$$

dunque

$$\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}$$

e quindi la matrice cercata è

$$\mathbf{A} = \mathbf{SBP}^{-1} = \begin{bmatrix} 2 & 2 & 1 \\ 1 & -1 & 1 \end{bmatrix}$$

Osservazione 176. L'importanza di conoscere la matrice associata rispetto alle basi canoniche risiede nel fatto che, per ogni $\mathbf{v} \in \mathbb{C}^3$

$$f(\mathbf{v}) = \mathrm{C}_{\mathscr{E}_2}(f(\mathbf{v})) = \mathbf{A}\,\mathrm{C}_{\mathscr{E}_3}(\mathbf{v}) = \mathbf{A}\mathbf{v}$$

cioè $f(\mathbf{v}) = \mathbf{A}\mathbf{v}$ e quindi f coincide con la pre-moltiplicazione per $\mathbf{A}$.

2.6.4 Spazi isomorfi e dimensione

Osservazione 177. Il fatto che due spazi vettoriali finitamente generati siano isomorfi è equivalente al fatto che abbiano la stessa dimensione.

Proposizione 2.6.6. *Siano V e W spazi vettoriali finitamente generati. Allora V e W sono isomorfi se e solo se* $\dim V = \dim W$

Dimostrazione. Supponiamo che $f : V \to W$ sia un isomorfismo e sia $\mathscr{B}$ una base di V; allora $f[\mathscr{B}]$ è un insieme linearmente indipendente in W (per 2.4.9, dato che f è iniettiva, essendo biettiva, e $\mathscr{B}$ è linearmente indipendente). Ma allora $\dim V \leq \dim W$ ($\dim V = \mathrm{Card}\,(\mathscr{B}) = \mathrm{Card}\,(f[\mathscr{B}])$, ma possono essere necessari altri vettori linearmente indipendenti oltre a quelli portati mediante f, al fine di formare una base di W). Viceversa, considerando f^{-1} otteniamo che $\dim W \leq \dim V$. Quindi complessivamente $\dim V = \dim W$.
Supponiamo invece che $\dim V = \dim W = n$, e siano $\mathscr{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$, $\mathscr{D} = \{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ basi di V e W rispettivamente. Per il teorema 2.4.15 esiste un'applicazione lineare $f : V \to W$ tale che:

$$f(\mathbf{v}_i) = \mathbf{w}_i, \quad (i = 1, \dots, n)$$

Allora:

TODO: non capisco, non dovrebbe essere $\mathrm{Im}\,(f) \subseteq \langle\mathscr{D}\rangle$?

- f è suriettiva perché $\mathrm{Im}\,(f) \supseteq \langle\mathscr{D}\rangle = W$;
- f è iniettiva per il teorema nullità + rango, infatti essendo $\dim \mathrm{Im}\,(f) = n$ (visto che $\mathrm{Im}\,(f)$ contiene la base $\mathscr{D}$):

$$\begin{aligned} \dim V &= \dim \mathrm{N}\,(f) + \dim \mathrm{Im}\,(f) \\ n &= \dim \mathrm{N}\,(f) + n \end{aligned}$$

  da cui $\dim \mathrm{N}\,(f) = 0$, per cui $\mathrm{N}\,(f) = \{\mathbf{0}\}$.

Quindi complessivamente f è biettiva e V e W sono isomorfi. □

Osservazione 178. In alcuni casi è possibile trovare isomorfismi fra due spazi che danno informazioni supplementari. Diamo alcuni esempi di isomorfismi importanti.

Esempio 2.6.15. Se $\mathbf{A}$ e $\mathbf{B}$ sono matrici $m \times n$ e $\mathbf{B} = \mathbf{FA}$ con $\mathbf{F}$ invertibile, allora $\mathrm{C}\,(\mathbf{A})$ è isomorfo a $\mathrm{C}\,(\mathbf{B})$.
Sia infatti $\mathbf{v} \in \mathrm{C}\,(\mathbf{A})$; allora $\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{a}_i$, dove abbiamo posto $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n]$.
Allora:

$$\mathbf{Fv} = \mathbf{F} \sum_{i=1}^{n} \alpha_i \mathbf{a}_i = \sum_{i=1}^{n} \alpha_i \mathbf{F} \mathbf{a}_i = \sum_{i=1}^{n} \alpha_i \mathbf{b}_i \in \mathrm{C}\,(\mathbf{B})$$

dove $\mathbf{B} = [\mathbf{b}_1 \ \dots \ \mathbf{b}_n]$. Quindi ponendo $f(\mathbf{v}) = \mathbf{Fv}$ per $\mathbf{v} \in \mathrm{C}\,(\mathbf{A})$, otteniamo un'applicazione $f : \mathrm{C}\,(\mathbf{A}) \to \mathrm{C}\,(\mathbf{B})$ che è evidentemente lineare (sempre per le caratteristiche della pre-moltiplicazione per matrici). Analogamente abbiamo l'applicazione $g : \mathrm{C}\,(\mathbf{B}) \to \mathrm{C}\,(\mathbf{A})$ definita da $g(\mathbf{w}) = \mathbf{F}^{-1}\mathbf{w}$ e, chiaramente, g è l'inversa di f.

Esempio 2.6.16. Sia $\mathbf{A}$ una matrice $m \times n$; siccome $\dim \mathrm{C}\left(\mathbf{A}^H\right) = \dim \mathrm{C}(\mathbf{A})$, gli spazi vettoriali $\mathrm{C}\left(\mathbf{A}^H\right)$ e $\mathrm{C}(\mathbf{A})$ sono isomorfi. Anche in questo caso esiste un isomorfismo definito in modo naturale. Consideriamo infatti $\mathbf{A} = [\mathbf{a}_1 \ \dots \ \mathbf{a}_n]$. Allora $\mathbf{A}^H\mathbf{a}_i \in \mathrm{C}\left(\mathbf{A}^H\right)$, quindi abbiamo l'applicazione $f : \mathrm{C}(\mathbf{A}) \to \mathrm{C}\left(\mathbf{A}^H\right)$ definita da $f(\mathbf{v}) = \mathbf{A}^H\mathbf{v}$.
Per il teorema nullità + rango ci basta dimostrare che f è iniettiva. Ora $\mathbf{v} \in \mathrm{C}(\mathbf{A})$ se e solo se $\mathbf{v} = \mathbf{A}\mathbf{u}$, per un opportuno $\mathbf{u} \in \mathbb{C}^n$. Se $f(\mathbf{v}) = \mathbf{0}$ abbiamo $\mathbf{0} = \mathbf{A}^H\mathbf{A}\mathbf{u}$ e quindi

$$0 = \mathbf{u}^H\mathbf{A}^H\mathbf{A}\mathbf{u} = (\mathbf{A}\mathbf{u})^H\mathbf{A}\mathbf{u} = \mathbf{v}^H\mathbf{v}$$

da cui ricaviamo, come in 1.6.10, che $\mathbf{v} = \mathbf{0}$.

Capitolo 3

Norme, prodotti interni e ortogonalità

Contents

3.1 Norme di vettori

Osservazione 179. Fissato su un piano un sistema di riferimento ortogonale monometrico Oxy si definisce una corrispondenza biunivoca tra l'insieme dei punti del piano e gli elementi di $\mathbb{R}^2$. Per ciascun punto P nel piano sarà possibile associare un vettore $\mathbf{v} = [a\,b]^T$, con a e b rispettivamente ascissa e l'ordinata di P rispetto a Oxy.

Definizione 3.1.1. La distanza di P da O (o lunghezza del segmento OP) è data da (teorema di Pitagora):

$$|OP| = \sqrt{|a|^2 + |b|^2} = \sqrt{a^2 + b^2} = \sqrt{\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} = \sqrt{\mathbf{v}^T\mathbf{v}}$$

Tale funzione $f : \mathbb{R}^2 \to \mathbb{R}$ esprime algebricamente la lunghezza di un vettore nel piano.

Proposizione 3.1.1. *Proprietà rilevanti di f sono:*

1. $f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \sqrt{a^2+b^2} \geq 0$, $\forall a, b \in \mathbb{R}$; $f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = 0$ *se e solo se* $a = b = 0$

2. *per ogni* $\alpha, a, b \in \mathbb{R}$
$$f\left(\alpha \begin{bmatrix} a \\ b \end{bmatrix}\right) = |\alpha|\, f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)$$

3. *per ogni* $a, b, c, d \in \mathbb{R}$ *(disuguaglianza triangolare, regola del parallelogramma):*
$$f\left(\begin{bmatrix} a+c \\ b+d \end{bmatrix}\right) \leq f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) + f\left(\begin{bmatrix} c \\ d \end{bmatrix}\right)$$

Osservazione 180. L'ultima proprietà esprime algebricamente il fatto che la lunghezza della diagonale di un parallelogramma è minore o uguale alla somma delle lunghezze di due suoi lati adiacenti

Dimostrazione. La prima è ovvia; per la seconda

$$f\left(\alpha \begin{bmatrix} a \\ b \end{bmatrix}\right) = f\left(\begin{bmatrix} \alpha a \\ \alpha b \end{bmatrix}\right) = \sqrt{(\alpha a)^2 + (\alpha b)^2} = |\alpha| \sqrt{a^2+b^2} = |\alpha|\, f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right)$$

Per la terza cominciamo osservando che, sviluppando $(ad+bc)^2 \geq 0$ e riorganizzandola, si ha $2acbd \leq b^2c^2 + a^2d^2$. Allora se si sviluppa $ac + bd$ attraverso la disuguaglianza $x \leq |x| = \sqrt{x^2}$:

$$\begin{aligned} ac + bd &\leq \sqrt{(ac+bd)^2} = \sqrt{a^2c^2 + 2acbd + b^2d^2} \leq \sqrt{a^2c^2 + b^2c^2 + a^2d^2 + b^2d^2} \\ &= \sqrt{(a^2+b^2)(c^2+d^2)} = \sqrt{\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} \sqrt{\begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}} \end{aligned}$$

Pertanto complessivamente:

$$ac + bd \leq \sqrt{\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} \sqrt{\begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}} \tag{3.1}$$

L'ultima disuguaglianza comporta

$$\begin{aligned} \begin{bmatrix} a+c & b+d \end{bmatrix} \begin{bmatrix} a+c \\ b+d \end{bmatrix} &= (a+c)^2 + (b+d)^2 = a^2 + c^2 + 2ac + b^2 + d^2 + 2bd \\ &= (a^2+b^2) + (c^2+d^2) + 2(ac+bd) \\ &\overset{(1)}{\leq} \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} + 2\sqrt{\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} \sqrt{\begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}} \\ &= \left(\sqrt{\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} + \sqrt{\begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}}\right)^2 \end{aligned}$$

dove (1) deriva dal confronto dell'ultimo dei tre membri alla luce di 3.1 (gli altri sono uguali). Pertanto qui si conclude che

$$\begin{bmatrix} a+c & b+d \end{bmatrix} \begin{bmatrix} a+c \\ b+d \end{bmatrix} \leq \left(\sqrt{\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} + \sqrt{\begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}} \right)^2$$

Ora estraendo la radice da quest'ultima si ottiene la disuguaglianza da dimostrare:

$$\sqrt{\begin{bmatrix} a+c & b+d \end{bmatrix} \begin{bmatrix} a+c \\ b+d \end{bmatrix}} \leq \sqrt{\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}} + \sqrt{\begin{bmatrix} c & d \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix}}$$

□

Osservazione 181. Introduciamo ora il concetto di norma come generalizzazione di quello di lunghezza visto per i vettori del piano.

Definizione 3.1.2 (Norma)**.** Sia V uno spazio vettoriale. Una funzione $\|\cdot\| : V \to \mathbb{R}_{\geq 0}$ si dice norma su V se soddisfa le tre seguenti proprietà

1. $\|\mathbf{v}\| > 0$ per ogni $\mathbf{v} \in V : \mathbf{v} \neq \mathbf{0}$; $\|\mathbf{0}\| = 0$;
2. $\|\alpha\mathbf{v}\| = |\alpha| \|\mathbf{v}\|$ per ogni scalare $\alpha \in \mathbb{C}$ e ogni $\mathbf{v} \in V$;
3. $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ per ogni $\mathbf{v}, \mathbf{w} \in V$ (disuguaglianza triangolare).

Esempio 3.1.1 (Norma euclidea di $\mathbb{R}^2$)**.** La funzione $f(\mathbf{v}) = \sqrt{\mathbf{v}^T\mathbf{v}}$ definita su $\mathbb{R}^2$ verifica le proprietà 1, 2 e 3 (come mostrato in precedenza) per cui è una norma sullo spazio vettoriale reale $V = \mathbb{R}^2$. Si chiama norma euclidea di $\mathbb{R}^2$

Proposizione 3.1.2 (Norma euclidea di $\mathbb{C}^n$)**.** *La funzione*

$$\|\cdot\|_2 : \mathbb{C}^n \to \mathbb{R}_{\geq 0} \quad \mathbf{v} \to \sqrt{\mathbf{v}^H\mathbf{v}} \tag{3.2}$$

è una norma sullo spazio vettoriale complesso $V = \mathbb{C}^n$*.*

Osservazione 182. Si osservi che $\mathbf{v}^H\mathbf{v}$ da un reale per la proprietà del prodotto di due coniugati; considerando un singolo elemento di $\mathbf{v}$, $a + ib \in \mathbb{C}$, nell'applicazione della funzione si ha che

$$(a + ib)(a - ib) = a^2 - i^2b^2 = a^2 + b^2 \in \mathbb{R}$$

pertanto $\mathbf{v}^H\mathbf{v}$ sarà la somma di tanti reali, dalla quale verrà estratta la radice (pertanto il risultato sarà reale).

Osservazione 183. Per provare che vale la disuguaglianza triangolare (e che quindi effettivamente la funzione definita è una norma) abbiamo bisogno di un risultato che ora enunciamo e che dimostreremo in seguito.

Proposizione 3.1.3 (Disuguaglianza di Cauchy-Schwarz)**.** *Siano* $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$*. Allora si ha*

$$\left|\mathbf{v}^H\mathbf{w}\right| \leq \|\mathbf{v}\|_2 \|\mathbf{w}\|_2 \tag{3.3}$$

Dimostrazione. Dimostriamo che la norma euclidea di $\mathbb{C}^n$ è effettivamente una norma. Per la prima proprietà da soddisfare, se:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{C}^n$$

allora

$$\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}^H \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \sum_{i=1}^{n} \overline{v_i} v_i = \sum_{i=1}^{n} |v_i|^2$$

dove l'ultima è dovuta al fatto che $(a+ib)(a-ib) = a^2+b^2$ Dunque $\|\cdot\|_2$ soddisfa la proprietà 1 poiché per ogni $v_i \in \mathbb{C}$ si ha $|v_i| \in \mathbb{R}_{\geq 0}$, per cui la somma di numeri reali non negativi è non negativa ed è nulla se e solo se ciascun addendo è nullo (che si verifica nel caso che $\mathbf{v} = \mathbf{0}$) .
Per verificare che soddisfa la proprietà 2 basta osservare che se $\mathbf{v} \in \mathbb{C}^n$ e $\alpha \in \mathbb{C}$ si ha

$$(\alpha\mathbf{v})^H(\alpha\mathbf{v}) = \overline{\alpha}\mathbf{v}^H\alpha\mathbf{v} = \overline{\alpha}\alpha\mathbf{v}^H\mathbf{v} = |\alpha|^2\,\mathbf{v}^H\mathbf{v} = |\alpha|^2\,\|\mathbf{v}\|^2$$

dalla quale cui estraendo la radice per il primo e ultimo membro (e considerando come è definita la funzione norma) si ha:

$$|\alpha|\,\|\mathbf{v}\| = \sqrt{(\alpha\mathbf{v})^H(\alpha\mathbf{v})} = \|\alpha\mathbf{v}\|$$

Per la proprietà 3: poiché per ogni $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$ si ha che $\|\mathbf{v}\|_2$, $\|\mathbf{w}\|_2$ e $\|\mathbf{v}+\mathbf{w}\|_2$ sono reali non negativi (lunghezze di vettori), verificare la proprietà 3 equivale a provare che (elevando al quadrato la proposizione da dimostrare):

$$\|\mathbf{v}+\mathbf{w}\|_2^2 \leq (\|\mathbf{v}\|_2 + \|\mathbf{w}\|_2)^2, \quad \forall \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$$

Pertanto, essendo

$$\|\mathbf{v}+\mathbf{w}\|_2^2 = (\mathbf{v}+\mathbf{w})^H(\mathbf{v}+\mathbf{w}) = (\mathbf{v}^H+\mathbf{w}^H)(\mathbf{v}+\mathbf{w}) = \mathbf{v}^H\mathbf{v}+\mathbf{v}^H\mathbf{w}+\mathbf{w}^H\mathbf{v}+\mathbf{w}^H\mathbf{w}$$

segue che $\mathbf{v}^H\mathbf{w}+\mathbf{w}^H\mathbf{v}$ è un numero reale: infatti, dato che $\|\mathbf{v}+\mathbf{w}\|_2^2$ lo è e $\mathbf{v}^H\mathbf{v}$ e $\mathbf{w}^H\mathbf{w}$ lo sono, lo deve esser per forza anche $\mathbf{v}^H\mathbf{w}+\mathbf{w}^H\mathbf{v}$. Essendo poi:

$$(\|\mathbf{v}\|_2 + \|\mathbf{w}\|_2)^2 = \|\mathbf{v}\|_2^2 + \|\mathbf{w}\|_2^2 + 2\,\|\mathbf{v}\|_2\,\|\mathbf{w}\|_2 = \mathbf{v}^H\mathbf{v} + \mathbf{w}^H\mathbf{w} + 2\,\|\mathbf{v}\|_2\,\|\mathbf{w}\|_2$$

espandiamo usando quanto ottenuto

$$\begin{gathered} \|\mathbf{v}+\mathbf{w}\|_2^2 \leq (\|\mathbf{v}\|_2 + \|\mathbf{w}\|_2)^2 \\ \mathbf{v}^H\mathbf{v} + \mathbf{v}^H\mathbf{w} + \mathbf{w}^H\mathbf{v} + \mathbf{w}^H\mathbf{w} \leq \mathbf{v}^H\mathbf{v} + \mathbf{w}^H\mathbf{w} + 2\,\|\mathbf{v}\|_2\,\|\mathbf{w}\|_2 \end{gathered}$$

e sottraendo $\mathbf{v}^H\mathbf{v} + \mathbf{w}^H\mathbf{w}$ da entrambi i membri ci possiamo ricondurre a verificare la seguente disuguaglianza tra numeri reali:

$$\mathbf{v}^H\mathbf{w} + \mathbf{w}^H\mathbf{v} \leq 2\,\|\mathbf{v}\|_2\,\|\mathbf{w}\|_2 \tag{3.4}$$

Si ottiene

$$\mathbf{v}^H\mathbf{w} + \mathbf{w}^H\mathbf{v} \overset{(1)}{\leq} \left|\mathbf{v}^H\mathbf{w} + \mathbf{w}^H\mathbf{v}\right| \overset{(2)}{\leq} \left|\mathbf{v}^H\mathbf{w}\right| + \left|\mathbf{w}^H\mathbf{v}\right| \tag{3.5}$$

con (1) da $a \leq |a|$ per $a \in \mathbb{R}$, e (2) dalla disuguaglianza triangolare del modulo di numeri complessi.
Ricordiamo ora che se $z \in \mathbb{C}$ allora $z^H = \overline{z}$ per cui $\left|z^H\right| = |\overline{z}|$ (un complesso e il suo coniugato hanno lo stesso modulo). In particolare si ha che il complesso $(\mathbf{w}^H\mathbf{v})^H = \mathbf{v}^H(\mathbf{w}^H)^H = \mathbf{v}^H\mathbf{w}$ e pertanto che i loro moduli sono uguali: $\left|\mathbf{w}^H\mathbf{v}\right| = \left|\mathbf{v}^H\mathbf{w}\right|$.
Assieme alla disuguaglianza di Cauchy-Schwarz ciò comporta

$$\left|\mathbf{v}^H\mathbf{w}\right| + \left|\mathbf{w}^H\mathbf{v}\right| = \left|\mathbf{v}^H\mathbf{w}\right| + \left|\mathbf{v}^H\mathbf{w}\right| = 2\left|\mathbf{v}^H\mathbf{w}\right| \overset{(1)}{\leq} 2\left\|\mathbf{v}\right\|_2\left\|\mathbf{w}\right\|_2 \tag{3.6}$$

con (1) dovuto al moltiplicare per 2 della disuguaglianza di Cauchy-Schwartz. Pertanto la 3.4 segue da 3.5 e 3.6.
La dimostrazione che $\|\cdot\|_2$ è una norma è così completata. □

Esempio 3.1.2.

$$\left\|\begin{bmatrix}1\\-2i\\2+i\end{bmatrix}\right\|_2 = \sqrt{\begin{bmatrix}1\\-2i\\2+i\end{bmatrix}^H\begin{bmatrix}1\\-2i\\2+i\end{bmatrix}} = \sqrt{\begin{bmatrix}1 & 2i & 2-i\end{bmatrix}\begin{bmatrix}1\\-2i\\2+i\end{bmatrix}}$$
$$= \sqrt{1+4+(2-i)(2+i)} = \sqrt{1+4+4+1} = \sqrt{10}$$

Osservazione 184. Anche nel piano, la norma euclidea non è l'unica norma adottata; il prossimo esempio, nel caso particolare di uno spazio vettoriale reale di dimensione 2, esprime la distanza minima che occorre percorrere da un punto a un altro di una città in cui le strade siano a due a due perpendicolari o parallele.

Proposizione 3.1.4 (Manhattan distance)**.** *Le funzioni*

$$\begin{aligned}\|\cdot\|_1 : \mathbb{C}^n \to \mathbb{R}_{\geq 0} \quad &\textit{definita da} \quad \|\mathbf{v}\|_1 = |v_1| + |v_2| + \ldots + |v_n|\\ \|\cdot\|_1 : \mathbb{R}^n \to \mathbb{R}_{\geq 0} \quad &\textit{definita da} \quad \|\mathbf{v}\|_1 = |v_1| + |v_2| + \ldots + |v_n|\end{aligned}$$

per ogni $\mathbf{v} = \begin{bmatrix}v_1 & v_2 & \ldots & v_n\end{bmatrix}^T \in \mathbb{C}^n$ *(o rispettivamente in* $\mathbb{R}^n$*) sono due norme sugli spazi vettoriali* $\mathbb{C}^n$ *e* $\mathbb{R}^n$*.*

Dimostrazione. Sia $\mathbf{v} = [v_1 \ldots v_n]^T$; si ha che $\|\mathbf{v}\|_1 \geq 0$ e $\|\mathbf{v}\|_1 = 0 \iff \mathbf{v} = \mathbf{0}$ facilmente in quanto $\|\mathbf{0}\| = |0| + \ldots + |0| = 0$ e nei rimanenti casi la norma è somma di termini positivi o nulli (essendo sotto valore assoluto).
Per la seconda occorre mostrare che $\|\alpha\mathbf{v}\| = |\alpha|\,\|\mathbf{v}\|$; si ha che:

$$\|\alpha\mathbf{v}\|_1 = \left\|\begin{bmatrix}\alpha v_1\\ \ldots\\ \alpha v_n\end{bmatrix}\right\|_1 = |\alpha v_1| + \ldots |\alpha v_n| = |\alpha|\,|v_1| + \ldots |\alpha|\,|v_n|$$
$$= |\alpha|\left(|v_1| + \ldots + |v_n|\right) = |\alpha|\,\|\mathbf{v}\|_1$$

Per la terza mostriamo che $\|\mathbf{v}+\mathbf{w}\|_1 \leq \|\mathbf{v}\|_1 + \|\mathbf{w}\|_1$; si ha che

$$\left\| \begin{bmatrix} v_1 + w_1 \\ \dots \\ v_n + w_n \end{bmatrix} \right\|_1 = |v_1 + w_1| + \ldots + |v_n + w_n|$$

mentre

$$\|\mathbf{v}\|_1 + \|\mathbf{w}\|_1 = |v_1| + |w_1| + \ldots + |v_n| + |w_n|$$

Ora dato che a livello di singolo elemento (per la disuguaglianza triangolare sia in $\mathbb{C}$ che in $\mathbb{R}$) si ha che $|v_i + w_i| \leq |v_i| + |w_i|$, allora complessivamente $\|\mathbf{v}+\mathbf{w}\|_1 \leq \|\mathbf{v}\|_1 + \|\mathbf{w}\|_1$ □

Esempio 3.1.3.

$$\left\| \begin{bmatrix} 1 \\ -2i \\ 2+i \end{bmatrix} \right\| = |1| + |-2i| + |2+i| = 1 + 2 + \sqrt{4+1} = 3 + \sqrt{5}$$

Osservazione 185. Quando si è interessati a misurare le variazioni massime da un valore medio si usa la norma seguente

Proposizione 3.1.5 (Maximum distance)**.** *Le funzioni*

$$\begin{aligned} \|\cdot\|_\infty &: \mathbb{C}^n \to \mathbb{R}_{\geq 0} \quad \textit{definita da} \quad \|\mathbf{v}\|_\infty = \max(|v_1|, \ldots, |v_n|) \\ \|\cdot\|_\infty &: \mathbb{R}^n \to \mathbb{R}_{\geq 0} \quad \textit{definita da} \quad \|\mathbf{v}\|_\infty = \max(|v_1|, \ldots, |v_n|) \end{aligned}$$

per ogni $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T \in \mathbb{C}^n$ *(o rispettivamente in* $\mathbb{R}^n$*) sono due norme sugli spazi vettoriali* $\mathbb{C}^n$ *e* $\mathbb{R}^n$*.*

Dimostrazione. Verifichiamo solo che vale la proprietà 3, essendo le altre due verifiche banali: siano $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T, \mathbf{w} = \begin{bmatrix} w_1 & w_2 & \dots & w_n \end{bmatrix}^T \in \mathbb{C}^n$ e $k \in \{1, \ldots, n\}$ tale che

$$\|\mathbf{v}+\mathbf{w}\|_\infty = \max\{|v_1 + w_1|, \ldots |v_n + w_n|\} = |v_k + w_k|$$

Applicando la disuguaglianza triangolare del modulo di numeri si ottiene:

$$\begin{aligned} \|\mathbf{v}+\mathbf{w}\|_\infty &= |v_k + w_k| \leq |v_k| + |w_k| \leq \max(|v_1|, \ldots, |v_n|) + \max(|w_1|, \ldots, |w_n|) \\ &= \|\mathbf{v}\|_\infty + \|\mathbf{w}\|_\infty \end{aligned}$$

e quindi la proprietà 3 è dimostrata (la disuguaglianza si trova nei membri intermedi). □

Esempio 3.1.4.

$$\left\| \begin{bmatrix} 1 \\ -2i \\ 2+i \end{bmatrix} \right\|_\infty = \max(|1|, |-2i|, |2+i|) = \max\left(1, 2, \sqrt{5}\right) = \sqrt{5}$$

Proposizione 3.1.6. *Se* $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix}^T \in \mathbb{C}^n$ *allora*

$$\|\mathbf{v}\|_\infty \leq \|\mathbf{v}\|_2 \leq \|\mathbf{v}\|_1 \tag{3.7}$$

Dimostrazione. Infatti

$$\max\{|v_1|,\ldots,|v_n|\} \leq \sqrt{\mathbf{v}^H\mathbf{v}} = \sqrt{\sum_{i=1}^{n} |v_i|^2} \leq \sum_{i=1}^{n} |v_i|$$

□

Proposizione 3.1.7. *Sia $p \in \mathbb{N} : p \geq 1$. La funzione*

$$\|\cdot\|_p : \mathbb{C}^n \to \mathbb{R}_{\geq 0} \quad \textit{definita da} \quad \|\mathbf{v}\|_p = \left(\sum_{i=1}^{n} |v_i|^p\right)^{\frac{1}{p}}$$

per ogni $\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \ldots & v_n \end{bmatrix}^T \in \mathbb{C}^n$ *è una norma sullo spazio vettoriale* $\mathbb{C}^n$*.*

Dimostrazione. Non fatta; per la dimostrazione serve la disuguaglianza di Minkowski (generalizzazione della disuguaglianza triangolare) e della disuguaglianza di Holder (generalizzazione del teorema di Schwartz). □

Osservazione 186. Se $p = 1$ o $p = 2$ si ottengono rispettivamente $\|\cdot\|_1$ e $\|\cdot\|_2$; pertanto questa norma ne costituisce una generalizzazione.

Esempio 3.1.5.

$$\left\| \begin{bmatrix} 1 \\ -2i \\ 2+i \end{bmatrix} \right\|_3 = \sqrt[3]{|1|^3 + |-2i|^3 + |2+i|^3} = \sqrt[3]{1 + 2^3 + (\sqrt{5})^3} = \sqrt[3]{9 + 5\sqrt{5}}$$

Osservazione 187. A partire dalla norma euclidea si possono costruire altre norme nel modo descritto nella prossima proposizione.

Proposizione 3.1.8. *Sia* $\mathbf{B}$ *una matrice complessa invertibile* $n \times n$ *e sia* $\mathbf{A}$ *la matrice hermitiana di ordine* n *definita da* $\mathbf{A} = \mathbf{B}\mathbf{B}^H$ *(è hermitiana perché* $(\mathbf{B}\mathbf{B}^H)^H = \mathbf{B}\mathbf{B}^H$*). La funzione* $\|\cdot\|_{\mathbf{A}} : \mathbb{C}^n \to \mathbb{R}_{\geq 0}$ *definita da:*

$$\|\mathbf{v}\|_{\mathbf{A}} = \sqrt{\mathbf{v}^H\mathbf{A}\mathbf{v}} = \sqrt{\mathbf{v}^H\mathbf{B}\mathbf{B}^H\mathbf{v}} = \sqrt{(\mathbf{B}^H\mathbf{v})^H\mathbf{B}^H\mathbf{v}} = \left\|\mathbf{B}^H\mathbf{v}\right\|_2$$

per ogni $\mathbf{v} \in \mathbb{C}^n$ *(*$\mathbf{A}$ *funge come parametro) è una norma sullo spazio vettoriale* $\mathbb{C}^n$*.*

Dimostrazione. Per provare la 1 si deve sfruttare l'ipotesi che $\mathbf{B}$ è invertibile. Se $\mathbf{v} = \mathbf{0}$ si ha che $\left\|\mathbf{B}^H\mathbf{v}\right\|_2 = 0$ quale che sia $\mathbf{B}^H$. Se $\mathbf{v} \neq \mathbf{0}$ si ha che $\left\|\mathbf{B}^H\mathbf{v}\right\|_2 > 0$ se e solo se $\mathbf{B}^H\mathbf{v} \neq \mathbf{0}$; ma dato che $\mathbf{B}^H$ è invertibile (essendolo $\mathbf{B}$) si ha che l'unico caso in cui $\mathbf{B}^H\mathbf{v} = \mathbf{0}$ è se $\mathbf{v} = \mathbf{0}$; ma qui siamo nell'ipotesi che $\mathbf{v} \neq \mathbf{0}$ quindi $\mathbf{B}^H\mathbf{v} \neq \mathbf{0}$. Per la seconda si ha che

$$\|\alpha\mathbf{v}\|_{\mathbf{A}} = \left\|\alpha\mathbf{B}^H\mathbf{v}\right\|_2 \overset{(1)}{=} |\alpha| \left\|\mathbf{B}^H\mathbf{v}\right\|_2 = |\alpha| \, \|\mathbf{v}\|_{\mathbf{A}}$$

con (1) per la seconda proprietà della norma euclidea. Per la terza proprietà

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\|_{\mathbf{A}} &= \left\|\mathbf{B}^H(\mathbf{v} + \mathbf{w})\right\|_2 = \left\|\mathbf{B}^H\mathbf{w} + \mathbf{B}^H\mathbf{v}\right\|_2 \\ &\overset{(2)}{\leq} \left\|\mathbf{B}^H\mathbf{w}\right\|_2 + \left\|\mathbf{B}^H\mathbf{v}\right\|_2 = \|\mathbf{v}\|_{\mathbf{A}} + \|\mathbf{w}\|_{\mathbf{A}} \end{aligned}$$

□

Proposizione 3.1.9. *Sia $\mathscr{C}([a,b])$ lo spazio vettoriale reale delle funzioni reali continue definite nell'intervallo $[a,b]$ con $a,b \in \mathbb{R}$ e $a < b$. La funzione*

$$\|\cdot\|_2 : \mathscr{C}([a,b]) \to \mathbb{R}_{\geq 0} \quad \textit{definita da} \quad \|f\|_2 = \sqrt{\int_a^b [f(t)]^2\, dt} \tag{3.8}$$

per ogni $f \in \mathscr{C}([a,b])$, è una norma su $\mathscr{C}([a,b])$

Osservazione 188. Per la verifica della terza proprietà si impiega la seguente disuguaglianza che corrisponde a quella di Cauchy-Schwarz, enunciata in seguito.

Proposizione 3.1.10 (Disuguaglianza di Schwarz per funzioni continue)**.** *Siano $f, g \in \mathscr{C}([a,b])$. Allora si ha:*

$$\left| \int_a^b f(t)g(t)\, dt \right| \leq \sqrt{\int_a^b [f(t)]^2\, dt} \sqrt{\int_a^b [g(t)]^2\, dt} \tag{3.9}$$

Dimostrazione. Si ha che $\|\mathbf{0}\|_2 = 0$ (la funzione sempre nulla ha integrale nullo); alternativamente se $\mathbf{v} \neq \mathbf{0}$ il quadrato della funzione fa si che in integrazione si sommino sempre valori positivi e il risultato è la radice (positiva) su un numero positivo.
Per la seconda mostriamo che $\|\alpha f\| = |\alpha|\, \|f\|$

$$\|\alpha f\| = \sqrt{\int_a^b [\alpha f(t)]^2\, dt} = \sqrt{\alpha^2 \int_a^b f(t)^2\, dt} = |\alpha| \sqrt{\int_a^b f(t)^2\, dt} = |\alpha|\, \|f\|$$

Per la terza dobbiamo mostrare che $\|f+g\| \leq \|f\| + \|g\|$:

$$\begin{aligned}
\|f+g\| &= \sqrt{\int_a^b [f+g(t)]^2\, dt} = \sqrt{\int_a^b [f(t)+g(t)]^2\, dt} \\
&= \sqrt{\int_a^b [f(t)]^2 + [g(t)]^2 + 2f(t)g(t)\, dt} \\
&= \sqrt{\int_a^b [f(t)]^2\, dt} + \sqrt{\int_a^b [g(t)]^2\, dt} + \sqrt{\int_a^b 2f(t)g(t)\, dt} \\
&= \|f\| + \|g\| + \sqrt{2\int_a^b f(t)g(t)\, dt}
\end{aligned}$$

TODO: fixme

Non torna: qui stiamo mostrando che $\|f+g\| \geq \|f\| + \|g\|$ □

Esempio 3.1.6. Verificare che $f : \mathbb{C}^3 \to \mathbb{R}^+ \cup \{0\}$ definita da

$$f\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right) = |2a - b| + |a + c| + |ib|$$

è una norma di $\mathbb{C}^3$.
Si ha che $f(\mathbf{v}) \geq 0$ essendo una somma di valori assoluti, inoltre $f(\mathbf{0}) = 0$ come

si può verificare algebricamente; pertanto la prima proprietà è verificata. Per la seconda

$$\begin{aligned} f(\alpha\mathbf{v}) = f\left(\begin{bmatrix}\alpha a\\ \alpha b\\ \alpha c\end{bmatrix}\right) &= |2\alpha a - \alpha b| + |\alpha a - \alpha c| + |i\alpha b| \\ &= |\alpha|\,|2a-b| + |\alpha|\,|a-c| + |\alpha|\,|ib| \\ &= |\alpha|\,(|2a-b| + |a+c| + |ib|) \\ &= |\alpha|\,f(\mathbf{v}) \end{aligned}$$

Per la terza bisogna mostrare l'equivalente di $\|\mathbf{v}+\mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$; siano

$$\mathbf{v} = \begin{bmatrix}a\\ b\\ c\end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix}d\\ e\\ f\end{bmatrix}, \quad \mathbf{v}+\mathbf{w} = \begin{bmatrix}a+d\\ b+e\\ c+f\end{bmatrix}$$

Allora

$$\begin{aligned} f(\mathbf{v}+\mathbf{w}) &= |2a+2d-b-e| + |a+d-c-f| + |i(b+e)| \\ f(\mathbf{v}) &= |2a-b| + |a-c| + |ib| \\ f(\mathbf{w}) &= |2d-e| + |d-f| + |ie| \end{aligned}$$

per ogni termine che compone la sommatoria si può applicare la disuguaglianza triangolare $|x+y| \leq |x| + |y|$ ossia

$$\begin{aligned} &|(2a-b)+(2d-e)| \leq |2a-b| + |2d-c| \\ &|(a-c)+(d-f)| \leq |a-c| + |d-f| \\ &|ib+ic| \leq |ib| + |ic| \end{aligned}$$

dunque dai vari confronti si conclude che deve essere $f(\mathbf{v}+\mathbf{w}) \leq f(\mathbf{v}) + f(\mathbf{w})$

3.2 Prodotti interni

Definizione 3.2.1 (Prodotto interno)**.** Sia V uno spazio vettoriale. Si chiama prodotto interno su V ogni funzione

$$(\cdot|\cdot) : V \times V \to \mathbb{C} \tag{3.10}$$

che soddisfi le seguenti proprietà:

1. $(\mathbf{v}|\mathbf{w}) = \overline{(\mathbf{w}|\mathbf{v})}$ per ogni $\mathbf{v}, \mathbf{w} \in V$ (funzione hermitiana)
2. $(\mathbf{v}|\alpha\mathbf{w} + \beta\mathbf{z}) = \alpha(\mathbf{v}|\mathbf{w}) + \beta(\mathbf{v}|\mathbf{z})$ per ogni $\mathbf{v}, \mathbf{w}, \mathbf{z} \in V$ e $\alpha, \beta \in \mathbb{C}$ (funzione lineare nella seconda componente)
3. $(\mathbf{v}|\mathbf{v}) \in \mathbb{R}_{>0}$ se $\mathbf{v} \neq \mathbf{0}$ (funzione definita positiva)

Proposizione 3.2.1. *Se $(\cdot|\cdot)$ è un prodotto interno su uno spazio vettoriale V su $\mathbb{C}$ allora si ha:*

$$(\alpha\mathbf{u} + \beta\mathbf{v}|\mathbf{w}) = \overline{\alpha}(\mathbf{u}|\mathbf{w}) + \overline{\beta}(\mathbf{v}|\mathbf{w}) \tag{3.11}$$

per ogni $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ e ogni $\alpha, \beta \in \mathbb{C}$

Dimostrazione.

$$\begin{aligned}(\alpha\mathbf{u}+\beta\mathbf{v}|\mathbf{w}) &= \overline{(\mathbf{w}|\alpha\mathbf{u}+\beta\mathbf{v})} = \overline{[\alpha(\mathbf{w}|\mathbf{u})+\beta(\mathbf{w}|\mathbf{v})]} \\ &= \overline{\alpha}\overline{(\mathbf{w}|\mathbf{u})}+\overline{\beta}\ \overline{(\mathbf{w}|\mathbf{v})} = \overline{\alpha}(\mathbf{u}|\mathbf{w})+\overline{\beta}(\mathbf{v}|\mathbf{w})\end{aligned}$$

□

Proposizione 3.2.2. *Si ha*

$$(\mathbf{0}|\mathbf{w}) = 0 = (\mathbf{v}|\mathbf{0}) \tag{3.12}$$

per ogni $\mathbf{v},\mathbf{w}\in V$ *(in particolare, anche se* $\mathbf{v}=\mathbf{0}$ *o* $\mathbf{w}=\mathbf{0}$*).*

Dimostrazione.

$$\begin{aligned}(\mathbf{0}|\mathbf{w}) &= (\mathbf{v}-\mathbf{v}|\mathbf{w}) \overset{(1)}{=} (\mathbf{v}|\mathbf{w})-(\mathbf{v}|\mathbf{w}) = 0 \\ (\mathbf{w}|\mathbf{0}) &= (\mathbf{v}|\mathbf{w}-\mathbf{w}) \overset{(2)}{=} (\mathbf{v}|\mathbf{w})-(\mathbf{v}|\mathbf{w}) = 0\end{aligned}$$

con (1) per la 3.11 e la (2) per la seconda proprietà della definizione di prodotto interno. □

Definizione 3.2.2 (Spazio vettoriale euclideo)**.** Spazio vettoriale su cui sia definito *un* prodotto interno $(\cdot|\cdot)$.

Proposizione 3.2.3 (Prodotto interno standard in $\mathbb{C}^n$ e $\mathbb{R}^n$)**.** *Siano* $\mathbf{v},\mathbf{w}\in\mathbb{C}^n$*; la funzione:*

$$(\cdot|\cdot):\mathbb{C}^n\times\mathbb{C}^n\to\mathbb{C} \quad \textit{definita da} \quad (\mathbf{v}|\mathbf{w})=\mathbf{v}^H\mathbf{w} \quad \textit{con } \mathbf{v},\mathbf{w}\in\mathbb{C}^n \tag{3.13}$$

$$(\cdot|\cdot):\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R} \quad \textit{definita da} \quad (\mathbf{v}|\mathbf{w})=\mathbf{v}^T\mathbf{w} \quad \textit{con } \mathbf{v},\mathbf{w}\in\mathbb{R}^n \tag{3.14}$$

soddisfano le proprietà del prodotto interno sugli spazi vettoriali $\mathbb{C}^n$ *e in* $\mathbb{R}^n$*, e sono chiamate prodotto interno standard in* $\mathbb{C}^n$ *e in* $\mathbb{R}^n$*.*

Dimostrazione. La proprietà 1 segue dal fatto che, essendo $\mathbf{v}^H\mathbf{w}\in\mathbb{C}$ si ha $(\mathbf{v}^H\mathbf{w})^H=\overline{\mathbf{v}^H\mathbf{w}}$ (l'H-trasposta di un numero consta nel suo coniugato) per cui

$$\mathbf{v}^H\mathbf{w} \overset{(1)}{=} \overline{(\mathbf{v}^H\mathbf{w})^H} \overset{(2)}{=} \overline{\mathbf{w}^H\mathbf{v}}$$

dove in (1) abbiamo coniugato due volte un numero mentre in (2) si sono applicate le proprietà dell'H-trasposizione.
Per proprietà 2, se $(\mathbf{v}|\mathbf{w})=\mathbf{v}^H\mathbf{w}$ allora:

$$(\mathbf{v}|\alpha\mathbf{w}+\beta\mathbf{z}) = \mathbf{v}^H(\alpha\mathbf{w}+\beta\mathbf{z}) = \alpha\mathbf{v}^H\mathbf{w}+\beta\mathbf{v}^H\mathbf{z} = \alpha(\mathbf{v}|\mathbf{w})+\beta(\mathbf{v}|\mathbf{z})$$

Infine la proprietà 3 si desume osservando che:

$$(\mathbf{v}|\mathbf{v}) = \|\mathbf{v}\|_2^2 > 0 \iff \mathbf{v}\neq\mathbf{0}$$

□

Esempio 3.2.1. Il prodotto interno standard dei due vettori $\begin{bmatrix} 1+4i \\ 2-i \\ i \end{bmatrix}, \begin{bmatrix} 3-i \\ 6+2i \\ 5 \end{bmatrix}$

$$\begin{aligned}\begin{bmatrix} 1+4i \\ 2-i \\ i \end{bmatrix}^H \begin{bmatrix} 3-i \\ 6+2i \\ 5 \end{bmatrix} &= \begin{bmatrix} 1-4i & 2+i & -i \end{bmatrix} \begin{bmatrix} 3-i \\ 6+2i \\ 5 \end{bmatrix} \\ &= (1-4i)(3-i)+(2+i)(6+2i)+5(-i) \\ &= 9-8i\end{aligned}$$

Esempio 3.2.2. Sia V il sottospazio di $\mathbb{C}^2$. Si verifichi che $(\cdot,\cdot) : V \times V \to \mathbb{C}$ definita da

$$\left(\begin{bmatrix} a \\ a \end{bmatrix}, \begin{bmatrix} b \\ b \end{bmatrix} = 3\overline{a}b \right)$$

è un prodotto interno.

Per la prima proprietà se $\mathbf{v} = \begin{bmatrix} a \\ a \end{bmatrix}, \mathbf{w} = \begin{bmatrix} b \\ b \end{bmatrix}$ si ha che

$$(\mathbf{w}|\mathbf{v}) = 3\overline{b}a$$

e dunque

$$\overline{(\mathbf{w}|\mathbf{v})} = \overline{3\overline{b}a} = 3\overline{a}b = (\mathbf{v}|\mathbf{w})$$

Per la seconda proprietà se inoltre si ha $\mathbf{z} = \begin{bmatrix} c \\ c \end{bmatrix}$; dobbiamo mostrare che $(\mathbf{v}|\alpha\mathbf{w}+\beta\mathbf{z}) = \alpha(\mathbf{v}|\mathbf{w}) + \beta(\mathbf{v}|\mathbf{z})$. Si ha che $\alpha\mathbf{w}+\beta\mathbf{z} = \begin{bmatrix} \alpha b+\beta c \\ \alpha b+\beta c \end{bmatrix}$ per cui

$$(\mathbf{v}|\alpha\mathbf{w}+\beta\mathbf{z}) = 3\overline{a}(\alpha b+\beta c) = 3\overline{a}\alpha b + 3\overline{a}\beta c = \alpha 3\overline{a}b + \beta 3\overline{a}c = \alpha(\mathbf{v}|\mathbf{w}) + \beta(\mathbf{v}|\mathbf{z})$$

Infine se $\mathbf{v} \neq \mathbf{0} \iff a \neq 0$ dunque

$$(\mathbf{v}|\mathbf{v}) = 3\overline{a}a \neq 0$$

Esempio 3.2.3. Sia V lo spazio vettoriale reale delle matrici 2×2 reali simmetriche: si verifichi che $(\cdot,\cdot) : V \times V \to \mathbb{R}$ definita da

$$\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \right) = aa' + 2bb' + cc'$$

è un prodotto interno.

Per la prima proprietà siano $\mathbf{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ e $\mathbf{w} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$. Si ha che

$$(\mathbf{w}|\mathbf{v}) = a'a + 2b'b + c'c$$

dunque

$$\overline{(\mathbf{w}|\mathbf{v})} = \overline{a'a + 2b'b + c'c} \overset{(1)}{=} a'a + 2b'b + c'c = (\mathbf{v}|\mathbf{w})$$

con (1) dato che un reale non è modificato dal coniugio. Per la seconda proprietà sia $\mathbf{z} = \begin{bmatrix} a'' & b'' \\ c'' & d'' \end{bmatrix}$ e dunque $\alpha\mathbf{w}+\beta\mathbf{z} = \begin{bmatrix} \alpha a'+\beta a'' & \alpha b'+\beta b'' \\ \alpha c'+\beta c'' & \alpha d'+\beta d'' \end{bmatrix}$, per cui

$$\begin{aligned}(\mathbf{v}|\alpha\mathbf{w}+\beta\mathbf{z}) &= a(\alpha a'+\beta a'') + 2b(\alpha b'+\beta b'') + c(\alpha c'+\beta c'') \\ &= \alpha aa' + \beta aa'' + \alpha 2bb' + \beta 2bb'' + \alpha cc' + \beta cc' \\ &= \alpha(\mathbf{v}|\mathbf{w}) + \beta(\mathbf{v}|\mathbf{z})\end{aligned}$$

Infine se $\mathbf{v} \neq \mathbf{0}$ si ha che $(\mathbf{v}|\mathbf{v}) = a^2 + 2b^2 + c^2 > 0$ (perché $a, b, c \neq 0$)

Proposizione 3.2.4. *Sia $\mathscr{C}([a,b])$ lo spazio vettoriale reale delle funzioni reali continue definite nell'intervallo $[a,b]$ con $a, b \in \mathbb{R}$, $a < b$. La funzione:*

$$(\cdot|\cdot) : \mathscr{C}([a,b]) \times \mathscr{C}([a,b]) \to \mathbb{R} \quad \textit{definita come} \quad (f|g) = \int_a^b f(t)g(t)\, dt$$

(l'integrale del prodotto delle 2 funzioni continue tra a e b) è un prodotto interno sullo spazio vettoriale reale $V = \mathscr{C}([a,b])$.

Dimostrazione. La verifica della proprietà 1 segue dal fatto che $f(t)g(t) = g(t)f(t)$ per ogni $t \in [a,b]$; la 2 dalla linearità dell'integrale:

$$\begin{aligned}
(f|\alpha g + \beta h) &= \int_a^b f(t) \cdot (\alpha g + \beta h)(t)\, dt \\
&= \int_a^b f(t) \cdot (\alpha g(t) + \beta h(t))\, dt \\
&= \int_a^b f(t) \cdot \alpha g(t)\, dt + \int_a^b f(t) \cdot \beta h(t))\, dt \\
&= \alpha(f|g) + \beta(f|h)
\end{aligned}$$

La 3 dall'osservazione che

$$(f|f) = \|f\|_2^2$$

dove $\|f\|_2$ è la norma di $f \in \mathscr{C}([a,b])$ definita in proposizione 3.1.9. □

Esempio 3.2.4. Il prodotto interno delle funzioni $f(t) = t$ e $g(t) = t^2 + 1$ ristrette all'intervallo $[0,2]$ è:

$$\int_0^2 t(t^2+1)\, dt = \int_0^2 t^3\, dt + \int_0^2 t\, dt = \frac{1}{4}2^4 + \frac{1}{2}2^2 = 2^2 + 2 = 6$$

Proposizione 3.2.5. *Sia $\mathbf{B}$ una matrice complessa $n \times n$ invertibile e sia $\mathbf{A}$ la matrice hermitiana di ordine n definita da $\mathbf{A} = \mathbf{B}\mathbf{B}^H$. Siano $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$; la funzione*

$$(\cdot|\cdot)_{\mathbb{A}} : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} \quad \textit{definita da} \quad (\mathbf{v}|\mathbf{w})_{\mathbf{A}} = \mathbf{v}^H\mathbf{A}\mathbf{w} \tag{3.15}$$

è un prodotto interno su $\mathbb{C}^n$.

Dimostrazione. La verifica che soddisfa le relative proprietà:

- per la prima si ha che $(\mathbf{w}|\mathbf{v}) = \mathbf{w}^H\mathbf{A}\mathbf{v}$ per cui dato che per un numero singolo la H-trasposizione coincide con il coniugio:

  $$\overline{(\mathbf{w}|\mathbf{v})} = (\mathbf{w}^H\mathbf{A}\mathbf{v})^H = \mathbf{v}^H\mathbf{A}^H\mathbf{w} = \mathbf{v}^H\mathbf{A}\mathbf{w} = (\mathbf{v}|\mathbf{w})$$

  dove si è sfruttato che $\mathbf{A}^H = \mathbf{A}$ essendo $\mathbf{A}$ hermitiana.

- per la seconda

  $$\begin{aligned}
  (\mathbf{v}|\alpha\mathbf{w} + \beta z) &= \mathbf{v}^H\mathbf{A}\,[\alpha\mathbf{w} + \beta\mathbf{z}] = \mathbf{v}^H\mathbf{A}\alpha\mathbf{w} + \mathbf{v}^H\mathbf{A}\beta\mathbf{z} = \alpha\mathbf{v}^H\mathbf{A}\mathbf{w} + \beta\mathbf{v}^H\mathbf{A}\mathbf{z} \\
  &= \alpha(\mathbf{v}|\mathbf{w}) + \beta(\mathbf{v}|\mathbf{z})
  \end{aligned}$$

- per la terza conviene osservare che

$$(\mathbf{v}|\mathbf{v})_{\mathbf{A}} = \mathbf{v}^H\mathbf{A}\mathbf{v} = \mathbf{v}^H\mathbf{B}\mathbf{B}^H\mathbf{v} = (\mathbf{B}^H\mathbf{v})^H\mathbf{B}^H\mathbf{v} = \left\|\mathbf{B}^H\mathbf{v}\right\|_2^2$$

e tale quadrato è sempre positivo, o al più nullo se la norma è nulla e ciò avviene se $\mathbf{v} = \mathbf{0}$.

□

Osservazione 189. Qualsiasi prodotto interno $(\cdot|\cdot)$ definito su $\mathbb{C}^n$ si può riscrivere come quello in proposizione 3.2.5, trovando l'opportuna matrice $\mathbf{A} = \mathbf{B}\mathbf{B}^H$. Questo fatto è un caso particolare (con $\mathscr{B} = \mathscr{E}$) del seguente risultato, valido per ogni spazio vettoriale euclideo di dimensione n.

Teorema 3.2.6. *Sia V uno spazio vettoriale euclideo con base $\mathscr{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Allora per ogni $\mathbf{v}, \mathbf{w} \in V$ risulta:*

$$(\mathbf{v}|\mathbf{w}) = \mathrm{C}_{\mathscr{B}}(\mathbf{v})^H \mathbf{A}\, \mathrm{C}_{\mathscr{B}}(\mathbf{w}) \tag{3.16}$$

con $\mathbf{A} = [(\mathbf{v}_i|\mathbf{v}_j)]$ (detta matrice di Gram*) matrice $n \times n$ del tipo $\mathbf{A} = \mathbf{B}\mathbf{B}^H$ e $\mathbf{B}$ invertibile.*

Osservazione 190. L'elemento di posto i, j della matrice di Gram, quindi, è il prodotto dei vettori i-esimo e j-esimo della base $\mathscr{B}$.

Dimostrazione. Se $\mathbf{v} = \sum_{i=1}^n \alpha_i\mathbf{v}_i, \mathbf{w} = \sum_{j=1}^n \beta_j\mathbf{v}_j \in V$, con $\mathbf{v}_i, \mathbf{v}_j \in \mathscr{B}$ sfruttando le proprietà del prodotto interno si ottiene:

$$\begin{aligned}
(\mathbf{v}|\mathbf{w}) = \left(\sum_{i=1}^n \alpha_i\mathbf{v}_i \middle| \sum_{j=1}^n \beta_j\mathbf{v}_j\right) &= (\alpha_1\mathbf{v}_1 + \ldots \alpha_n\mathbf{v}_n|\beta_1\mathbf{v}_1 + \ldots + \beta_n\mathbf{v}_n) \\
&= \overline{\alpha_1}(\mathbf{v}_1|\beta_1\mathbf{v}_1 + \ldots + \beta_{\mathbf{n}}\mathbf{v}_n) + \ldots + \overline{\alpha_n}(\mathbf{v}_n|\beta_1\mathbf{v}_1 + \ldots + \beta_{\mathbf{n}}\mathbf{v}_n) \\
&= \overline{\alpha_1}\beta_1(\mathbf{v}_1|\mathbf{v}_1) + \ldots \overline{\alpha_1}\beta_n(\mathbf{v}_1|\mathbf{v}_n) + \ldots + \overline{\alpha_n}\beta_1(\mathbf{v}_n|\mathbf{v}_1) + \ldots + \overline{\alpha_n}\beta_n(\mathbf{v}_n|\mathbf{v}_n) \\
&= \sum_{i=1}^n\sum_{j=1}^n \overline{\alpha_i}\beta_j(\mathbf{v}_i|\mathbf{v}_j)
\end{aligned}$$

Ora, una volta fissata la base (e quindi conosciuti α_i, β_j e i $\mathbf{v}_i, \mathbf{v}_j$), se si costruisce la matrice $n \times n$ $\mathbf{A} = [a_{ij}]$ dove $a_{ij} = (\mathbf{v}_i|\mathbf{v}_j)$, il prodotto interno può essere ulteriormente espresso come:

$$(\mathbf{v}|\mathbf{w}) = \sum_{i=1}^n\sum_{j=1}^n \overline{\alpha_i}\beta_j(\mathbf{v}_i|\mathbf{v}_j) = \sum_{i=1}^n\sum_{j=1}^n \overline{\alpha_i}\beta_j a_{ij} = \begin{bmatrix}\overline{\alpha_1} & \overline{\alpha_2} & \ldots & \overline{\alpha_n}\end{bmatrix} \mathbf{A} \begin{bmatrix}\beta_1 \\ \beta_2 \\ \ldots \\ \beta_n\end{bmatrix}$$

$$= \begin{bmatrix}\alpha_1 \\ \alpha_2 \\ \ldots \\ \alpha_n\end{bmatrix}^H \mathbf{A} \begin{bmatrix}\beta_1 \\ \beta_2 \\ \ldots \\ \beta_n\end{bmatrix} = \mathrm{C}_{\mathscr{B}}(\mathbf{v})^H \mathbf{A}\, \mathrm{C}_{\mathscr{B}}(\mathbf{w}) = (\mathrm{C}_{\mathscr{B}}(\mathbf{v})|\, \mathrm{C}_{\mathscr{B}}(\mathbf{w}))_{\mathbf{A}}$$

□

Proposizione 3.2.7 (Proprietà della matrice di Gram)**.** *Si ha che:*

1. $\mathbf{A}$ *è hermitiana;*
2. $\mathbf{x}^H\mathbf{A}\mathbf{x} \in \mathbb{R}_{>0}$ *(il prodotto interno di un vettore con se stesso è positivo) per ogni* $\mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{0}$*;*
3. $\mathbf{A}$ *ammette una fattorizzazione del tipo* $\mathbf{A} = \mathbf{B}\mathbf{B}^H$*, dove* $\mathbf{B}$ *è un'opportuna matrice invertibile* $n \times n$ *(come nella proposizione 3.2.5).*

Dimostrazione. Rispettivamente

1. si ha perché $\overline{a_{ij}} = a_{ji}$, dato che per la proprietà del prodotto interno:
$$\overline{a_{ij}} = \overline{(\mathbf{v}_i|\mathbf{v}_j)} = (\mathbf{v}_j|\mathbf{v}_i) = a_{ji} \quad \text{per ogni } i, j = 1, \dots, n$$
2. per ogni $\mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{0}$ esiste $\mathbf{v} \in V, \mathbf{v} \neq \mathbf{0}$ tale che $\mathrm{C}_{\mathscr{B}}(\mathbf{v}) = \mathbf{x}$ e pertanto
$$\mathbf{x}^H\mathbf{A}\mathbf{x} = (\mathrm{C}_{\mathscr{B}}(\mathbf{v})|\, \mathrm{C}_{\mathscr{B}}(\mathbf{v}))_{\mathbf{A}} \overset{(1)}{>} 0$$
dove (1) deriva dalle proprietà del prodotto interno;
3. la dimostrazione verrà data solo nel teorema 3.5 del capitolo 6 (dipende dalle due proprietà precedenti di $\mathbf{A}$)

TODO: fixme

□

Esempio 3.2.5. In particolare possiamo prendere $V = \mathbb{C}^n$ e la base canonica $\mathscr{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$. Allora $\mathrm{C}_{\mathscr{E}}(\mathbf{v}) = \mathbf{v}$ per ogni $\mathbf{v} \in V$. Quindi per ogni prodotto interno $(\cdot|\cdot)$ definito su $\mathbb{C}^n$ si ha
$$(\mathbf{v}|\mathbf{w})_{\mathbf{A}} = \mathbf{v}^H\mathbf{A}\mathbf{w} = \mathbf{v}^H\mathbf{I}\mathbf{w} = \mathbf{v}^H\mathbf{w} = (\mathbf{v}|\mathbf{w}) \quad \text{per ogni } \mathbf{v}, \mathbf{w} \in \mathbb{C}^n$$
con $\mathbf{A} = \mathbf{I}$ la matrice di Gram di $(\cdot|\cdot)$ rispetto alla base canonica $\mathscr{E}$ di $\mathbb{C}^n$ (dato che $(\mathbf{e}_i|\mathbf{e}_j) = \mathbf{e}_i^H\mathbf{e}_j = 1$ solo se $i = j$, ossia sulla diagonale, altrimenti è $= 0$).

Osservazione 191. Enunciamo ora la versione generale della disuguaglianza di Cauchy-Schwarz

Teorema 3.2.8 (Disuguaglianza di Cauchy-Schwarz)**.** *Sia* V *uno spazio vettoriale euclideo, e* $\mathbf{v}, \mathbf{w} \in V$
$$|(\mathbf{v}|\mathbf{w})| \leq \sqrt{(\mathbf{v}|\mathbf{v})}\sqrt{(\mathbf{w}|\mathbf{w})} \tag{3.17}$$

Dimostrazione. Proviamo il teorema nel caso in cui V sia uno spazio vettoriale complesso.
La disuguaglianza che vogliamo provare è vera se $(\mathbf{v}|\mathbf{w}) = 0$: infatti il prodotto $\sqrt{(\mathbf{v}|\mathbf{v})}\sqrt{(\mathbf{w}|\mathbf{w})}$ è positivo (se $\mathbf{v} \neq \mathbf{0}$ e $\mathbf{w} \neq \mathbf{0}$) o al più nullo (almeno uno dei due è $\mathbf{0}$), ma in entrambi i casi la disuguaglianza è soddisfatta.
Supponiamo invece che $(\mathbf{v}|\mathbf{w}) \neq 0$. In particolare per la proprietà del prodotto interno possiamo riscrivere $\mathbf{v} = [v_1 \dots v_n]^T$ come combinazione lineare di elementi di $\mathscr{E}$, $\mathbf{v} = v_1\mathbf{e}_1 + \dots + v_n\mathbf{e}_n$ per cui applicando le consuete proprietà:
$$0 \neq (v_1\mathbf{e}_1 + \dots + v_n\mathbf{e}_n|\mathbf{w}) = \overline{v_1}(\mathbf{e}_1|\mathbf{w}) + \dots + \overline{v_n}(\mathbf{e}_n|\mathbf{w})$$
e ciò implica che non può essere che $v_1 = \dots = v_n = 0$ per cui possiamo supporre che $\mathbf{v} \neq \mathbf{0}$. Pertanto si ha che $(\mathbf{v}|\mathbf{v}) > 0$ (per la terza proprietà del prodotto

interno).
Sia $\alpha \in \mathbb{C}$; dalle proprietà del prodotto interno:

$$\begin{aligned}0 &\leq (\mathbf{v}+\alpha\mathbf{w}|\mathbf{v}+\alpha\mathbf{w}) = 1(\mathbf{v}+\alpha\mathbf{w}|\mathbf{v}) + \alpha(\mathbf{v}+\alpha\mathbf{w}|\mathbf{w}) \\ &= 1\left[1(\mathbf{v}|\mathbf{v}) + \overline{\alpha}(\mathbf{w}|\mathbf{v})\right] + \alpha\left[1(\mathbf{v}|\mathbf{w}) + \overline{\alpha}(\mathbf{w}|\mathbf{w})\right] \\ &= (\mathbf{v}|\mathbf{v}) + \alpha(\mathbf{v}|\mathbf{w}) + \overline{\alpha}(\mathbf{w}|\mathbf{v}) + \alpha\overline{\alpha}(\mathbf{w}|\mathbf{w}) \\ &= (\mathbf{v}|\mathbf{v}) + \alpha(\mathbf{v}|\mathbf{w}) + \overline{\alpha}(\mathbf{w}|\mathbf{v}) + |\alpha|^2(\mathbf{w}|\mathbf{w}) \end{aligned} \tag{3.18}$$

In particolare se si pone $\alpha = -\frac{(\mathbf{v}|\mathbf{v})}{(\mathbf{v}|\mathbf{w})}$ dalle proprietà del prodotto interno si ottiene che il coniugato $\overline{\alpha}$:

$$\overline{\alpha} = -\frac{\overline{(\mathbf{v}|\mathbf{v})}}{\overline{(\mathbf{v}|\mathbf{w})}} = -\frac{(\mathbf{v}|\mathbf{v})}{(\mathbf{w}|\mathbf{v})}$$

(ricordando che il coniugato di un rapporto è rapporto dei coniugati) per cui

$$|\alpha|^2 = \alpha\overline{\alpha} = \frac{(\mathbf{v}|\mathbf{v})}{(\mathbf{v}|\mathbf{w})}\frac{(\mathbf{v}|\mathbf{v})}{(\mathbf{w}|\mathbf{v})} = \frac{(\mathbf{v}|\mathbf{v})}{(\mathbf{v}|\mathbf{w})}\frac{(\mathbf{v}|\mathbf{v})}{\overline{(\mathbf{v}|\mathbf{w})}} = \frac{(\mathbf{v}|\mathbf{v})^2}{|(\mathbf{v}|\mathbf{w})|^2}$$

Sostituendo opportunamente in 3.18:

$$0 \leq (\mathbf{v}|\mathbf{v}) - \frac{(\mathbf{v}|\mathbf{v})}{(\mathbf{v}|\mathbf{w})}(\mathbf{v}|\mathbf{w}) - \frac{(\mathbf{v}|\mathbf{v})}{(\mathbf{w}|\mathbf{v})}(\mathbf{w}|\mathbf{v}) + \frac{(\mathbf{v}|\mathbf{v})^2}{|(\mathbf{v}|\mathbf{w})|^2}(\mathbf{w}|\mathbf{w}) = -(\mathbf{v}|\mathbf{v}) + \frac{(\mathbf{v}|\mathbf{v})^2}{|(\mathbf{v}|\mathbf{w})|^2}(\mathbf{w}|\mathbf{w})$$

e quindi

$$(\mathbf{v}|\mathbf{v}) \leq \frac{(\mathbf{v}|\mathbf{v})^2}{|(\mathbf{v}|\mathbf{w})|^2}(\mathbf{w}|\mathbf{w})$$

La disuguaglianza di Cauchy-Schwarz segue ora da quest'ultima moltiplicandone ambo i membri per il numero reale positivo $|(\mathbf{v}|\mathbf{w})|^2/(\mathbf{v}|\mathbf{v})$ ed estraendo la radice quadrata. □

Osservazione 192. Le disuguaglianze di Cauchy-Schwarz enunciate in 3.1.3 e 3.1.10 si ottengono da quella appena dimostrata specificando i prodotti interni di proposizione 3.2.3 e 3.2.4: basta osservare che $(\mathbf{v}|\mathbf{v}) = \|\mathbf{v}\|_2^2$ per ogni $\mathbf{v} \in \mathbb{C}^n$ e $(f|f) = \|f\|_2^2$ per ogni $f \in \mathscr{C}([a,b])$.

Osservazione 193. Alcune norme (non tutte) si possono ottenere come prodotto interno di un elemento con se stesso. Ad esempio le norme $\|\cdot\|_2$ in $\mathbb{C}^n$ (proposizione 3.1.2) e $\|\cdot\|_2$ in $\mathscr{C}([a,b])$ (proposizione 3.1.9) si possono definire a partire dai prodotti interni definiti nelle proposizioni 3.2.3 e 3.2.4, ponendo:

$$\begin{aligned} \|\mathbf{v}\|_2 &= \sqrt{(\mathbf{v}|\mathbf{v})} \quad \text{per ogni } \mathbf{v} \in \mathbb{C}^n, \text{ se } (\mathbf{v}|\mathbf{v}) = \mathbf{v}^H\mathbf{v} \\ \|f\|_2 &= \sqrt{(f|f)} \quad \text{per ogni } \mathscr{C}([a,b]), \text{ se } (f|f) = \int_a^b [f(t)]^2\, dt \end{aligned}$$

Per tali motivi si dice che la norma euclidea di $\mathbb{C}^n$ è indotta dai prodotti interni di proposizione 3.2.3 mentre $\|\cdot\|_2$ di $\mathscr{C}([a,b])$ è indotta dal prodotto interno di proposizione 3.2.4.

Proposizione 3.2.9 (Norma indotta dal prodotto interno)**.** *In generale se V è uno spazio vettoriale euclideo con prodotto interno $(\cdot|\cdot)$, si chiama così la funzione*

$$\|\cdot\| : V \to \mathbb{R}_{\geq 0} \quad \textit{definita da } \|\mathbf{v}\| = \sqrt{(\mathbf{v}|\mathbf{v})} \tag{3.19}$$

Tale funzione è (appunto) una norma su V.

Dimostrazione. La verifica che essa soddisfa le proprietà delle norme si ottiene ricalcando quella che abbiamo fatto nelle proposizioni 3.1.2 e 3.1.8, tenendo conto delle proprietà del prodotto interno e della disuguaglianza di Schwarz. □

Esempio 3.2.6. Sia $V = M_2(\mathbb{C})$. Si verifichi che $(\cdot|\cdot) : V \times V \to \mathbb{C}$ definita da

$$\left(\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}, \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}\right) = \sum_{i=1}^{4} \overline{a_i} b_i$$

è un prodotto interno. Inoltre sia $\|\cdot\| : M_2(\mathbb{C}) \to \mathbb{R}_{\geq 0}$ la norma indotta da tale prodotto interno. Si calcoli $\left\| \begin{bmatrix} 1 & i \\ 1+i & -2+3i \end{bmatrix} \right\|$.

Per la prima proprietà del prodotto interno si ha che

$$(\mathbf{w}|\mathbf{v}) = \sum_{i=1}^{4} \overline{b_i} a_i$$

da cui $\overline{(\mathbf{w}|\mathbf{v})} = \overline{\sum_{i=1}^{4} \overline{b_i} a_i} = \sum_{i=1}^{4} = \overline{a_i} b_i = (\mathbf{v}|\mathbf{w})$. In vista della seconda proprietà se definiamo

$$\alpha\mathbf{w} + \beta\mathbf{z} = \begin{bmatrix} \alpha b_1 + \beta c_1 & \alpha b_2 + \beta c_2 \\ \alpha b_3 + \beta c_3 & \alpha b_4 + \beta c_4 \end{bmatrix}$$

allora

$$\begin{aligned}(\mathbf{v}|\alpha\mathbf{w} + \beta\mathbf{z}) &= \sum_{i=1}^{4} \overline{a_i}(\alpha b_i + \beta c_i) = \sum_{i=1}^{4} \alpha \overline{a_i} b_i + \sum_{i=1}^{4} \beta \overline{a_i} c_i = \alpha \sum_{i=1}^{4} \overline{a_i} b_i + \beta \sum_{i=1}^{4} \overline{a_i} c_i \\ &= \alpha(\mathbf{v}|\mathbf{w}) + \beta(\mathbf{v}|\mathbf{z})\end{aligned}$$

Per la terza proprietà se $\mathbf{v} \neq \mathbf{0}$

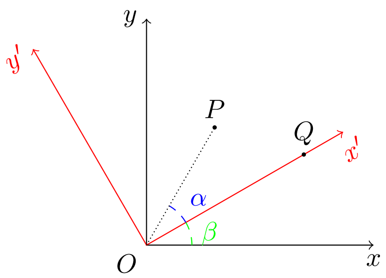
$$(\mathbf{v}|\mathbf{v}) = \sum_{i=1}^{4} \overline{a_i} a_i = \sum_{i=1}^{4} |a_i|^2 > 0$$

poiché somma di positivi.
Infine per il calcolo della norma indotta dal prodotto interno si ha che:

$$\|\mathbf{v}\| = \sqrt{\sum_{i=1}^{4} |v_i|^2}$$

dunque

$$\left\| \begin{bmatrix} 1 & i \\ 1+i & -2+3i \end{bmatrix} \right\| = \sqrt{|1|^2 + |i|^2 + |1+i|^2 + |-2+3i|^2} = \sqrt{1+1+2+13} = \sqrt{17}$$



3.2.1 Angolo fra due vettori

Sia Oxy un sistema di riferimento ortogonale e monometrico nel piano, e P e Q due punti del piano (diversi da O); ruotiamo in senso antiorario dell'angolo β il sistema di riferimento, ottenendo $Ox'y'$, in modo che il semiasse positivo delle ascisse x' contenga il punto Q.
Se $\mathbf{v} = [x_P\, y_P]^T$ e $\mathbf{w} = [x_Q\, y_Q]^T$ sono i vettori delle coordinate di P e Q rispetto al sistema di riferimento Oxy e $\mathbf{v}' = [x'_P\, y'_P]^T$ e $\mathbf{w}' = [x'_Q\, y'_Q]^T$ quelli rispetto a $Ox'y'$, la relazione che intercorre tra le coordinate (ad esempio di P) tra i due sistemi è:

$$\begin{cases} x_P = x'_P \cos\beta - y'_P \sin\beta \\ y_P = x'_P \sin\beta + y'_P \cos\beta \end{cases} \rightarrow \begin{bmatrix} x_P \\ y_P \end{bmatrix} = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} x'_P \\ y'_P \end{bmatrix} \rightarrow \mathbf{v} = \mathbf{R}_\beta \mathbf{v}'$$

come applicazioni della trigonometria di base. In generale, quindi:

$$\mathbf{v} = \mathbf{R}_\beta \mathbf{v}' \qquad \mathbf{w} = \mathbf{R}_\beta \mathbf{w}' \tag{3.20}$$

dove $\mathbf{R}_\beta = \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix}$ è detta matrice di rotazione dell'angolo β.
Si ha che:

$$\begin{aligned} \mathbf{R}_\beta^T \mathbf{R}_\beta &= \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix} \begin{bmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\beta + \sin^2\beta & \cos\beta(-\sin\beta) + \sin\beta\cos\beta \\ -\sin\beta\cos\beta + \cos\beta\sin\beta & \sin^2\beta + \cos^2\beta \end{bmatrix} \\ &= \mathbf{I}_2 \end{aligned}$$

dato che $\cos^2\beta + \sin^2\beta = 1$. Notiamo come in termini di vettori, il "prodotto" del primo punto (trasposto) per il secondo coincidano, nonostante il cambio del sistema di riferimento:

$$\mathbf{v}^T \mathbf{w} = (\mathbf{R}_\beta \mathbf{v}')^T (\mathbf{R}_\beta \mathbf{w}') = (\mathbf{v}')^T \underbrace{\mathbf{R}_\beta^T \mathbf{R}_\beta}_{\mathbf{I}_2} \mathbf{w}' = (\mathbf{v}')^T \mathbf{w}' \tag{3.21}$$

D'altra parte, se α è l'angolo tra la retta passante per O e Q e la retta passante per O e P, allora, indicando con $|OP|$ e $|OQ|$ le misure dei segmenti OP e OQ rispettivamente, si ha anche

$$\mathbf{v}' = \begin{bmatrix} |OP|\cos\alpha \\ |OP|\sin\alpha \end{bmatrix} \quad \text{e} \quad \mathbf{w}' = \begin{bmatrix} |OQ| \\ 0 \end{bmatrix}$$

e quindi

$$(\mathbf{v}')^T\mathbf{w}' = \begin{bmatrix} |OP|\cos\alpha \\ |OP|\sin\alpha \end{bmatrix}^T \begin{bmatrix} |OQ| \\ 0 \end{bmatrix} = |OP|\,|OQ|\cos\alpha = \|\mathbf{v}\|_2\,\|\mathbf{w}\|_2\cos\alpha \quad (3.22)$$

Poiché $\mathbf{v}, \mathbf{w} \neq \mathbf{0}$ allora $\|\mathbf{v}\|_2, \|\mathbf{w}\|_2 \neq 0$ e da 3.21 e 3.22 si ricava la seguente relazione:

$$\cos\alpha = \frac{\mathbf{v}^T\mathbf{w}}{\|\mathbf{v}\|_2\,\|\mathbf{w}\|_2}, \qquad \text{per ogni } \mathbf{v}, \mathbf{w} \in \mathbb{R}^2 \setminus \{\mathbf{0}\} \quad (3.23)$$

In particolare se P e Q sono punti del piano appartenenti alla circonferenza di centro O e raggio 1, si ottiene:

$$\cos\alpha = \mathbf{v}^T\mathbf{w}$$

Pertanto il prodotto interno standard $f(\mathbf{v}, \mathbf{w}) = (\mathbf{v}|\mathbf{w}) = \mathbf{v}^T\mathbf{w}$ su $\mathbb{R}^2 \times \mathbb{R}^2$, calcola il coseno dell'angolo compreso tra due vettori di lunghezza unitaria.
Ritornando alla più generica (in termini di lunghezza dei vettori coinvolti) 3.23 possiamo impiegarla per introdurre una generalizzazione del concetto di angolo tra due vettori.

Definizione 3.2.3. Siano V uno spazio vettoriale euclideo e $\|\cdot\|$ la norma indotta dal prodotto interno $(\cdot|\cdot)$ definito su V. Allora per ogni $\mathbf{v}, \mathbf{w} \in V \setminus \{\mathbf{0}\}$ si pone:

$$\cos\widehat{\mathbf{v}\mathbf{w}} = \frac{(\mathbf{v}|\mathbf{w})}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|} \quad (3.24)$$

$\cos\widehat{\mathbf{v}\mathbf{w}}$ è un numero complesso (in quanto rapporto tra $(\mathbf{v}|\mathbf{w}) \in \mathbb{C}$ e due reali), e per la disuguaglianza di Schwarz ha modulo ≤ 1.
Solo nel caso in cui V sia uno spazio vettoriale reale ($(\mathbf{v}|\mathbf{w})$ è reale e si può applicare arccos) si può porre:

$$\alpha = \widehat{\mathbf{v}\mathbf{w}} = \arccos\left(\frac{(\mathbf{v}|\mathbf{w})}{\|\mathbf{v}\| \cdot \|\mathbf{w}\|}\right) \quad (3.25)$$

e α si chiama *angolo* tra i vettori $\mathbf{v}$ e $\mathbf{w}$.

Esempio 3.2.7. Sia V lo spazio vettoriale reale delle matrici 2×2 reali simmetriche e il prodotto interno considerato sia

$$\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right) = aa' + 2bb' + cc'$$

Trovare l'angolo tra

$$\begin{bmatrix} 1 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{bmatrix}$$

e $\mathbf{I}_2$. Si ha che

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}|\mathbf{x})} = \sqrt{a^2 + 2b^2 + c^2}$$

quindi

$$\|\mathbf{v}\| = \sqrt{1 + 2\frac{1}{2} + \frac{1}{2}} = \sqrt{\frac{5}{2}}$$
$$\|\mathbf{w}\| = \sqrt{1^2 + 0^2 + 0^2} = 1$$
$$(\mathbf{v}|\mathbf{w}) = 1 \cdot 1 + 2\frac{1}{\sqrt{2}}0 + \frac{1}{\sqrt{2}}0 = 1$$

pertanto

$$\cos \widehat{\mathbf{vw}} = \frac{1}{\sqrt{5/2}} = \sqrt{2/5}$$

da cui l'angolo

$$\widehat{\mathbf{vw}} = \arccos \sqrt{2/5}$$

Esempio 3.2.8. Sia V spazio vettoriale reale delle matrici 2×2 reali simmetriche e il prodotto interno considerato sia

$$\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}, \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \right) = aa' + 2bb' + cc'$$

Sia

$$W = \left\langle \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\rangle$$

Trovare $W^{\perp}$. Se $\mathbf{z} = \begin{bmatrix} x & y \\ y & z \end{bmatrix}$. Deve essere a tal fine $(\mathbf{v}|\mathbf{z}) = 0$ ossia

$$\begin{aligned} &1 \cdot x + 2 \cdot 1 \cdot y + 1 \cdot w = 0 \\ &x + 2y + w = 0 \\ &x = -2y - w \end{aligned}$$

quindi le matrici che compongono $W^{\perp}$ debbono essere del tipo

$$\mathbf{z} = \begin{bmatrix} -2y - w & y \\ w & z \end{bmatrix}$$

3.3 Ortogonalità e proiezioni ortogonali

Definizione 3.3.1 (Vettore ortogonale ad un altro vettore)**.** Sia V uno spazio vettoriale euclideo. Si dice che il vettore $\mathbf{w} \in V$ è ortogonale al vettore $\mathbf{u} \in V$ e si scrive $\mathbf{w} \perp \mathbf{u}$, se alternativamente:

- il coseno dell'angolo compreso tra $\mathbf{w}$ e $\mathbf{u}$ è 0 (e quindi $(\mathbf{w}|\mathbf{v}) = 0$ per la definizione 3.24);
- almeno uno tra $\mathbf{w}$ ed $\mathbf{u}$ è il vettore nullo.

Esempio 3.3.1. In $\mathbb{C}^3$ e $\mathbb{C}^2$, considerando il prodotto interno standard si ha che $\mathbf{a}_1 \perp \mathbf{a}_2$, $\mathbf{b}_1 \perp \mathbf{b}_2$ con

$$\mathbf{a}_1 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \quad \mathbf{b}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Proposizione 3.3.1. *L'ortogonalità è riflessiva se* $\mathbf{w}, \mathbf{u} \in V$

$$\mathbf{w} \perp \mathbf{u} \iff \mathbf{u} \perp \mathbf{w}$$

e si dice che i due vettori sono tra loro ortogonali.

Dimostrazione.

$$\mathbf{w} \perp \mathbf{u} \iff (\mathbf{w}|\mathbf{u}) = 0 \overset{(1)}{\iff} (\mathbf{u}|\mathbf{w}) = 0 \iff \mathbf{u} \perp \mathbf{w}$$

dove in (1) si prende il coniugato di entrambi i membri e si applicano le proprietà del prodotto interno. □

Definizione 3.3.2 (Vettore ortogonale a sottospazio)**.** Il vettore $\mathbf{w} \in V$ è ortogonale al sottospazio U di V, e si scrive $\mathbf{w} \perp U$, se $\mathbf{w}$ è ortogonale a ogni vettore $\mathbf{u} \in U$.

Proposizione 3.3.2. *Per verificare che un vettore $\mathbf{w} \in V$ sia ortogonale a un sottospazio U di V, è sufficiente verificare che sia ortogonale ad un insieme di generatori di U.*

Dimostrazione. Se $\{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ è un insieme di generatori di U, allora

$$\begin{aligned}
\mathbf{w} \in U^\perp &\overset{(1)}{\iff} 0 = (\mathbf{w}|\alpha_1\mathbf{u}_1 + \dots + \alpha_k\mathbf{u}_k) \overset{(2)}{=} \alpha_1(\mathbf{w}|\mathbf{u}_1) + \dots + \alpha_k(\mathbf{w}|\mathbf{u}_k) \quad \forall \alpha_1, \dots, \alpha_k \\
&\iff (\mathbf{w}|\mathbf{u}_i) = 0 \quad \forall i = 1, \dots, k \\
&\iff \mathbf{w} \perp \mathbf{u}_i \quad \forall i = 1, \dots, k
\end{aligned}$$

dove (1) è dovuta al fatto che il coseno che $\mathbf{w}$ forma con un elemento di U è nullo e (2) dalle proprietà del prodotto interno. □

Definizione 3.3.3 (Complemento ortogonale di U in V)**.** È l'insieme di tutti i vettori di V ortogonali a U; si indica con $U^\perp$.

Proposizione 3.3.3. *$U^\perp$ è un sottospazio di V*

Dimostrazione. Innanzitutto $\mathbf{0}$ è ortogonale a qualsiasi elemento di U, quindi $0 \in U^\perp$; inoltre per ogni $\mathbf{w}_1, \mathbf{w}_2 \in U^\perp$, $\mathbf{u} \in U$ e α_1, α_2 scalari si ha:

$$(\mathbf{u}|\alpha_1\mathbf{w}_1 + \alpha_2\mathbf{w}_2) = \alpha_1(\mathbf{u}|\mathbf{w}_1) + \alpha_2(\mathbf{u}|\mathbf{w}_2) = \alpha_1 \cdot 0 + \alpha_2 \cdot 0 = 0$$

Ossia una combinazione lineare di elementi ortogonali a U è ancora ortogonale a U, e quindi appartiene ancora a $U^\perp$. □

Osservazione 194. Ogni vettore è ortogonale a quello nullo; il vettore nullo è l'unico vettore ortogonale a tutti gli altri. Rispettivamente:

$$\{\mathbf{0}\}^\perp = V, \qquad V^\perp = \{\mathbf{0}\}$$

Esempio 3.3.2. Siano $V = \mathbb{R}^3$, dotato del prodotto interno standard $(\mathbf{v}|\mathbf{z}) = \mathbf{v}^T\mathbf{z}$, $\mathscr{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base canonica di V, e $U = \langle \mathbf{e}_1 \rangle$. Si ha che:

$$\begin{aligned}
U^\perp &= \left\{\mathbf{w} \in \mathbb{R}^3 : \mathbf{w} \perp \mathbf{e}_1\right\} = \left\{\mathbf{w} \in \mathbb{R}^3 : \mathbf{w}^T\mathbf{e}_1 = 0\right\} \\
&= \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 : \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 0 \right\} = \left\{ \begin{bmatrix} 0 \\ y \\ z \end{bmatrix} : y, z \in \mathbb{R} \right\} \\
&= \langle \mathbf{e}_2; \mathbf{e}_3 \rangle
\end{aligned}$$

Per una interpretazione geometrica nello spazio dove si fissa un riferimento $Oxyz$, in questo caso ad $\langle \mathbf{e}_1 \rangle$ corrisponde l'insieme dei punti dell'asse x, mentre

a $U^{\perp}$ corrisponde l'insieme dei punti del piano Oyz, che è il piano perpendicolare all'asse x contenente O.
Analogamente se $Z = \langle \mathbf{e}_1; \mathbf{e}_3 \rangle$:

$$\begin{aligned} Z^{\perp} &= \left\{ \mathbf{w} \in \mathbb{R}^3 : \mathbf{w} \perp \mathbf{e}_1 \wedge \mathbf{w} \perp \mathbf{e}_3 \right\} = \left\{ \mathbf{w} \in \mathbb{R}^3 : \mathbf{w}^T \mathbf{e}_1 = 0 = \mathbf{w}^T \mathbf{e}_3 \right\} \\ &= \left\{ \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} : y \in \mathbb{R} \right\} = \langle \mathbf{e}_2 \rangle \end{aligned}$$

Qui Z corrisponde all'insieme dei punti del piano Oxz e a $Z^{\perp}$ l'insieme dei punti dell'asse y ossia la retta passante per O perpendicolare al piano Oxz.

Osservazione 195. In generale, se V è un sottospazio di dimensione 1 di $\mathbb{R}^3$ corrispondente all'insieme dei punti di una retta r passante per O il suo complemento ortogonale $V^{\perp}$ è il sottospazio di dimensione 2 di R^3 corrispondente all'insieme dei punti del piano perpendicolare a r e contenente O.
Viceversa se V è un sottospazio di dimensione 2 di $\mathbb{R}^3$ corrispondente all'insieme dei punti di un piano π contenente O, il suo complemento ortogonale $V^{\perp}$ è il sottospazio di dimensione 1 di R^3 corrispondente ai punti della retta r passante per O e perpendicolare a π.

Esempio 3.3.3. Sia $U \subseteq \mathbb{R}^4$ il sottospazio generato da $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$ e $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$.

Si trovi una base di $U^{\perp}$. Si ha che:

$$U^{\perp} = \left\{ \mathbf{w} \in \mathbb{R}^4 : \mathbf{w} \perp \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \wedge \mathbf{w} \perp \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Se $\mathbf{w} = \begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix}$ allora deve essere $\begin{cases} 1x + 1y + 1w + 0z = 0 \\ 0x + 1y + 0w + 1z = 0 \end{cases}$ $\begin{cases} x + y + w = 0 \\ y + z = 0 \end{cases}$ $\begin{cases} x = z - w \\ y = -z \end{cases}$
e dunque $\mathbf{w}$ deve essere del tipo

$$\mathbf{w} = \begin{bmatrix} z - w \\ -z \\ w \\ z \end{bmatrix} = z \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

e dunque una base è

$$U^{\perp} = \left\langle \begin{bmatrix} 1 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\rangle$$

Proposizione 3.3.4. *Sia* $\mathbf{A}$ *una matrice complessa* $m \times n$*. Allora:*

1. $\mathrm{N}(\mathbf{A})$ *è il complemento ortogonale di* $\mathrm{C}\left(\mathbf{A}^H\right)$ *in* $\mathbb{C}^n$*;*

2. *$\mathrm{N}\left(\mathbf{A}^H\right)$ è il complemento ortogonale di $\mathrm{C}(\mathbf{A})$ in $\mathbb{C}^m$;*

Osservazione 196. Per cui ad esempio gli elementi di $\mathrm{N}(\mathbf{A})$ sono ortogonali a quelli di $\mathrm{C}\left(\mathbf{A}^H\right)$ e viceversa.

Dimostrazione. Suddividiamo $\mathbf{A}$ in blocchi riga ed evidenziamo le colonne di $\mathbf{A}^H$

$$\mathbf{A} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix}, \quad \mathbf{A}^H = \begin{bmatrix} \overline{\mathbf{r}_1} & \overline{\mathbf{r}_2} & \dots & \overline{\mathbf{r}_m} \end{bmatrix}, \quad \mathbf{r}_i \in \mathbb{C}^m, i = 1, \dots, m$$

Allora se $\mathbf{v} \in V$ si ha:

$$\begin{aligned} \mathbf{v} \in \mathrm{N}(\mathbf{A}) &\iff \mathbf{0} = \mathbf{A}\mathbf{v} = \begin{bmatrix} \mathbf{r}_1^T \\ \mathbf{r}_2^T \\ \dots \\ \mathbf{r}_m^T \end{bmatrix} \mathbf{v} = \begin{bmatrix} \mathbf{r}_1^T\mathbf{v} \\ \mathbf{r}_2^T\mathbf{v} \\ \dots \\ \mathbf{r}_m^T\mathbf{v} \end{bmatrix} \\ &\overset{(1)}{\iff} 0 = \mathbf{r}_i^T\mathbf{v} \overset{(2)}{=} \overline{\mathbf{r}_i}^H\mathbf{v} \overset{(3)}{=} (\overline{\mathbf{r}_i}|\mathbf{v}) \quad i = 1, \dots, m \\ &\overset{(4)}{\iff} \mathbf{v} \in \mathrm{C}\left(\mathbf{A}^H\right)^\perp \end{aligned}$$

dove in (1) ci si focalizza sul singolo elemento del vettore risultato ($\mathbf{A}\mathbf{v} = \mathbf{0}$ solo se tutti i suoi elementi sono 0); in (2) si sostituisce $\mathbf{r}_i^T = \overline{\mathbf{r}_i}^H$ (legittimo poiché con il coniugio annulliamo quello dell'H-trasposizione, che torna una trasposizione normale); in (3) si applica la definizione di prodotto interno standard; in (4) si ha che $(\overline{\mathbf{r}_i}|\mathbf{v}) = 0$ se e solo se $\cos\widehat{\mathbf{r}_i\mathbf{v}} = 0$ ossia se $\mathbf{v}$ è ortogonale ai generatori $\mathbf{r}_i$ dello spazio $\mathrm{C}\left(\mathbf{A}^H\right)$, quindi per proposizione 3.3.2 $\mathbf{v} \in \mathrm{C}\left(\mathbf{A}^H\right)^\perp$.
Per quanto appena dimostrato, considerando poi $\mathbf{A}^H$ al posto di $\mathbf{A}$ si ottiene che $\mathrm{N}\left(\mathbf{A}^H\right)$ è il complemento ortogonale in $\mathbb{C}^m$ di $\mathrm{C}\left((\mathbf{A}^H)^H\right) = \mathrm{C}(\mathbf{A})$. □

Esempio 3.3.4. In particolare se $U = \langle \mathbf{u}_1; \dots; \mathbf{u}_n \rangle$ è un sottospazio di $\mathbb{C}^m$ dotato del prodotto interno standard, la matrice $m \times n$ $\mathbf{A} = [\mathbf{u}_1\, \mathbf{u}_2 \dots \mathbf{u}_n]$ è tale che $U = \mathrm{C}(\mathbf{A})$ (per definizione dello spazio colonne dato che le colonne di A sono i vettori di U), e pertanto si ha che

$$U^\perp = [\mathrm{C}(\mathbf{A})]^\perp = \mathrm{N}\left(\mathbf{A}^H\right)$$

per il punto 2 della proposizione 3.3.4.

Proposizione 3.3.5 (Proprietà complemento ortogonale di un sottospazio)**.** *Sia V uno spazio vettoriale euclideo:*

1. *$U \cap U^\perp = \{\mathbf{0}\}$ per ogni sottospazio U di V;*

2. *se U e W sono sottospazi di V, e U è contenuto in W, allora $W^\perp$ è contenuto in $U^\perp$;*

3. *se U_1 e U_2 sono sottospazi di V allora $(U_1 + U_2)^\perp = U_1^\perp \cap U_2^\perp$*

Dimostrazione. Tentativo di dimostrazione; rispettivamente:

1. essendo U un sottospazio $\mathbf{0} \in U$; inoltre $\mathbf{0}$ è perpendicolare a qualsiasi vettore di U quindi $\mathbf{0} \in U^\perp$. E qualsiasi altro vettore $\mathbf{v}$ non può sia appartenere a U che a $U^\perp$ per il semplice fatto che non è perpendicolare con se stesso, dato che $(\mathbf{v}|\mathbf{v}) > 0$

2. poiché ci possono essere vettori $\mathbf{v}$ che non sono perpendicolari a tutti gli elementi di W ($\mathbf{v} \notin W^\perp$) ma lo sono a tutti quelli di U ($\mathbf{v} \in U^\perp$) si ha che $W^\perp \subseteq U^\perp$

3. un singolo vettore $\mathbf{v}$ è ortogonale a tutti i vettori di $U_1 + U_2$ (ossia $\mathbf{v} \in (U_1 + U_2)^\perp$) se e solo se è ortogonale ai vettori di U_1 (quindi $\mathbf{v} \in U_1^\perp$) che ai vettori di U_2 (quindi $\mathbf{v} \in U_2^\perp$). Ossia appartiene all'intersezione dei due insiemi: $\mathbf{v} \in U_1^\perp \cap U_2^\perp$

□

Esempio 3.3.5. Sia $V = \mathbb{C}^3$ dotato del prodotto interno standard $(\mathbf{v}|\mathbf{w}) = \mathbf{v}^H\mathbf{w}$. Abbiamo visto nell'esempio 3.3.2 che

$$\langle \mathbf{e}_1; \mathbf{e}_3 \rangle^\perp = \left\{ \begin{bmatrix} x \\ 0 \\ z \end{bmatrix} : x, z \in \mathbb{C} \right\}^\perp = \langle \mathbf{e}_2 \rangle = \left\{ \begin{bmatrix} 0 \\ y \\ 0 \end{bmatrix} : y \in \mathbb{C} \right\} \subseteq \langle \mathbf{e}_2; \mathbf{e}_3 \rangle = \left\{ \begin{bmatrix} 0 \\ y \\ z \end{bmatrix} : y, z \in \mathbb{C} \right\} = \langle \mathbf{e}_1 \rangle^\perp$$

È facile verificare specularmente (per $\mathbf{e}_3$ al posto di $\mathbf{e}_1$) che $\langle \mathbf{e}_3 \rangle^\perp = \langle \mathbf{e}_1; \mathbf{e}_2 \rangle$. I vettori perpendicolari sia a $\begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}$ che a $\begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix}$ ossia

$$(\langle \mathbf{e}_1 \rangle + \langle \mathbf{e}_3 \rangle)^\perp = \langle \mathbf{e}_1; \mathbf{e}_3 \rangle^\perp = \langle \mathbf{e}_2 \rangle = \langle \mathbf{e}_2; \mathbf{e}_3 \rangle \cap \langle \mathbf{e}_1; \mathbf{e}_2 \rangle = \langle \mathbf{e}_1 \rangle^\perp \cap \langle \mathbf{e}_3 \rangle^\perp$$

dove l'uguaglianza tra il primo e l'ultimo membro è una versione particolare del punto 3 di proposizione 3.3.5.

Osservazione 197. Il prossimo obiettivo è dimostrare che ogni spazio vettoriale euclideo è somma diretta (definizione 2.1.8) di un suo qualunque sottospazio e del suo completamento ortogonale. La dimostrazione sarà per induzione sulla dimensione del sottospazio; il lemma seguente fornisce il primo passo dell'induzione.

Lemma 3.3.6. *Per ogni* $\mathbf{u} \in V$*,* $\mathbf{u} \neq \mathbf{0}$ *si ha*

$$V = \langle \mathbf{u} \rangle \oplus \langle \mathbf{u} \rangle^\perp \tag{3.26}$$

Dimostrazione. Per ogni $\mathbf{v} \in V$ vogliamo spezzare $\mathbf{v}$ nella somma di due vettori, uno appartenente allo spazio generato da $\mathbf{u}$ e l'altro ortogonale ad esso. Iniziamo notando che il vettore

$$\mathbf{v} - \frac{(\mathbf{u}|\mathbf{v})}{(\mathbf{u}|\mathbf{u})}\mathbf{u} \in \langle \mathbf{u} \rangle^\perp$$

è perpendicolare a $\mathbf{u}$, poiché l'angolo formato con $\mathbf{u}$ è nullo, in quanto:

$$\left(\mathbf{v} - \frac{(\mathbf{u}|\mathbf{v})}{(\mathbf{u}|\mathbf{u})}\mathbf{u}\middle|\mathbf{u}\right) = (\mathbf{v}|\mathbf{u}) - \frac{(\mathbf{v}|\mathbf{u})}{(\mathbf{u}|\mathbf{u})}(\mathbf{u}|\mathbf{u}) = 0$$

Inoltre si noti che

$$\frac{(\mathbf{u}|\mathbf{v})}{(\mathbf{u}|\mathbf{u})}\mathbf{u} \in \langle \mathbf{u} \rangle \tag{3.27}$$

poiché stiamo moltiplicando $\mathbf{u}$ per una costante. Pertanto il generico vettore $\mathbf{v}$ può essere riscritto algebricamente come

$$\mathbf{v} = \frac{(\mathbf{u}|\mathbf{v})}{(\mathbf{u}|\mathbf{u})}\mathbf{u} + \left(\mathbf{v} - \frac{(\mathbf{u}|\mathbf{v})}{(\mathbf{u}|\mathbf{u})}\mathbf{u}\right)$$

(ci si è limitati ad aggiungere e sottrarre 3.27 al secondo membro). La tesi (ossia il fatto che l'unico vettore che appartiene a $\langle \mathbf{u} \rangle$ e $\langle \mathbf{u} \rangle^{\perp}$ sia $\mathbf{0}$, per cui si usi $\oplus$) segue dal punto 1 della proposizione 3.3.5. □

Osservazione 198. Questo era il passo base, dato che $\langle \mathbf{u} \rangle$ ha dimensione 1; il risultato del lemma 3.3.6 si estende da sottospazi di dimensione 1 di V a sottospazi di dimensione qualunque

Proposizione 3.3.7. *Per ogni sottospazio U di uno spazio euclideo V si ha*

$$V = U \oplus U^{\perp} \tag{3.28}$$

da cui in particolare si ottiene che

$$\dim U^{\perp} = \dim V - \dim U \tag{3.29}$$

Dimostrazione. È sufficiente provare che $V = U \oplus U^{\perp}$ perché la formula sulle dimensioni segue dalla proposizione 2.3.11.
Procediamo per induzione su $\dim V$ e osserviamo che non è restrittivo supporre $U \neq \{\mathbf{0}\}$ (se fosse $U = \{\mathbf{0}\}$ si avrebbe $U^{\perp} = V$ per cui $\mathbf{v} = \mathbf{0} + \mathbf{v}$ e dato che $U \cap U^{\perp} = \{\mathbf{0}\}$ possiamo scrivere $V = U \oplus V = U \oplus U^{\perp}$, per cui la proposizione è dimostrata).
Se $\dim V = 1$ (base di un elemento), allora $V = U$ (in quanto un sottospazio deve esser generato da una base ad almeno 1 elemento, ed avendone V 1 U ne ha 1) e non c'è nulla da dimostrare. Si supponga ora che la proposizione sia vera per ogni spazio di dimensione minore di $\dim V$ e si fissi $\mathbf{u} \in U$ con $\mathbf{u} \neq 0$.
Poiché per il lemma 3.3.6 è $V = \langle \mathbf{u} \rangle \oplus \langle \mathbf{u} \rangle^{\perp}$, posto $W = \langle \mathbf{u} \rangle^{\perp}$ e considerato che

$$U \cap W = W = \langle \mathbf{u} \rangle^{\perp}$$

si ha

$$U = \langle \mathbf{u} \rangle \oplus (U \cap W) \tag{3.30}$$

e quindi applicando l'ortogonalità ad entrambi i membri

$$U^{\perp} = (\langle \mathbf{u} \rangle \oplus (U \cap W))^{\perp} \overset{(1)}{=} \langle \mathbf{u} \rangle^{\perp} \cap (U \cap W)^{\perp} = W \cap (U \cap W)^{\perp} \tag{3.31}$$

dove (1) deriva dal punto 3 di proposizione 3.3.5.
Poiché $\dim W < \dim V$ (W è sottospazio di V) e $W \cap (U \cap W)^{\perp}$ è il complemento ortogonale di $U \cap W$ in W (il complemento ortogonale di $(U \cap W)$ in W sono elementi di W ortogonali a $(U \cap W)$, ossia $W \cap (U \cap W)^{\perp}$) dall'ipotesi induttiva segue che

$$W = (U \cap W) \oplus (W \cap (U \cap W)^{\perp}) \overset{(1)}{=} (U \cap W) \oplus U^{\perp}$$

dove (1) deriva da 3.31. Pertanto si ottiene

$$V = \langle \mathbf{u} \rangle \oplus W = \langle \mathbf{u} \rangle \oplus (U \cap W) \oplus U^{\perp} = U \oplus U^{\perp}$$

dove all'ultimo passaggio si è sfruttata l'uguaglianza della 3.30. □

Corollario 3.3.8. *Per ogni sottospazio U di V si ha $U^{\perp\perp} = U$*

Dimostrazione. Poiché si ha banalmente che $U^{\perp}$ è un sottospazio di V, dalla proposizione 3.3.7 segue

$$\dim V = \dim U^{\perp} + \dim (U^{\perp})^{\perp}$$

e quindi

$$\dim U^{\perp\perp} = \dim V - \dim U^{\perp} \overset{(1)}{=} \dim V - (\dim V - \dim U) = \dim U$$

dove in (1) si è applicata proposizione 3.3.7. La conclusione segue da proposizione 2.3.10 (su uguaglianza delle dimensioni e uguaglianza degli spazi). □

Proposizione 3.3.9. *Una volta fissato un sottospazio U di V ogni vettore $\mathbf{v} \in V$ individua completamente due vettori $\mathbf{u} \in U$ e $\mathbf{w} \in U^{\perp}$, tali che $\mathbf{v} = \mathbf{u} + \mathbf{w}$.*

Dimostrazione. Deriva dal fatto che lo spazio vettoriale V è somma diretta di due suoi sottospazi W e Z se e solo se ogni vettore di V si scrive in uno e un solo modo come somma di un vettore di W e uno di Z (per la proposizione 2.3.12, punti 1 e 2). □

Esempio 3.3.6. Sia $V = \mathbb{R}^3$ dotato del prodotto interno standard, $U = \langle \mathbf{e}_1; \mathbf{e}_3 \rangle$ e pertanto (esempio 3.3.2) $U^{\perp} = \langle \mathbf{e}_2 \rangle$. Ogni vettore $\mathbf{v} = [x_0\, y_0 z_0]^T \in \mathbb{R}^3$ si può scomporre nella somma del vettore $\mathbf{w} = [x_0\, 0\, z_0]^T \in U$ (che rappresenta la proiezione di $\mathbf{v}$ sul piano Oxz) con il vettore $\mathbf{z} = [0\, y_0\, 0]^T \in U^{\perp}$ (proiezione sull'asse y).

Osservazione 199. Possiamo generalizzare il concetto geometrico di proiezione ortogonale di un vettore di $\mathbb{R}^3$ su di un piano o su di una retta.

Proposizione 3.3.10 (Linearità applicazione scelta del sottospazio)**.** *Sia U un sottospazio di V. L'applicazione $P_U : V \to V$ che associa a ogni $\mathbf{v} \in V$ l'unico vettore $\mathbf{u} \in U$*

$$P_U(\mathbf{v}) = \mathbf{u} \tag{3.32}$$

e tale che $\mathbf{v} = \mathbf{u} + \mathbf{w}$ con $\mathbf{w} \in U^{\perp}$ è un'applicazione lineare.

Dimostrazione. Da fare osservando che $V = U \oplus U^{\perp}$ per la proposizione 3.3.7. Dovrebbe comunque essere un'applicazione della funzione di scelta delle componenti, che è lineare, o qualcosa di simile. □ **TODO**: fixme

Osservazione 200. P_U è l'unica applicazione lineare di V la cui restrizioni a U è l'identità, mentre a $U^{\perp}$ è la funzione identicamente nulla.

Definizione 3.3.4 (Proiezione ortogonale)**.** Sia U un sottospazio dello spazio vettoriale euclideo V. L'applicazione lineare $P_U : V \to V$ tale che

1. $P_U(\mathbf{u}) = \mathbf{u}$ per ogni $\mathbf{u} \in U$

2. $P_U(\mathbf{u}) = \mathbf{0}$ per ogni $\mathbf{w} \in U^\perp$

si chiama proiezione ortogonale di V su U. L'immagine $P_U(\mathbf{v})$ tramite P_U di un vettore $\mathbf{v} \in V$ si chiama la proiezione ortogonale del vettore $\mathbf{v}$ sul sottospazio U di V.

Proposizione 3.3.11. *Siano U un sottospazio dello spazio vettoriale euclideo V e $P_U : V \to V$ la proiezione ortogonale di V su U. Per ogni $\mathbf{v} \in V$ si ha:*

1. $P_U(\mathbf{v}) = \mathbf{v} \iff \mathbf{v} \in U$
2. $\mathbf{v} - P_U(\mathbf{v}) = \mathbf{w} \in U^\perp$

Dimostrazione. Dalla linearità di P_U si ottiene che se $\mathbf{v} = \mathbf{u} + \mathbf{w}$ con $\mathbf{v} \in V$, $\mathbf{u} \in U$ e $\mathbf{w} \in U^\perp$ allora:

$$P_U(\mathbf{v}) = P_U(\mathbf{u} + \mathbf{w}) = P_U(\mathbf{u}) + P_U(\mathbf{w}) = \mathbf{u} + \mathbf{0} = \mathbf{u}$$

ciò prova il primo punto. Per il secondo, semplicemente considerando che $\mathbf{v} = \mathbf{u} + \mathbf{w} = P_U(\mathbf{v}) + \mathbf{w}$, si ha che

$$\mathbf{v} - P_U(\mathbf{v}) = \mathbf{w} \in U^\perp$$

□

Osservazione 201. Nel caso di $V = \mathbb{R}^3$ dotato del prodotto interno standard e U un sottospazio di dimensione 2 (si veda l'esempio 3.3.6) identificando i vettori con i segmenti orientati uscenti dall'origine che li rappresentano, $P_U(\mathbf{v})$ è quindi proprio la proiezione ortogonale in senso geometrico di $\mathbf{v}$ sul piano che rappresenta U.

Esempio 3.3.7 (Distanza di un punto da un piano)**.** Ricordiamo che in geometria la distanza di un punto P da un piano π è la lunghezza minima di un segmento congiungente P a un punto di π, e coincide con quella del segmento perpendicolare al piano (punto di incontro in H) e passante per P, ossia la lunghezza del segmento PH.
Per il calcolo della distanza tra $\mathbf{v} \in V$ e il piano rappresentato dal sottospazio U si applica la 3.3.11, in particolare calcolando la norma di $\mathbf{v} - P_U(\mathbf{v})$.
Per esempio se $V = \mathbb{R}^3$, $U = \langle \mathbf{e}_1; \mathbf{e}_3 \rangle$ (ossia è il piano xz) e $\mathbf{v} = [5\,4\,3]$ allora possiamo scomporre $\mathbf{v} = \mathbf{u} + \mathbf{w}$ con

$$\mathbf{u} = \begin{bmatrix} 5 \\ 0 \\ 3 \end{bmatrix} \in U, \quad \mathbf{w} = \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix} \in U^\perp = \langle \mathbf{e}_2 \rangle$$

la distanza del punto P corrispondente a $\mathbf{v}$ dal piano corrispondente a U è

$$\|\mathbf{v} - P_U(\mathbf{v})\|_2 = \|\mathbf{w}\|_2 = \left\| \begin{bmatrix} 0 \\ 4 \\ 0 \end{bmatrix} \right\|_2 = 4$$

Osservazione 202. Generalizziamo ora quanto visto in $\mathbb{R}^3$.

Teorema 3.3.12. *Siano U un sottospazio dello spazio vettoriale euclideo V e $P_U : V \to V$ la proiezione ortogonale di V su U; sia inoltre $\|\cdot\|$ la norma indotta dal prodotto interno di V. Per ogni $\mathbf{v} \in V$ si ha:*

1.
$$\begin{cases} \|P_U(\mathbf{v})\| < \|\mathbf{v}\| & \text{se } \mathbf{v} \notin U \\ \|\mathbf{v} - P_U(\mathbf{v})\| < \|\mathbf{v}\| & \text{se } \mathbf{v} \notin U^\perp \end{cases}$$

2. $\|\mathbf{v} - P_U(\mathbf{v})\| < \|\mathbf{v} - \mathbf{z}\|$ *per ogni* $\mathbf{z} \in U$ *con* $\mathbf{z} \neq P_U(\mathbf{v})$

3. $P_U(P_U(\mathbf{v})) = P_U(\mathbf{v})$

4. $(\mathbf{v}|P_U(\mathbf{z})) = (P_U(\mathbf{v})|\mathbf{z})$

Dimostrazione. Rispettivamente:

1. per proposizione 3.3.7 esistono $\mathbf{u} \in U$ e $\mathbf{w} \in U^\perp$ tali che $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Da
$$\begin{aligned} \|\mathbf{v}\|^2 = \|\mathbf{u}+\mathbf{w}\|^2 &= (\mathbf{u}+\mathbf{w}|\mathbf{u}+\mathbf{w}) = (\mathbf{u}+\mathbf{w}|\mathbf{u}) + (\mathbf{u}+\mathbf{w}|\mathbf{w}) \\ &= (\mathbf{u}|\mathbf{u}) + (\mathbf{u}|\mathbf{w}) + (\mathbf{w}|\mathbf{u}) + (\mathbf{w}|\mathbf{w}) \\ &\overset{(1)}{=} (\mathbf{u}|\mathbf{u}) + (\mathbf{w}|\mathbf{w}) \\ &= \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2 \end{aligned}$$
dove (1) deriva dal fatto che $\mathbf{u}$ e $\mathbf{w}$ sono perpendicolari quindi $(\mathbf{w}|\mathbf{u}) = (\mathbf{u}|\mathbf{w}) = 0$. Pertanto, dato che $\|\mathbf{u}+\mathbf{w}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2$ si deduce che:
$$\begin{aligned} \|\mathbf{u}+\mathbf{w}\|^2 &> \|\mathbf{u}\|^2 \quad \text{se } \mathbf{w} \neq \mathbf{0} \\ \|\mathbf{u}+\mathbf{w}\|^2 &> \|\mathbf{w}\|^2 \quad \text{se } \mathbf{u} \neq \mathbf{0} \end{aligned}$$
ossia che (estraendo la radice di due positivi):
$$\begin{aligned} \|\mathbf{u}+\mathbf{w}\| &> \|\mathbf{u}\| \quad \text{se } \mathbf{v} \notin U^\perp \\ \|\mathbf{u}+\mathbf{w}\| &> \|\mathbf{w}\| \quad \text{se } \mathbf{v} \notin U^\perp \end{aligned}$$
A parte la radice mostriamo però a titolo di esempio per la prima che $\mathbf{w} \neq \mathbf{0} \iff \mathbf{v} \notin U$; lo facciamo mostrando l'equivalenza contraria corrispondente, ossia $\mathbf{w} = \mathbf{0} \iff \mathbf{v} \in U$. Se $\mathbf{w} = \mathbf{0}$ allora $\mathbf{v} = \mathbf{u} \in U$; se viceversa $\mathbf{v} \in U$ allora tenendo conto che $\mathbf{v} = \mathbf{u} + \mathbf{w}$ deve essere $\mathbf{v} = \mathbf{u}$ ossia $\mathbf{w} = \mathbf{0}$.
Per concludere dimostrando il punto 1 si effettuano opportune sostituzioni alla luce del fatto che $P_U(\mathbf{v}) = \mathbf{u}$ e $\mathbf{v} = \mathbf{u} + \mathbf{w}$.

2. per il punto 2, se $\mathbf{z} \in U$ e $\mathbf{z} \neq P_U(\mathbf{v})$ si ha che il vettore che unisce un generico $\mathbf{v} \in V$ a $\mathbf{z}$ non è perpendicolare a U ossia $\mathbf{v}_1 = \mathbf{v} - \mathbf{z} \notin U^\perp$, e un'applicazione del punto 1 segue: se $\mathbf{v}_1 = \mathbf{v} - \mathbf{z} \notin U^\perp$ si ha che
$$\|\mathbf{v}_1 - P_U(\mathbf{v}_1)\| < \|\mathbf{v}_1\| \tag{3.33}$$
Ma dalla $\mathbf{v}_1 = \mathbf{v} - \mathbf{z}$ abbiamo, applicando la proiezione ad entrambi i membri che
$$P_U(\mathbf{v}_1) = P_U(\mathbf{v} - \mathbf{z}) = P_U(\mathbf{v}) - P_U(\mathbf{z}) = P_U(\mathbf{v}) - \mathbf{z}$$
dove l'ultima uguaglianza è giustificata dal fatto che $\mathbf{z} \in U$. Dunque sostituendo vari pezzi nella 3.33:
$$\begin{aligned} \|\mathbf{v}_1 - P_U(\mathbf{z})\| &< \|\mathbf{v}_1\| \\ \|\mathbf{v} - \mathbf{z} - P_U(\mathbf{v}) + \mathbf{z}\| &< \|\mathbf{v} - \mathbf{z}\| \\ \|\mathbf{v} - P_U(\mathbf{v})\| &< \|\mathbf{v} - \mathbf{z}\| \end{aligned}$$

3. il punto 3 segue da $P_U(\mathbf{v}) \in U$ e dal fatto che definizione di $P_U(\mathbf{x}) = \mathbf{x}$ se $\mathbf{x} \in U$;

4. il punto 4: se

   - $\mathbf{v}, \mathbf{z} \in V$
   - $\mathbf{v} = \mathbf{u} + \mathbf{w}$ e $\mathbf{z} = \mathbf{u}' + \mathbf{w}'$,
   - $\mathbf{u}, \mathbf{u}' \in U$ e $\mathbf{w}, \mathbf{w}' \in U^\perp$

   ci basta verificare la seguente uguaglianza:

$$(\mathbf{u} + \mathbf{w}|P_U(\mathbf{u}' + \mathbf{w}')) = (P_U(\mathbf{u} + \mathbf{w})|\mathbf{u}' + \mathbf{w}')$$
$$(\mathbf{u} + \mathbf{w}|P_U(\mathbf{u}') + P_U(\mathbf{w}')) = (P_U(\mathbf{u}) + P_U(\mathbf{w})|\mathbf{u}' + \mathbf{w}')$$

   Il primo membro si semplifica come:

$$(\mathbf{u}+\mathbf{w}|P_U(\mathbf{u}')+P_U(\mathbf{w}')) = (\mathbf{u}|P_U(\mathbf{u}'))+\underbrace{(\mathbf{u}|P_U(\mathbf{w}'))}_{=0}+\underbrace{(\mathbf{w}|P_U(\mathbf{u}'))}_{=0}+\underbrace{(\mathbf{w}|P_U(\mathbf{w}'))}_{=0}$$

   Il secondo come:

$$(P_U(\mathbf{u})+P_U(\mathbf{w})|\mathbf{u}'+\mathbf{w}') = (P_U(\mathbf{u})|\mathbf{u}')+\underbrace{(P_U(\mathbf{u})|\mathbf{w})}_{=0}+\underbrace{(P_U(\mathbf{w})|\mathbf{u}')}_{=0}+\underbrace{(P_U(\mathbf{w})|\mathbf{w})}_{=0}$$

   Proseguendo nell'uguaglianza dalla quale si è partiti:

$$(\mathbf{u}|P_U(\mathbf{u}')) = (P_U(\mathbf{u})|\mathbf{u}')$$
$$(\mathbf{u}|\mathbf{u}') = (\mathbf{u}|\mathbf{u}')$$

   e quest'ultima, essendo una identità, è sempre verificata.

$\square$

Osservazione 203. Dal punto 2 del teorema 3.3.12 si deduce che se per $\mathbf{z}_1 \in U$ si ha $\|\mathbf{v} - \mathbf{z}_1\| \leq \|\mathbf{v} - \mathbf{z}\|$ per ogni $\mathbf{z} \in U$ allora $\mathbf{z}_1 = P_U(\mathbf{v})$.
Per questo motivo si dice che $P_U(\mathbf{v})$ è il vettore di U che *meglio approssima* $\mathbf{v}$ rispetto alla norma indotta dal prodotto interno di V; se tale norma coincide con la norma euclidea (per cui $\|\mathbf{v}\|_2 = \sqrt{\left(\sum |v_i|^2\right)}$ si dice che $P_U(\mathbf{v})$ meglio approssima $\mathbf{v}$ ai minimi quadrati.

Osservazione 204. In analogia con il caso $\mathbb{R}^3$ si pone la seguente definizione.

Definizione 3.3.5 (Distanza del vettore $\mathbf{v}$ dal sottospazio U di V)**.** Nelle notazioni del teorema 3.3.12 viene chiamato così il numero $\|\mathbf{v} - P_U(\mathbf{v})\|$ e si indica con $d(\mathbf{v}, U)$.

Esempio 3.3.8. Si consideri $V = \mathbb{R}^2$ con la norma $\|\cdot\|_1$. Sia $d : V \times V \to \mathbb{R}_{\geq 0}$ la distanza tra due vettori definita da $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_1$. Sia U il sottospazio di V generato dal vettore $\mathbf{w} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Trovare $d(\mathbf{u}, \mathbf{v})$ per $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ al variare di

$u \in U$. Dedurre che non esiste la proiezione ortogonale di v su U.
Se $U = \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle$ un generico $u \in U$ è del tipo $\begin{bmatrix} \alpha \\ \alpha \end{bmatrix}$ quindi

$$d(\mathbf{u}, \mathbf{v}) = \left\| \begin{bmatrix} \alpha \\ \alpha \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\|_1 = \left\| \begin{bmatrix} \alpha - 1 \\ \alpha \end{bmatrix} \right\|_1 = |\alpha - 1| + |\alpha|$$

Vediamo quali valori può assumere la norma in relazione al valore di α; si ha che

$$|\alpha - 1| = \begin{cases} \alpha - 1 & \text{se } \alpha \geq 1 \\ 1 - \alpha & \text{se } \alpha < 1 \end{cases}, \quad |\alpha| = \begin{cases} \alpha & \text{se } \alpha \geq 0 \\ -\alpha & \text{se } \alpha < 0 \end{cases}$$

Dunque

$$d(\mathbf{u}, \mathbf{v}) = |\alpha - 1| + |\alpha| = \begin{cases} 1 - \alpha - \alpha = 1 - 2\alpha > 1 & \text{se } \alpha < 0 \\ \alpha - 1 + \alpha = 2\alpha - 1 > 1 & \text{se } \alpha \geq 1 \\ 1 - \alpha + \alpha = 1 & \text{se } 0 \leq \alpha < 1 \end{cases}$$

La distanza è minimizzata se $0 \leq \alpha < 1$, ma vi sono infiniti vettori che la rispettano quindi non ce n'è uno unico che la minimizza e la proiezione ortogonale non esiste.

3.4 Basi ortogonali e basi ortonormali

3.4.1 Basi ortogonali

Osservazione 205. Sia V uno spazio vettoriale euclideo. Generalizziamo il concetto di ortogonalità tra due vettori, introdotto nel paragrafo precedente, a un insieme finito di vettori.

Definizione 3.4.1. Un insieme di vettori $\{\mathbf{v}_1, \dots \mathbf{v}_k\}$ di $k \geq 2$ elementi di V si dice ortogonale se $\mathbf{v}_i \perp \mathbf{v}_j$ per ogni $i, j = 1, \dots, k$ con $i \neq j$ (ossia se è costituito da vettori a due a due ortogonali).

Osservazione 206. Si conviene che ogni insieme costituito da un unico vettore è un insieme ortogonale.

Definizione 3.4.2 (Base ortogonale)**.** Una base ortogonale di V è una base di V che sia anche un insieme ortogonale.

Esempio 3.4.1. La base canonica $\mathscr{E} = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di $\mathbb{C}^n$ è una base ortogonale di $\mathbb{C}^n$ dotato del prodotto interno standard. Infatti se $n > 1$:

$$(\mathbf{e}_i | \mathbf{e}_j) = \mathbf{e}_i^H \mathbf{e}_j = 0, \quad \forall i \neq j$$

Osservazione 207. L'ortogonalità è una condizione più forte dell'indipendenza lineare.

Proposizione 3.4.1. *Un insieme ortogonale $\mathscr{S} = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ costituito da vettori non nulli è linearmente indipendente.*

Dimostrazione. Si consideri una combinazione lineare nulla degli elementi di $\mathscr{S}$:

$$\mathbf{0} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k$$

Per ogni $j \in \{1, \ldots, k\}$ dalla linearità del prodotto interno e da $\mathbf{v}_j \perp \mathbf{v}_i$ per ogni $i \neq j$ si ottiene

$$\begin{aligned} 0 = (\mathbf{v}_j|\mathbf{0}) &= (\mathbf{v}_j|\alpha_1 \mathbf{v}_1 + \ldots + \alpha_k \mathbf{v}_k) \\ &= \alpha_1(\mathbf{v}_j|\mathbf{v}_1) + \ldots + \alpha_k(\mathbf{v}_j|\mathbf{v}_k) \\ &= \alpha_j(\mathbf{v}_j|\mathbf{v}_j) \end{aligned}$$

Da ciò si deduce, essendo $(\mathbf{v}_j|\mathbf{v}_j) \neq 0$ ($\mathbf{v}_j$ è supposto non nullo) che $\alpha_j = 0$, questo per tutti i $j \in \{1, \ldots, k\}$, ossia che $\mathscr{S}$ è linearmente indipendente. □

Osservazione 208. La seguente proposizione prova l'esistenza di basi ortogonali di V. Nel prossimo paragrafo illustreremo un algoritmo, detto algoritmo di Gram-Schmidt, che permette di costruire una base ortogonale di V a partire da una base generica. Per tutto il capitolo, anche se non sarà detto esplicitamente, ogni spazio vettoriale considerato sarà euclideo (con norma interna).

Proposizione 3.4.2. *Se $V \neq \{\mathbf{0}\}$, allora V ammette una base ortogonale.*

Dimostrazione. Per induzione su $n = \dim V \geq 1$:

- se $n = 1$ ogni base di V è costituita da un unico vettore, e quindi per definizione/convenzione è una base ortogonale di V;
- supponiamo ora $n > 1$ e che ogni spazio vettoriale euclideo U di dimensione $n - 1$ ammetta una base ortogonale (ipotesi induttiva).
  Siano $\mathbf{0} \neq \mathbf{v} \in V$ e $U = \langle \mathbf{v} \rangle^\perp$ il completamento ortogonale di $\langle \mathbf{v} \rangle$ in V. Dalla proposizione 3.3.6 si ottiene che $V = \langle \mathbf{v} \rangle \oplus U$ e

  $$\dim U = \dim V - 1 = n - 1$$

  Per ipotesi induttiva esiste una base ortogonale $\{\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}\}$ di U; i vettori che la compongono (se in numero maggiore di uno) sono a due a due ortogonali e inoltre ciascuno di essi (stando nel complemento ortogonale di $\langle \mathbf{v} \rangle$) è ortogonale a $\mathbf{v}$.
  L'insieme $\mathscr{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_{n-1}, \mathbf{v}\}$ è dunque un insieme ortogonale costituito da vettori non nulli, che quindi, per la proposizione 3.4.1 è linearmente indipendente. Poiché $\mathscr{B}$ ha $n = \dim V$ elementi, esso è una base di V, e concludendo è una base ortogonale di V.

□

Osservazione 209. Il vantaggio dell'impiego di basi ortogonali (rispetto a basi generiche) sta nella possibilità di un calcolo immediato dei coefficienti delle combinazioni lineari dei loro vettori. Più precisamente vale il risultato seguente

Proposizione 3.4.3. *Siano U un sottospazio di V e $\mathscr{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ una base ortogonale di U. Allora per ogni vettore $\mathbf{u}$ di U si ha*

$$\mathbf{u} = \alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k$$

con

$$\alpha_i = \frac{(\mathbf{u}_i|\mathbf{u})}{(\mathbf{u}_i|\mathbf{u}_i)}, \quad i = 1, \ldots, k$$

Osservazione 210. Ossia dato $\mathbf{u}$ e conosciuta $\mathscr{B}$, possiamo determinare facilmente la combinazione (coefficienti) che genera $\mathbf{u}$.

Dimostrazione. Poiché $\mathscr{B}$ è una base di U e $\mathbf{u} \in U$, esistono scalari $\alpha_1, \dots, \alpha_k$ tali che

$$\mathbf{u} = \alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k$$

Con un procedimento analogo a quello per la dimostrazione di proposizione 3.4.1, per ogni $i \in \{1, \dots, k\}$ dalla linearità del prodotto interno e da $\mathbf{u}_j \perp \mathbf{u}_i$ per ogni $j \neq i$ si ottiene

$$\begin{aligned}(\mathbf{u}_i|\mathbf{u}) &= (\mathbf{u}_i|\alpha_1 \mathbf{u}_1 + \ldots + \alpha_k \mathbf{u}_k) \\ &= \alpha_1(\mathbf{u}_i|\mathbf{u}_1) + \ldots + \alpha_k(\mathbf{u}_i|\mathbf{u}_k) \\ &= \alpha_i(\mathbf{u}_i|\mathbf{u}_i)\end{aligned}$$

Se ne deduce, essendo $(\mathbf{u}_i|\mathbf{u}_i) \neq 0$ che

$$\alpha_i = \frac{(\mathbf{u}_i|\mathbf{u})}{(\mathbf{u}_i|\mathbf{u}_i)}$$

□

Esempio 3.4.2. Sia $\mathbb{R}^3$ dotato del prodotto interno standard e consideriamo

$$\mathscr{B} = \left\{ \mathbf{v}_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}1\\0\\-1\end{bmatrix} \right\}$$
$$\mathscr{B}' = \left\{ \mathbf{v}'_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}, \mathbf{v}'_2 = \begin{bmatrix}1\\2\\0\end{bmatrix}, \mathbf{v}'_3 = \begin{bmatrix}1\\1\\-1\end{bmatrix} \right\}$$

$\mathscr{B}$ e $\mathscr{B}'$ sono due basi di V (se si vuole, verificarlo); $\mathscr{B}$ è ortogonale, a differenza di $\mathscr{B}'$. Infatti ad esempio $(\mathbf{v}'_1|\mathbf{v}'_2) = \mathbf{v}'^T_1 \mathbf{v}'_2 = 1 \neq 0$.
Vogliamo esprimere il vettore $\mathbf{v} = [3\,4\,5]^T$ come combinazione lineare degli elementi di $\mathscr{B}$ e di $\mathscr{B}'$ rispettivamente. Cerchiamo quindi gli scalari α_i, β_i tali che:

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 = \beta_1 \mathbf{v}'_1 + \beta_2 \mathbf{v}'_2 + \beta_3 \mathbf{v}'_3$$

Per trovare i β_i (caso di base non ortogonale) non abbiamo altra scelta che risolvere un sistema lineare, che può richiedere anche calcoli laboriosi. In questo caso il sistema lineare ha come matrice aumentata

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 0 & 2 & 1 & 4 \\ 1 & 0 & -1 & 5 \end{array}\right]$$

e come unica soluzione $[7/3 \; 10/3 \; -8/3]$.
Per trovare gli α_i invece possiamo procedere direttamente calcolando

$$\alpha_1 = \frac{\mathbf{v}_1^T \mathbf{v}}{\mathbf{v}_1^T \mathbf{v}_1} = 4, \quad \alpha_2 = \frac{\mathbf{v}_2^T \mathbf{v}}{\mathbf{v}_2^T \mathbf{v}_2} = 4, \quad \alpha_3 = \frac{\mathbf{v}_3^T \mathbf{v}}{\mathbf{v}_3^T \mathbf{v}_3} = -1$$

Esempio 3.4.3. Sia

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} \right\}$$

Provare che $\mathscr{B}$ è una base ortogonale e ricavarne una base ortonormale; inoltre dato $\mathbf{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$, trovare $\mathrm{C}_{\mathscr{B}}(\mathbf{v})$.

Proviamo che $\mathscr{B}$ sia base (di $\mathbb{R}^3$) facendo la riduzione e notando che la matrice è triangolare superiore

$$\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 5 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 6 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 6 \\ 0 & -3 & -4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 6 & 18 \\ 0 & -6 & -12 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 6 & 18 \\ 0 & 0 & 6 \end{bmatrix}$$

Ora vediamo che sia ortogonale, ossia che

$$\mathbf{v}_i \perp \mathbf{v}_j, \forall i \neq j \iff (\mathbf{v}_i|\mathbf{v}_j) = \mathbf{v}_i^T \mathbf{v}_j = 0$$

Si ha

$$\mathbf{v}_1^T \mathbf{v}_2 = \begin{bmatrix} 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} = 2 \cdot 1 - 0 - 2 = 0$$

$$\mathbf{v}_1^T \mathbf{v}_3 = \begin{bmatrix} 1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} = 1 \cdot 1 - 5 + 2 \cdot 2 = 0$$

$$\mathbf{v}_2^T \mathbf{v}_3 = \begin{bmatrix} 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} = 2 \cdot 1 - 2 \cdot 1 = 0$$

Quindi sì, è ortogonale. Per ricavare una base ortonormale si ha (come si vedrà in seguito)

$$\mathscr{B}' = \left\{ \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} \right\}$$

con

$$\|\mathbf{v}_1\| = \sqrt{(1)^2 + (-1)^2 + (2)^2 4} = \sqrt{1+1+4} = \sqrt{6}$$
$$\|\mathbf{v}_2\| = \sqrt{(2)^2 + 0^2 + (-1)^2} = \sqrt{4+1} = \sqrt{5}$$
$$\|\mathbf{v}_3\| = \sqrt{1^2 + 5^2 + 2^2} = \sqrt{1+25+40} = \sqrt{30}$$

quindi

$$\mathscr{B}' = \left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}, \frac{1}{\sqrt{30}} \begin{bmatrix} 1 \\ 5 \\ 2 \end{bmatrix} \right\}$$

Infine per trovare $\mathrm{C}_{\mathscr{B}}\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix}\right)$ si avrà che

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \frac{\mathbf{v}_1^T \mathbf{v}}{\mathbf{v}_1^T \mathbf{v}_1} \mathbf{v}_1 + \frac{\mathbf{v}_2^T \mathbf{v}}{\mathbf{v}_2^T \mathbf{v}_2} \mathbf{v}_2 + \frac{\mathbf{v}_3^T \mathbf{v}}{\mathbf{v}_3^T \mathbf{v}_3} \mathbf{v}_3$$

per cui la risposta è

$$\begin{bmatrix} \frac{\mathbf{v}_1^T\mathbf{v}}{\mathbf{v}_1^T\mathbf{v}_1} \\ \frac{\mathbf{v}_2^T\mathbf{v}}{\mathbf{v}_2^T\mathbf{v}_2} \\ \frac{\mathbf{v}_3^T\mathbf{v}}{\mathbf{v}_3^T\mathbf{v}_3} \end{bmatrix} = \begin{bmatrix} \frac{a-b+2c}{6} \\ \frac{2a+c}{5} \\ \frac{a+5b-2c}{30} \end{bmatrix}$$

Osservazione 211. Che significato ha la scrittura $\alpha_1\mathbf{u}_1 + \ldots + \alpha_n\mathbf{u}_n$ della proposizione 3.4.3 (con $\alpha_i = \frac{(\mathbf{u}_i|\mathbf{u})}{(\mathbf{u}_i|\mathbf{u}_i)}$) se $\mathbf{u}$ non appartiene al sottospazio U?

Proposizione 3.4.4. *Siano U un sottospazio di V e $\mathscr{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ una base ortogonale di U. Per ogni $\mathbf{v} \in V$ si ha che la sua proiezione ortogonale su U si ottiene facilmente come:*

$$P_U(\mathbf{v}) = \frac{(\mathbf{u}_1|\mathbf{v})}{(\mathbf{u}_1|\mathbf{u}_1)}\mathbf{u}_1 + \ldots + \frac{(\mathbf{u}_k|\mathbf{v})}{(\mathbf{u}_k|\mathbf{u}_k)}\mathbf{u}_k$$

Dimostrazione. Applicando la proposizione 3.4.3 a $P_U(\mathbf{v})$ (dato che $P_U(\mathbf{v}) \in U$):

$$P_U(\mathbf{v}) = \frac{(\mathbf{u}_1|P_U(\mathbf{v}))}{(\mathbf{u}_1|\mathbf{u}_1)}\mathbf{u}_1 + \ldots + \frac{(\mathbf{u}_k|P_U(\mathbf{v}))}{(\mathbf{u}_k|\mathbf{u}_k)}\mathbf{u}_k \tag{3.34}$$

Poiché abbiamo visto che $\mathbf{v} = P_U(\mathbf{v}) + \mathbf{w}$ per un opportuno $\mathbf{w} \in U^\perp$ possiamo riscrivere:

$$(\mathbf{u}_i|\mathbf{v}) = (\mathbf{u}_i|P_U(\mathbf{v}) + \mathbf{w}) = (\mathbf{u}_i|P_U(\mathbf{v})) + (\mathbf{u}_i|\mathbf{w}) = (\mathbf{u}_i|P_U(\mathbf{v}))$$

per ogni $i = 1, \ldots, k$. Quindi, dato che $(\mathbf{u}_i|P_U(\mathbf{v})) = (\mathbf{u}_i|\mathbf{v})$, andando a sostituire quest'ultima in 3.34 si ottiene la tesi. □

Osservazione 212. Siano U un sottospazio non nullo di V ed $\mathscr{S} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ un insieme ortogonale di generatori di U. Allora l'insieme che si ottiene da $\mathscr{S}$ togliendo gli eventuali vettori nulli è una base ortogonale di U e per ogni $\mathbf{v} \in V$ si ha (cfr Esercizio 3.17)

TODO: fixme

$$P_U(\mathbf{v}) = \alpha_1\mathbf{u}_1 + \ldots + \alpha_k\mathbf{u}_k$$

dove

$$\alpha_i = \begin{cases} 0 & \text{se } \mathbf{u}_i = \mathbf{0} \text{ (ossia se } \mathbf{u}_i \text{ non fa parte della base ortogonale)} \\ \frac{(\mathbf{u}_i|\mathbf{v})}{(\mathbf{u}_i|\mathbf{u}_i)} & \text{se } \mathbf{u}_i \neq \mathbf{0} \text{ (ossia per gli altri vettori della base, non nulli)} \end{cases}$$

per ogni $i = 1, \ldots, k$

Osservazione 213. Per il calcolo della proiezione ortogonale di un vettore $\mathbf{v} \in V$ su un sottospazio U di V occorre quindi saper costruire una base ortogonale di U, la cui esistenza è assicurata dalla proposizione 3.4.2.

3.4.2 Basi ortonormali

Osservazione 214. Se i calcoli possono esser semplificati dall'impiego di basi ortogonali, lo sono ancor di più quando si usano *basi ortonormali*.

Definizione 3.4.3 (Vettore normalizzato)**.** Diremo che un vettore $\mathbf{v} \in V$ viene normalizzato quando al posto di $\mathbf{v}$ si considera il vettore di norma unitaria ottenuto come:

$$\frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$$

dove $\|\cdot\|$ è la norma indotta dal prodotto interno.

Osservazione 215. Si ha che:

$$\left\langle \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\rangle = \langle \mathbf{v} \rangle$$

Definizione 3.4.4 (Base ortonormale)**.** Una base $\mathscr{B} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di V si dice una base ortonormale se è ortogonale e ogni suo vettore è normalizzato, ossia se

$$\mathbf{v}_i \perp \mathbf{v}_j \quad \text{per } i \neq j, \quad \|\mathbf{v}_i\| = 1 \quad \text{per } i = 1, \dots, n$$

Esempio 3.4.4. La base canonica $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ di $\mathbb{C}^n$ è una base ortonormale di $\mathbb{C}^n$ dotato del prodotto interno standard, essendo ortogonale (esempio 3.4.1) ed essendo $\|\mathbf{e}_i\|_2 = 1$ per ogni $i = 1, \dots, n$

Osservazione 216. Ovviamente, per ottenere una base ortonormale, normalizzarne una ortogonale.

Osservazione 217. Siano U un sottospazio di V e $\mathscr{B} = \{\mathbf{u}_1, \dots, \mathbf{u}_k\}$ una base ortonormale di U. Dal momento che per ogni $i = 1, \dots, k$ si ha che $(\mathbf{u}_i|\mathbf{u}_i) = \|\mathbf{u}_i\|^2 = 1$ la proposizione 3.4.3 si semplifica nel seguente modo:

$$\mathbf{u} = (\mathbf{u}_1|\mathbf{u})\mathbf{u}_1 + \ldots + (\mathbf{u}_k|\mathbf{u})\mathbf{u}_k$$

Analogamente, se $\mathbf{v} \in V$ dalla proposizione 3.4.4 si ottiene:

$$P_U(\mathbf{v}) = (\mathbf{u}_1|\mathbf{v})\mathbf{u}_1 + \ldots + (\mathbf{u}_k|\mathbf{v})\mathbf{u}_k$$

3.5 L'algoritmo di Gram-Schmidt

Osservazione 218. Sia V uno spazio vettoriale euclideo. L'algoritmo applicato ad un insieme di generatori di V, fornisce un insieme ortogonale di generatori. E se $V \neq \{\mathbf{0}\}$ l'insieme che si ottiene togliendo da quest'ultimo gli eventuali vettori nulli è una base ortogonale di V (osservazione 212).

Lemma 3.5.1. *Siano U un sottospazio di V e $\mathscr{S}_U = \{\mathbf{u}_1, \dots \mathbf{u}_r\}$ un insieme ortogonale di generatori di U. Siano poi $\mathbf{v} \in V$ e $W = U + \langle \mathbf{v} \rangle$. Allora posto*

$$\mathbf{u}_{r+1} = \mathbf{v} - P_U(\mathbf{v}) = \mathbf{v} - \alpha_1\mathbf{u}_1 - \ldots - \alpha_r\mathbf{u}_r$$

(da cui $\mathbf{u}_{r+1}$ è perpendicolare a U) dove

$$\alpha_i = \begin{cases} 0 & \text{se } \mathbf{u}_i = \mathbf{0} \\ \frac{(\mathbf{u}_i|\mathbf{v})}{(\mathbf{u}_i|\mathbf{u}_i)} & \text{se } \mathbf{u}_i \neq \mathbf{0} \end{cases}$$

per ogni $i = 1, \dots, r$ si ha che $\mathscr{S}_W = \{\mathbf{u}_1, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}\}$ è un insieme ortogonale di generatori di W

Dimostrazione. Se $U = \{\mathbf{0}\}$ l'asserto è ovvio: in tal caso $\mathscr{S}_U = \{\}$ e $\mathscr{S}_W = \{\mathbf{u}_{r+1}\}$ è un insieme di 1 elemento (quindi ortogonale).
Se invece $U \neq \{\mathbf{0}\}$, per l'osservazione 212 si ha che

$$P_U(\mathbf{v}) = \alpha_1\mathbf{u}_1 + \ldots + \alpha_r\mathbf{u}_r$$

per cui se poniamo

$$\mathbf{u}_{r+1} = \mathbf{v} - P_U(\mathbf{v}) = \mathbf{v} - \alpha_1\mathbf{u}_1 - \ldots - \alpha_r\mathbf{u}_r$$

si ha che $\mathbf{u}_{r+1} \in U + \langle\mathbf{v}\rangle$ e che $\langle\mathscr{S}_W\rangle = \langle\mathbf{u}_1, \ldots, \mathbf{u}_{r+1}\rangle$ è un sottospazio di $W = U + \langle\mathbf{v}\rangle$ (ossia $\langle\mathscr{S}_W\rangle \subseteq W$).
Viceversa, poiché $\mathscr{S}_U$ è un insieme ortogonale e sappiamo che $\mathbf{v} - P_U(\mathbf{v}) \in U^\perp$, allora anche $\mathscr{S}_W$ è un insieme ortogonale (dato che $\mathbf{u}_1, \ldots, \mathbf{u}_r \in U$ ed $\mathbf{u}_{r+1}$ è ortogonale a tutti gli elementi di U). Infine da:

$$\mathbf{v} = P_U(\mathbf{v}) + \mathbf{u}_{r+1} = \alpha_1\mathbf{u}_1 + \ldots + \alpha_r\mathbf{u}_r + \mathbf{u}_{r+1}$$

otteniamo che $W = U + \langle\mathbf{v}\rangle \subseteq \langle\mathscr{S}_W\rangle = \langle\mathbf{u}_1, \ldots, \mathbf{u}_{r+1}\rangle$.
Complessivamente dunque si conclude che $W = \langle\mathscr{S}_W\rangle$ ossia $\mathscr{S}_W$ è un insieme di generatori di W. $\square$

Teorema 3.5.2 (Algoritmo di Gram-Schmidt). *Sia $\mathscr{S} = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ un insieme di generatori di V. Si ponga*

$$V_0 = \{\mathbf{0}\}, \quad V_k = \langle\mathbf{v}_1, \ldots, \mathbf{v}_k\rangle, \quad \textit{per ogni } 1 \leq k \leq m$$

e sia $P_{V_k} : V \to V$ la proiezione ortogonale di V su V_k per ogni $0 \leq k \leq m$. Allora l'insieme $\mathscr{S}^* = \{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ dei vettori

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1 - P_{V_0}(\mathbf{v}_1) = \mathbf{v}_1 \\
\mathbf{u}_2 &= \mathbf{v}_2 - P_{V_1}(\mathbf{v}_2) = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 \\
\mathbf{u}_3 &= \mathbf{v}_3 - P_{V_2}(\mathbf{v}_3) = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 \\
&\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - P_{V_{k-1}}(\mathbf{v}_k) = \mathbf{v}_k - \alpha_{1k}\mathbf{u}_1 - \alpha_{2k}\mathbf{u}_2 - \ldots - \alpha_{k-1,k}\mathbf{u}_{k-1} \\
&\vdots \\
\mathbf{u}_m &= \mathbf{v}_m - P_{V_{m-1}}(\mathbf{v}_m) = \mathbf{v}_m - \alpha_{1m}\mathbf{u}_1 - \alpha_{2m}\mathbf{u}_2 - \ldots - \alpha_{m-1,m}\mathbf{u}_{m-1}
\end{aligned}$$

dove per ogni $i, j = 1, \ldots, m$, con $i < j$, gli α_{ij} sono definiti da

$$\alpha_{ij} = \begin{cases} 0 & \textit{se } \mathbf{u}_i = \mathbf{0} \\ \frac{(\mathbf{u}_i|\mathbf{v}_j)}{(\mathbf{u}_i|\mathbf{u}_i)} & \textit{se } \mathbf{u}_i \neq \mathbf{0} \end{cases}$$

è un insieme ortogonale di generatori di V.

Dimostrazione. Poiché $\mathbf{u}_1 = \mathbf{v}_1$ allora $\mathscr{S}_{V_1} = \{\mathbf{u}_1\}$ è un insieme ortogonale (un unico elemento) di generatori di V_1 (dato che $V_1 = \langle\mathbf{v}_1\rangle = \langle\mathbf{u}_1\rangle$).
La dimostrazione del teorema consiste ora in $m - 1$ passaggi, ciascuno dei quali è un'applicazione del lemma 3.5.1.

Al primo passaggio lo si applica a $U = V_1$, $\mathscr{S}_{V_1} = \{\mathbf{u}_1\}$ e $\mathbf{v} = \mathbf{v}_2$ ottenendo così che $\mathscr{S}_{V_2} = \{\mathbf{u}_1, \mathbf{u}_2\}$ è un insieme ortogonale di generatori di V_2.
Al secondo passaggio lo si applica a $U = V_2$, $\mathscr{S}_{V_2} = \{\mathbf{u}_1, \mathbf{u}_2\}$ (che per il passaggio precedente è un insieme ortogonale di generatori di V_2) e $\mathbf{v} = \mathbf{v}_3$ ottenendo così che $\mathscr{S}_{V_3} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ è un insieme ortogonale di generatori di V_3.
Procedendo in questo modo all'$(m-1)$-esimo passaggio, sapendo dal passaggio precedente che $\mathscr{S}_{V_{m-1}} = \{\mathbf{u}_1, \dots, \mathbf{u}_{m-1}\}$ è un insieme ortogonale di generatori di V_{m-1}, si applica il lemma 3.5.1 a $U = V_{m-1}$, $\mathscr{S}_{V_{m-1}} = \{\mathbf{u}_1, \dots, \mathbf{u}_{m-1}\}$ e $\mathbf{v} = \mathbf{v}_m$ ottenendo così che $\mathscr{S}_{V_m} = \{\mathbf{u}_1, \dots, \mathbf{u}_m\}$ è un insieme ortogonale di generatori di $V_m = V$. □

Osservazione 219. L'insieme $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ che si ottiene applicando l'algoritmo di Gram-Schmidt a un insieme linearmente indipendente $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ (ad esempio una base) non contiene vettori nulli.
Infatti se, con le notazioni del lemma 3.5.1 esistesse $1 < i \leq n$ tale che $\mathbf{u}_i = \mathbf{0}$, da $\mathbf{u}_i = \mathbf{v}_i - P_{V_{i-1}}(\mathbf{v}_i)$ si otterrebbe $\mathbf{v}_i \in V_{i-1}$. Infatti

$$\mathbf{u}_i = \mathbf{0} \iff \mathbf{v}_i = P_{V_{i-1}}(\mathbf{v}_i) \iff \mathbf{v}_i \in V_{i-1}$$

da cui $\mathbf{v}_i$ può essere riscritto come combinazione lineare di $\{\mathbf{v}_1, \dots, \mathbf{v}_{i-1}\}$.
In tal caso, come anche nel caso in cui a essere nullo fosse il vettore $\mathbf{u}_1 = \mathbf{v}_1$ si avrebbe che $\{\mathbf{v}_1, \dots, \mathbf{v}_i\}$ sarebbe linearmente dipendente, mentre in realtà è un sottoinsieme dell'insieme linearmente indipendente $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$.

Esempio 3.5.1. Sia V il sottospazio di $\mathbb{C}^3$, dotato del prodotto interno standard, generato dall'insieme di vettori

$$\mathscr{S} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} i \\ -2 \\ i \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} 8i \\ 8 \\ 8i \end{bmatrix} \right\}$$

Applicheremo l'algoritmo di Gram-Schmidt ai vettori di $\mathscr{S}$. Otterremo i vettori $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ ciascuno dei quali è legato a quelli che lo precedono e al corrispondente elemento di $\mathscr{S}$ dalle seguenti relazioni

$$\begin{aligned}
&\mathbf{u}_1 = \mathbf{v}_1, \\
&\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1, \\
&\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2, \\
&\mathbf{u}_4 = \mathbf{v}_4 - \alpha_{14}\mathbf{u}_1 - \alpha_{24}\mathbf{u}_2 - \alpha_{34}\mathbf{u}_3
\end{aligned}$$

dove per $1 \leq i < j \leq 4$ i coefficienti α_{ij}sono dati dalla formula

$$\alpha_{ij} = \begin{cases} 0 & \text{se } \mathbf{u}_i = \mathbf{0} \\ \frac{(\mathbf{u}_i|\mathbf{v}_j)}{(\mathbf{u}_i|\mathbf{u}_i)} & \text{se } \mathbf{u}_i \neq \mathbf{0} \end{cases}$$

Si ha:

1. poniamo innanzitutto

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}$$

2. per costruire $\mathbf{u}_2$ dobbiamo calcolare il coefficiente α_{12} che essendo $\mathbf{u}_1 \neq \mathbf{0}$ è dato da

$$\alpha_{12} = \frac{(\mathbf{u}_1|\mathbf{v}_2)}{(\mathbf{u}_1|\mathbf{u}_1)} = \frac{\mathbf{u}_1^H\mathbf{v}_2}{\mathbf{u}_1^H\mathbf{u}_1} = \frac{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} i \\ -2 \\ i \end{bmatrix}}{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}} = \frac{i+4i+i}{1-4i^2+1} = \frac{6i}{6} = i$$

Dunque si ha

$$\mathbf{u}_2 = \mathbf{v}_2 - i\mathbf{u}_1 = \begin{bmatrix} i \\ -2 \\ i \end{bmatrix} - i \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

3. per costruire $\mathbf{u}_3$ dobbiamo essendo già a conoscenza di $\mathbf{u}_1$ e $\mathbf{u}_2$, calcolare i coefficienti α_{13} e α_{23}. Essendo $\mathbf{u}_1 \neq 0$ il primo è dato da

$$\alpha_{13} = \frac{(\mathbf{u}_1|\mathbf{v}_3)}{(\mathbf{u}_1|\mathbf{u}_1)} = \frac{\mathbf{u}_1^H\mathbf{v}_3}{\mathbf{u}_1^H\mathbf{u}_1} = \frac{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}}{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}} = \frac{12}{6} = 2$$

mentre, essendo $\mathbf{u}_2 = 0$ si prende $\alpha_{23} = 0$. Dunque si ha

$$\mathbf{u}_3 = \mathbf{v}_3 - 2\mathbf{u}_1 = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix}$$

4. infine per costruire $\mathbf{u}_4$ dopo aver già determinato i precedenti tre dobbiamo calcolare i coefficienti $\alpha_{14}, \alpha_{24}, \alpha_{34}$. Essendo $\mathbf{u}_1 \neq \mathbf{0}$ e $\mathbf{u}_3 \neq \mathbf{0}$ per ottenere il primo e il terzo si calcola

$$\alpha_{14} = \frac{(\mathbf{u}_1|\mathbf{v}_4)}{(\mathbf{u}_1|\mathbf{u}_1)} = \frac{\mathbf{u}_1^H\mathbf{v}_4}{\mathbf{u}_1^H\mathbf{u}_1} = \frac{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} 8i \\ 8 \\ 8i \end{bmatrix}}{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}} = \frac{0}{6} = 0$$

$$\alpha_{34} = \frac{(\mathbf{u}_3|\mathbf{v}_4)}{(\mathbf{u}_3|\mathbf{u}_3)} = \frac{\mathbf{u}_3^H\mathbf{v}_4}{\mathbf{u}_3^H\mathbf{u}_4} = \frac{\begin{bmatrix} 4 \\ 4i \\ 4 \end{bmatrix}^T \begin{bmatrix} 8i \\ 8 \\ 8i \end{bmatrix}}{\begin{bmatrix} 4 \\ 4i \\ 4 \end{bmatrix}^T \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix}} = \frac{96i}{48} = 2i$$

mentre essendo $\mathbf{u}_2 = \mathbf{0}$ si prende $\alpha_{24} = 0$. In definitiva si ricava

$$\mathbf{u}_4 = \mathbf{v}_4 - 2i\mathbf{u}_3 = \begin{bmatrix} 8i \\ 8 \\ 8i \end{bmatrix} - 2i \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Pertanto

$$\mathscr{S}^* = \left\{ \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix}, \mathbf{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

è un insieme ortogonale di generatori di V e l'insieme che si ottiene togliendo il vettore nullo:

$$\left\{ \mathbf{u}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix} \right\}$$

è una base ortogonale di V.

Osservazione 220. Il procedimento illustrato verrà impiegato soprattutto nel prossimo paragrafo per trovare una decomposizione $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$.

Osservazione 221. Nel caso però in cui l'obiettivo si limiti ad ottenere una base ortogonale di $V = \langle \mathscr{S} \rangle$ a partire da $\mathscr{S}$ è meglio (meno calcoli) trovare innanzitutto una generica base e poi applicare l'algoritmo per ottenerne una ortogonale.

Esempio 3.5.2. Riprendendo il sottoinsieme $\mathscr{S}$ di $\mathbb{C}^3$ dell'esempio 3.5.1, cambiando per comodità la notazione

$$\mathscr{S} = \left\{ \mathbf{w}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} i \\ -2 \\ i \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}, \mathbf{w}_4 = \begin{bmatrix} 8i \\ 8 \\ 8i \end{bmatrix} \right\}$$

Troviamo una generica base $\mathscr{B}$ di $V = \langle \mathscr{S} \rangle$; ossia, equivalentemente, una base di $\mathrm{C}(\mathbf{A})$ della matrice $\mathbf{A}$ avente per colonne gli elementi di $\mathscr{S}$ (per cui $\mathrm{C}(\mathbf{A}) = \langle \mathscr{S} \rangle = V$). Se

$$\mathbf{A} = \begin{bmatrix} 1 & i & 6 & 8i \\ 2i & -2 & 0 & 8 \\ 1 & i & 6 & 8i \end{bmatrix}$$

con un'eliminazione di Gauss su $\mathbf{A}$ si ottiene una sua forma ridotta $\mathbf{U}$

$$\begin{bmatrix} 1 & i & 6 & 8i \\ 2i & -2 & 0 & 8 \\ 1 & i & 6 & 8i \end{bmatrix} \begin{bmatrix} 1 & i & 6 & 8i \\ 0 & 0 & -12i & 24 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & i & 6 & 8i \\ 0 & 0 & 1 & i \\ 0 & 0 & 0 & 0 \end{bmatrix} = \mathbf{U}$$

Poiché le colonne dominanti di U sono la prima e la terza, allora una base $\mathscr{B}$ di $\mathrm{C}(\mathbf{A})$ (e quindi di V) è:

$$\mathscr{B} = \left\{ \mathbf{w}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix} \right\}$$

Applichiamo l'algoritmo di Gram-Schmidt all'insieme ottenuto per ottenere una base ortogonale; dopo aver posto $\mathbf{v}_1 = \mathbf{w}_1$ e $\mathbf{v}_2 = \mathbf{w}_3$:

1. prendiamo innanzitutto
$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}$$

2. per determinare $\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1$ dobbiamo calcolare α_{12} che essendo $\mathbf{u}_1 \neq \mathbf{0}$ è dato da
$$\alpha_{12} = \frac{(\mathbf{u}_1|\mathbf{v}_2)}{(\mathbf{u}_1|\mathbf{u}_1)} = \frac{\mathbf{u}_1^H\mathbf{v}_2}{\mathbf{u}_1^H\mathbf{u}_1} = \frac{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix}}{\begin{bmatrix} 1 \\ -2i \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}} = \frac{12}{6} = 2$$
per cui
$$\mathbf{u}_2 = \mathbf{v}_2 - 2\mathbf{u}_1 = \begin{bmatrix} 6 \\ 0 \\ 6 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix}$$
e quindi
$$\left\{ \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix} \right\}$$
è una base ortogonale di V.
Infine, volendo trovare una base ortonormale di V si normalizzano i vettori ricavati, dividendo ciascuno per la sua norma:
$$\|\mathbf{u}_1\|_2 = \sqrt{\mathbf{u}_1^H\mathbf{u}_1} = \sqrt{\begin{bmatrix} 1 & -2i & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}} = \sqrt{1+4+1} = \sqrt{6}$$
$$\|\mathbf{u}_2\|_2 = \sqrt{\mathbf{u}_2^H\mathbf{u}_2} = \sqrt{\begin{bmatrix} 4 & 4i & 4 \end{bmatrix} \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix}} = \sqrt{16+16+16} = 4\sqrt{3}$$
Pertanto:
$$\left\{ \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|_2} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix} ; \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|_2} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -i \\ 1 \end{bmatrix} \right\}$$
è una base ortonormale di V.

3.5.0.1 L'utilizzo di maxima

La funzione `gramschmidt` del pacchetto `eigen` effettua l'ortogonalizzazione fornendogli una matrice (lavorerà sulle righe) o una lista di liste (lavorerà sulle liste contenute).
Riproduciamo l'ortogonalizzazione della base dell'ultimo esempio:

```
##
## load("eigen")
## B:matrix([1,2*%i,1],[6,0,6])
## res:expand(gramschmidt(B))
##                          [[1, 2 %i, 1], [4, - 4 %i, 4]]
```

3.5.0.2 Esercizi

Esempio 3.5.3. Sia $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$; trovare una base ortogonale di $\mathrm{C}(\mathbf{A})$; determinare la proiezione ortogonale di $\mathbf{v} = \begin{bmatrix} -2 & 4 & 0 & 4 \end{bmatrix}^T$ su $\mathrm{C}(\mathbf{A})$.
Partiamo dalla riduzione per trovare le colonne dominanti:

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Tutte e tre le colonne sono dominanti, procediamo allora con l'algoritmo di Gram-Schmidt

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 0 \\ 0 \\ 1/2 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ 0 \\ -1/2 \end{bmatrix}$$

Quindi la base ortogonale è

$$\mathscr{B} = \left\{ \mathbf{u}_1 = \begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}0\\1\\1\\0\end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix}1/2\\0\\0\\-1/2\end{bmatrix} \right\}$$

Per determinare la proiezione ortogonale di $\mathbf{v}$ su $\mathrm{C}(\mathbf{A})$:

$$P_U\left(\begin{bmatrix}-2\\4\\0\\4\end{bmatrix}\right) = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 = \frac{\mathbf{u}_1^H\mathbf{v}}{\mathbf{u}_1^H\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{u}_2^H\mathbf{v}}{\mathbf{u}_2^H\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{u}_3^H\mathbf{v}}{\mathbf{u}_3^H\mathbf{u}_3}\mathbf{u}_3$$

$$= \frac{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}^T \begin{bmatrix}-2\\4\\0\\4\end{bmatrix}}{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}^T \begin{bmatrix}1\\0\\0\\1\end{bmatrix}} \begin{bmatrix}1\\0\\0\\1\end{bmatrix} + \frac{\begin{bmatrix}0\\1\\1\\0\end{bmatrix}^T \begin{bmatrix}-2\\4\\0\\4\end{bmatrix}}{\begin{bmatrix}0\\1\\1\\0\end{bmatrix}^T \begin{bmatrix}0\\1\\1\\0\end{bmatrix}} \begin{bmatrix}0\\1\\1\\0\end{bmatrix} + \frac{\begin{bmatrix}1/2\\0\\0\\-1/2\end{bmatrix}^T \begin{bmatrix}-2\\4\\0\\4\end{bmatrix}}{\begin{bmatrix}1/2\\0\\0\\-1/2\end{bmatrix}^T \begin{bmatrix}1/2\\0\\0\\-1/2\end{bmatrix}} \begin{bmatrix}1/2\\0\\0\\-1/2\end{bmatrix}$$

$$= \frac{2}{2}\begin{bmatrix}1\\0\\0\\1\end{bmatrix} + \frac{4}{2}\begin{bmatrix}0\\1\\1\\0\end{bmatrix} + \frac{-3}{1/2}\begin{bmatrix}1/2\\0\\0\\-1/2\end{bmatrix} = \begin{bmatrix}1\\0\\0\\1\end{bmatrix} + \begin{bmatrix}0\\2\\2\\0\end{bmatrix} + \begin{bmatrix}-6/2\\0\\0\\6/2\end{bmatrix}$$

$$= \begin{bmatrix}-2\\2\\2\\4\end{bmatrix}$$

Esempio 3.5.4. Sia

$$U = \left\langle \mathbf{v}_1 = \begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}0\\2\\0\\3\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}-2\\2\\0\\1\end{bmatrix} \right\rangle$$

1. si determini una base ortogonale di U
2. sia $\{\mathbf{u}_1, \mathbf{u}_2\}$ la base trovata. Per ogni $\mathbf{x} = \begin{bmatrix}a & b & c & d\end{bmatrix}^T \in U$ si trovino t_1, t_2 tali che $\mathbf{x} = t_1\mathbf{u}_1 + t_2\mathbf{u}_2$
3. si completi $\{\mathbf{u}_1, \mathbf{u}_2\}$ a una base ortogonale di $\mathbb{R}^4$

Per il primo punto partiamo dalla riduzione

$$\begin{bmatrix}1&0&-2\\0&2&2\\0&0&0\\1&3&1\end{bmatrix} \begin{bmatrix}1&0&-2\\0&2&2\\0&3&3\\0&0&0\end{bmatrix} \begin{bmatrix}1&0&-2\\0&6&6\\0&6&6\\0&0&0\end{bmatrix} \begin{bmatrix}1&0&-2\\0&6&6\\0&0&0\\0&0&0\end{bmatrix}$$

Quindi le prime due colonne sono dominanti e formano una base. Per ottenere la base ortogonale $\mathscr{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3 \end{bmatrix} - \frac{\mathbf{u}_1^H \mathbf{v}_2}{\mathbf{u}_1^H \mathbf{u}_1}\mathbf{u}_1 = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \\ 3 \end{bmatrix} - \frac{3}{2}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 2 \\ 0 \\ 3/2 \end{bmatrix}$$

Quindi una base è

$$\mathscr{B} = \left\{ \mathbf{u}_1' = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2' = \begin{bmatrix} -3/2 \\ 2 \\ 0 \\ 3/2 \end{bmatrix} \right\}$$

Ora per semplificare i calcoli trasformiamo il secondo elemento moltiplicandolo per due (rimane una base, rimane ortogonale) e adottiamo questa nel prosieguo:

$$\mathscr{B}' = \left\{ \mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -3 \\ 4 \\ 0 \\ 3 \end{bmatrix} \right\}$$

Per il secondo punto si ha che

$$\mathbf{x} = t_1\mathbf{u}_1 + t_2\mathbf{u}_2 = \frac{\mathbf{u}_1^H \mathbf{x}}{\mathbf{u}_1^H \mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{u}_2^H \mathbf{x}}{\mathbf{u}_2^H \mathbf{u}_2}\mathbf{u}_2$$

per cui

$$t_1 = \frac{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}} = \frac{a+d}{2}$$

mentre

$$t_2 = \frac{\begin{bmatrix}-3\\4\\0\\3\end{bmatrix}^T \begin{bmatrix}a\\b\\c\\d\end{bmatrix}}{\begin{bmatrix}-3\\4\\0\\3\end{bmatrix}^T \begin{bmatrix}-3\\4\\0\\3\end{bmatrix}} = \frac{-3a+4b+3d}{9+16+9} = \frac{-3a+4b+3d}{34}$$

Per il terzo punto aggiungiamo $\mathbf{I}_4$ alle colonne $\mathbf{u}_1$ e $\mathbf{u}_2$ ed effettuiamo la riduzione trovando le colonne dominanti

$$\begin{bmatrix}1&-3&1&0&0&0\\0&4&0&1&0&0\\0&0&0&0&1&0\\1&3&0&0&0&1\end{bmatrix}\begin{bmatrix}0&-6&1&0&0&-1\\0&4&0&1&0&0\\0&0&0&0&1&0\\1&3&0&0&0&1\end{bmatrix}\begin{bmatrix}0&12&2&0&0&-2\\0&12&0&3&0&0\\0&0&0&0&1&0\\1&3&0&0&0&1\end{bmatrix}\begin{bmatrix}0&0&2&-3&0&-2\\0&12&0&3&0&0\\0&0&0&0&1&0\\1&3&0&0&0&1\end{bmatrix}$$

Quindi le colonne dominanti sono quelle corrispondenti a $\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_1, \mathbf{e}_3$; applichiamo l'algoritmo alle ultime due (in quanto le prime due sono già state fatte

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{12}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix}1\\0\\0\\0\end{bmatrix} - \frac{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}^T \begin{bmatrix}1\\0\\0\\0\end{bmatrix}}{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}^T \begin{bmatrix}1\\0\\0\\1\end{bmatrix}} \begin{bmatrix}1\\0\\0\\1\end{bmatrix} - \frac{\begin{bmatrix}-3\\4\\0\\3\end{bmatrix}^T \begin{bmatrix}1\\0\\0\\0\end{bmatrix}}{\begin{bmatrix}-3\\4\\0\\3\end{bmatrix}^T \begin{bmatrix}-3\\4\\0\\3\end{bmatrix}} \begin{bmatrix}-3\\4\\0\\3\end{bmatrix}$$

$$= \begin{bmatrix}1\\0\\0\\0\end{bmatrix} - \frac{1}{2}\begin{bmatrix}1\\0\\0\\1\end{bmatrix} + \frac{3}{9+16+9}\begin{bmatrix}-3\\4\\0\\3\end{bmatrix} = \ldots = \begin{bmatrix}4/17\\6/17\\0\\-4/17\end{bmatrix}$$

Ora, sfruttando 2.3.8, moltiplico il vettore per 17, ottenendo $\begin{bmatrix}4\\6\\0\\-4\end{bmatrix}$ e poi per

1/2 ottenendo $\begin{bmatrix}2\\3\\0\\-2\end{bmatrix}$. Venendo all'ultimo vettore della base ortogonale

$\mathbf{u}_4 = \mathbf{v}_4 - \alpha_{12}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 - \alpha_{34}\mathbf{u}_3$

$$= \begin{bmatrix}0\\0\\1\\0\end{bmatrix} - \frac{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}^T \begin{bmatrix}0\\0\\1\\0\end{bmatrix}}{\begin{bmatrix}1\\0\\0\\1\end{bmatrix}^T \begin{bmatrix}1\\0\\0\\1\end{bmatrix}} \begin{bmatrix}1\\0\\0\\1\end{bmatrix} - \frac{\begin{bmatrix}-3\\4\\0\\3\end{bmatrix}^T \begin{bmatrix}0\\0\\1\\0\end{bmatrix}}{\cdots} \begin{bmatrix}-3\\4\\0\\3\end{bmatrix} - \frac{\begin{bmatrix}2\\3\\0\\-2\end{bmatrix}^T \begin{bmatrix}0\\0\\1\\0\end{bmatrix}}{\begin{bmatrix}2\\3\\0\\-2\end{bmatrix}^T \begin{bmatrix}2\\3\\0\\-2\end{bmatrix}} \begin{bmatrix}2\\3\\0\\-2\end{bmatrix} = \begin{bmatrix}0\\0\\1\\0\end{bmatrix} - 0 - 0 - 0 = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}$$

Quindi la base ortogonale è formata come

$$\mathscr{B} = \left\{ \begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}-3\\4\\0\\3\end{bmatrix}, \begin{bmatrix}2\\3\\0\\2\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix} \right\}$$

Esempio 3.5.5. Si consideri il sottoinsieme S di $\mathbb{C}^3$ definito da

$$S = \left\{ \begin{bmatrix}x\\y\\z\end{bmatrix} : 2x - y + z = 0 \right\}$$

1. si provi che S è un sottospazio vettoriale di $\mathbb{C}^3$
2. si trovi una base ortonormale di S
3. estendere tale base a una base ortonormale di $\mathbb{C}^3$

Per il primo punto $\mathbf{0} \in S$ in quanto

$$\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix} \implies 2 \cdot 0 - 0 + 0 = 0$$

Inoltre se $\mathbf{x} = \begin{bmatrix}x\\y\\z\end{bmatrix}, \mathbf{y} = \begin{bmatrix}x'\\y'\\z'\end{bmatrix} \in S$ si ha che $\alpha\mathbf{x} + \beta\mathbf{y} = \begin{bmatrix}\alpha x + \beta x'\\ \alpha y + \beta y'\\ \alpha z + \beta z'\end{bmatrix}$ e

$$\begin{aligned} 2(\alpha x + \beta x') - (\alpha y + \beta y') + \alpha z + \beta z' &= 2\alpha x + 2\beta x' - \alpha y - \beta y' + \alpha z + \beta z \\ &= \alpha \underbrace{(2x - y + z)}_{=0} + \beta \underbrace{(2x' - y' + z)}_{=0} \\ &= 0 \end{aligned}$$

per cui $\alpha\mathbf{x} + \beta\mathbf{y} \in S$.

Per il secondo punto se riarrangiamo l'equazione che definisce l'insieme come

$y = 2x + z$ (farlo in maniera alternativa è equivalente, si perverrà ad altre basi sempre ortonormali; procediamo così per check con le soluzioni degli esercizi) si ha che possiamo riscrivere i vettori dell'insieme come

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ 2x+z \\ z \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Con i vettori posti in evidenza a formare una base dello spazio considerato

$$\mathscr{B} = \left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Ortogonalizziamo la base mediante l'algoritmo GS

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - a_{12}\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{5} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 1/5 \\ 1 \end{bmatrix}$$

Moltiplico infine $\mathbf{u}_2$ per 5 ottenendo $\begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix}$. Per normalizzare la base ortogonale:

$$\|\mathbf{u}_1\| = \sqrt{1^2 + 2^2} = \sqrt{5}$$
$$\|\mathbf{u}_2\| = \sqrt{4 + 1 + 25} = \sqrt{30}$$

per cui la base ortonormale è

$$\mathscr{B}' = \left\{ \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{30}} \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} \right\}$$

Infine per il terzo punto partiamo dalla base non normalizzata, per facilitare il calcolo; effettuando la riduzione

$$\begin{bmatrix} 1 & -2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 0 & 5 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -4 & 2 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 0 & 5 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -4 & 2 & 0 & 0 \\ 0 & 5 & -2 & 1 & 0 \\ 0 & 5 & 0 & 0 & 1 \end{bmatrix}$$

concludiamo che le prime tre colonne (corrispondenti a $\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_1$ sono dominanti, quindi procediamo ad ortogonalizzare $\mathbf{e}_1$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{12}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{\begin{bmatrix}1\\2\\0\end{bmatrix}^T \begin{bmatrix}1\\0\\0\end{bmatrix}}{\begin{bmatrix}1\\2\\0\end{bmatrix}^T \begin{bmatrix}1\\2\\0\end{bmatrix}} \begin{bmatrix}1\\2\\0\end{bmatrix} - \frac{\begin{bmatrix}-2\\1\\5\end{bmatrix}^T \begin{bmatrix}1\\0\\0\end{bmatrix}}{\begin{bmatrix}-2\\1\\5\end{bmatrix}^T \begin{bmatrix}-2\\1\\5\end{bmatrix}} \begin{bmatrix}-2\\1\\5\end{bmatrix}$$

$$= \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{1}{5}\begin{bmatrix}1\\2\\0\end{bmatrix} - \frac{-2}{4+1+25}\begin{bmatrix}-2\\1\\5\end{bmatrix} = \ldots = \begin{bmatrix}2/3\\-1/3\\1/3\end{bmatrix}$$

Lo moltiplico per 3 ottenendo $\begin{bmatrix}2\\-1\\1\end{bmatrix}$ la cui norma è $\|\mathbf{u}_3\| = \sqrt{4+1+1} = \sqrt{6}$ per cui la base ortonormale di C^3 si ottiene aggiungendo a $\mathscr{B}'$ quest'ultimo normalizzato:

$$\mathscr{B}' = \left\{ \frac{1}{\sqrt{5}}\begin{bmatrix}1\\2\\0\end{bmatrix}, \frac{1}{\sqrt{30}}\begin{bmatrix}-2\\1\\5\end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix}2\\-1\\1\end{bmatrix} \right\}$$

Esempio 3.5.6. Sia

$$\mathbf{A} = \begin{bmatrix}1 & -2 & -1 & 3\\ -1 & 3 & -3 & 1\\ 1 & -1 & 1 & 1\\ 0 & 2 & -2 & 2\end{bmatrix}$$

1. si trovi una base ortogonale di $\mathrm{C}(\mathbf{A})$
2. si trovi una base ortogonale di $\mathrm{C}(\mathbf{A})^\perp$
3. si costruisca una base ortonormale di $\mathbb{C}^4$ a partire dalle due basi ortogonali di $\mathrm{C}(\mathbf{A})$ e $\mathrm{C}(\mathbf{A})^\perp$

Per il primo punto partiamo dal trovare una base effettuando la riduzione di gauss

$$\begin{bmatrix}1 & -2 & -1 & 3\\ -1 & 3 & -3 & 1\\ 1 & -1 & 1 & 1\\ 0 & 2 & -2 & 2\end{bmatrix} \begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 1 & -4 & 4\\ 0 & 1 & 2 & -2\\ 0 & 2 & -2 & 2\end{bmatrix} \begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 1 & -4 & 4\\ 0 & 2 & 4 & -4\\ 0 & 2 & -2 & 2\end{bmatrix} \begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 3 & 0 & 0\\ 0 & 2 & 4 & -4\\ 0 & 4 & -4 & 4\end{bmatrix}$$

$$\begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 3 & 0 & 0\\ 0 & 6 & 0 & 0\\ 0 & 4 & -4 & 4\end{bmatrix} \begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 6 & 0 & 0\\ 0 & 6 & 0 & 0\\ 0 & 4 & -4 & 4\end{bmatrix} \begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 6 & 0 & 0\\ 0 & 4 & -4 & 4\\ 0 & 0 & 0 & 0\end{bmatrix} \begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 2 & 0 & 0\\ 0 & 2 & -2 & 2\\ 0 & 0 & 0 & 0\end{bmatrix} \begin{bmatrix}1 & -2 & -1 & 3\\ 0 & 2 & -2 & 2\\ 0 & 0 & -2 & 2\\ 0 & 0 & 0 & 0\end{bmatrix}$$

Per cui sono dominanti le prime 3; trovo la base ortogonale basandomi su queste:

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1$$

$$= \begin{bmatrix} -2 \\ 3 \\ -1 \\ 2 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -2 \\ 3 \\ -1 \\ 2 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ -1 \\ 2 \end{bmatrix} - \frac{-2-3-1}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \\ -1 \\ 2 \end{bmatrix} + 2 \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} -1 \\ -3 \\ 1 \\ -2 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -1 \\ -3 \\ 1 \\ -2 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}^T \begin{bmatrix} -1 \\ 3 \\ 1 \\ -2 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ -3 \\ 1 \\ -2 \end{bmatrix} - \frac{-1+3+1}{1+1+1} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} - \frac{-3+1-4}{0+1+1+4} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ -3 \\ 1 \\ -2 \end{bmatrix} - \frac{3}{3} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} - \frac{-6}{6} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$

Quindi la base ortogonale è

$$\mathscr{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Per la base ortogonale di $\mathrm{C}(\mathbf{A})^{\perp}$ si ha che

$$\mathbf{w} = \begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix} \in \mathrm{C}(\mathbf{A})^{\perp} \iff (\mathbf{w}|\mathbf{v}_1) = 0 \wedge (\mathbf{w}|\mathbf{v}_2) = 0 \wedge (\mathbf{w}|\mathbf{v}_3) = 0$$

Che si traduce nel sistema

$$\begin{cases} x - y + w = 0 \\ y + w + 2z = 0 \\ -2x - y + w = 0 \end{cases} \rightarrow \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ -2 & -1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Risolviamo per $\mathbf{w}$ facendo la riduzione

$$\begin{bmatrix} 1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ -2 & -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -2 & 2 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ -2 & -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -3 & 3 & 0 & 0 \\ 0 & 1 & 1 & 2 & 0 \\ -2 & -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -3 & 3 & 0 & 0 \\ 0 & 3 & 3 & 6 & 0 \\ -2 & -1 & 1 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & -3 & 3 & 0 & 0 \\ 0 & 0 & 6 & 6 & 0 \\ -2 & -1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 1/2 & -1/2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & -1/2 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Che conduce a

$$\begin{cases} w + h = 0 \\ y - w = 0 \\ x + (1/2)y - (1/2)w = 0 \end{cases} \qquad \begin{cases} w = -h \\ y = -h \\ x - \frac{1}{2}h + \frac{1}{2}h = 0 \end{cases} \qquad \begin{cases} x = 0 \\ y = -h \\ w = -h \\ z = h \end{cases} \rightarrow h \begin{bmatrix} 0 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

Per la base ortonormale di $\mathbb{C}^4$ unisco le 2 basi e le normalizzo (dato che la dimensione della base ottenuta è 4, pari a quella di $\mathbb{C}^4$ e inoltre sono vettori ortogonali tra loro quindi formano una base di $\mathbb{C}^4$)

$$\left\{ \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \frac{1}{\sqrt{6}} \begin{bmatrix} -2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{3}} \begin{bmatrix} 0 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

3.6 Matrici di proiezione e decomposizioni QR

3.6.1 Matrici di proiezione

Definizione 3.6.1 (Matrice idempotente)**.** Se $\mathbf{A}^r = \mathbf{A}$, $\mathbf{A}$ si dice idempotente di ordine r; nel caso speciale $\mathbf{A}^2 = \mathbf{A}$, $\mathbf{A}$ si dice idempotente (tout court)

Definizione 3.6.2 (Matrice di proiezione)**.** Matrice (complessa o reale) idempotente ed hermitiana, ossia matrice $\mathbf{P}$ tale che:

$$\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^H \tag{3.35}$$

Osservazione 222. La motivazione della scelta del nome è data dal seguente teorema, che caratterizza le matrici di proiezione come quelle matrici che se usate come pre-moltiplicatore restituiscono la proiezione ortogonale di $\mathbb{C}^n$ su di un opportuno sottospazio U.

Teorema 3.6.1. *Una matrice quadrata* $\mathbf{P}$ *complessa di ordine* n *è una matrice di proiezione se e solo se esiste un sottospazio* U *di* $\mathbb{C}^n$ *tale che per ogni* $\mathbf{v} \in \mathbb{C}^n$ *si abbia* $\mathbf{P}\mathbf{v} = P_U(\mathbf{v})$*. In tal caso risulta* $U = \mathrm{C}(\mathbf{P})$*.*

Dimostrazione. Ordinatamente:

1. sia $\mathbf{P}$ una matrice di proiezione, ossia tale che $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^H$.
   Dato che $\mathbf{P} = \mathbf{P}^H$ segue che $\mathrm{N}\left(\mathbf{P}^H\right) = \mathrm{N}(\mathbf{P})$, per cui, essendo $\mathrm{C}(\mathbf{P})^\perp = \mathrm{N}\left(\mathbf{P}^H\right)$ (proposizione 3.3.4) si ha che

   $$\mathbb{C}^n = \mathrm{C}(\mathbf{P}) \oplus \mathrm{C}(\mathbf{P})^\perp = \mathrm{C}(\mathbf{P}) \oplus \mathrm{N}\left(\mathbf{P}^H\right) = \mathrm{C}(\mathbf{P}) \oplus \mathrm{N}(\mathbf{P})$$

   per la proposizione 3.3.7.
   Da $\mathbf{P}^2 = \mathbf{P}$ segue poi che $\mathbf{P}(\mathbf{P}\mathbf{e}_i) = (\mathbf{P}\mathbf{P})\mathbf{e}_i = \mathbf{P}\mathbf{e}_i$ per ogni $i = 1, \ldots, n$ dove $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ è la base canonica di $\mathbb{C}^n$. Dunque ricordando che $\mathbf{P}\mathbf{e}_i$ è la i-esima colonna di $\mathbf{P}$ si ottiene che la pre-moltiplicazione per $\mathbf{P}$ è una applicazione lineare di $\mathbb{C}^n$ che:

   - ristretta a $\mathrm{C}(\mathbf{P})$ è l'identità: se pre-moltiplichiamo per $\mathbf{P}$ un elemento dello spazio delle colonne di $\mathbf{P}$, ad esempio $\mathbf{v} = \alpha_1\mathbf{P}\mathbf{e}_1 + \ldots + \alpha_n\mathbf{P}\mathbf{e}_n \in \mathrm{C}(\mathbf{P})$, si ottiene $\mathbf{v}$ stesso in quanto:

     $$\begin{aligned}\mathbf{P}\mathbf{v} &= \mathbf{P}(\alpha_1\mathbf{P}\mathbf{e}_1 + \ldots + \alpha_n\mathbf{P}\mathbf{e}_n) = \alpha_1\mathbf{P}\mathbf{P}\mathbf{e}_1 + \ldots + \alpha_n\mathbf{P}\mathbf{P}\mathbf{e}_n \\ &= \alpha_1\mathbf{P}\mathbf{e}_1 + \ldots + \alpha_n\mathbf{P}\mathbf{e}_n \\ &= \mathbf{v}\end{aligned}$$

   - ristretta a $\mathrm{N}(\mathbf{P})$ è l'applicazione nulla: se pre-moltiplichiamo un elemento di $\mathrm{N}(\mathbf{P})$ per $\mathbf{P}$ per definizione di $\mathrm{N}(\mathbf{P})$ si ha $\mathbf{0}$ sempre, quindi è l'applicazione nulla.

   Pertanto la pre-moltiplicazione per $\mathbf{P}$ è la proiezione ortogonale di $\mathbb{C}^n$ su $\mathrm{C}(\mathbf{P})$, in base alla definizione 3.3.4 (con $\mathrm{C}(\mathbf{P})$ e $\mathrm{N}(\mathbf{P})$ a fungere da U e $U^\perp$ rispettivamente);

2. viceversa siano U un sottospazio di $\mathbb{C}^n$ e $\mathbf{P}$ la (unica) matrice tale che $\mathbf{P}\mathbf{v} = P_U(\mathbf{v})$ per ogni $\mathbf{v} \in \mathbb{C}^n$, definita come

   $$\mathbf{P} = \begin{bmatrix}\mathbf{P}\mathbf{e}_1 & \mathbf{P}\mathbf{e}_2 & \ldots & \mathbf{P}\mathbf{e}_n\end{bmatrix} = \begin{bmatrix}P_U(\mathbf{e}_1) & P_U(\mathbf{e}_2) & \ldots & P_U(\mathbf{e}_n)\end{bmatrix}$$

   Prendendo $V = \mathbb{C}^n$ con il prodotto interno standard si ottiene che $\mathbf{P}^2 = \mathbf{P}$, infatti:

   $$\begin{aligned}\mathbf{P}\mathbf{P} &= \mathbf{P}\begin{bmatrix}P_U(\mathbf{e}_1) & \ldots & P_U(\mathbf{e}_n)\end{bmatrix} = \begin{bmatrix}\mathbf{P}P_U(\mathbf{e}_1) & \ldots & \mathbf{P}P_U(\mathbf{e}_n)\end{bmatrix} \\ &\overset{(1)}{=} \begin{bmatrix}P_U(P_U(\mathbf{e}_1)) & \ldots & P_U(P_U(\mathbf{e}_n))\end{bmatrix} \overset{(2)}{=} \begin{bmatrix}P_U(\mathbf{e}_1) & \ldots & P_U(\mathbf{e}_n)\end{bmatrix} \\ &= \mathbf{P}\end{aligned}$$

   dove (1) dato che $\mathbf{P}\mathbf{v} = P_U(\mathbf{v})$ mentre (2) è giustificato da 3.3.12 (punto 3).

Similmente si dimostra che $\mathbf{P} = \mathbf{P}^H$, in quanto il prodotto $\mathbf{P}^H\mathbf{P} = \mathbf{P}^H\mathbf{P}^H$

$$\begin{aligned}
\mathbf{P}^H\mathbf{P} &= \begin{bmatrix}\mathbf{P}^H P_U(\mathbf{e}_1) & \dots & \mathbf{P}^H P_U(\mathbf{e}_n)\end{bmatrix} \\
&\overset{(1)}{=} \begin{bmatrix} \mathbf{v}_1^H P_U(\mathbf{e}_1) & \dots & \mathbf{v}_1^H P_U(\mathbf{e}_n) \\ \mathbf{v}_2^H P_U(\mathbf{e}_1) & \dots & \mathbf{v}_2^H P_U(\mathbf{e}_n) \\ \dots & \dots & \dots \\ \mathbf{v}_n^H P_U(\mathbf{e}_1) & \dots & \mathbf{v}_n^H P_U(\mathbf{e}_n) \end{bmatrix} \overset{(2)}{=} \begin{bmatrix} P_U(\mathbf{v}_1)^H \mathbf{e}_1 & \dots & P_U(\mathbf{v}_1)^H \mathbf{e}_n \\ P_U(\mathbf{v}_2)^H \mathbf{e}_1 & \dots & P_U(\mathbf{v}_2)^H \mathbf{e}_n \\ \dots & \dots & \dots \\ P_U(\mathbf{v}_n)^H \mathbf{e}_1 & \dots & P_U(\mathbf{v}_n)^H \mathbf{e}_n \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{v}_1^H \mathbf{P}^H \mathbf{e}_1 & \dots & \mathbf{v}_1^H \mathbf{P}^H \mathbf{e}_n \\ \mathbf{v}_2^H \mathbf{P}^H \mathbf{e}_1 & \dots & \mathbf{v}_2^H \mathbf{P}^H \mathbf{e}_n \\ \dots & \dots & \dots \\ \mathbf{v}_n^H \mathbf{P}^H \mathbf{e}_1 & \dots & \mathbf{v}_n^H \mathbf{P}^H \mathbf{e}_n \end{bmatrix} \\
&\overset{(3)}{=} \mathbf{P}^H\mathbf{P}^H
\end{aligned}$$

dove per (1), ad esempio $\mathbf{v}_1$ è la prima colonna di $\mathbf{P}$ e $\mathbf{v}_1^H$ è la prima riga di $\mathbf{P}^H$; in (2) si è sfruttato il teorema 3.3.12 punto 4; in 3 si presti attenzione al fatto che (ad esempio) $\mathbf{v}_1^H$ è la prima riga di $\mathbf{P}^H$, mentre $\mathbf{P}^H\mathbf{e}_1$ è la prima colonna di $\mathbf{P}_H$, quindi quello che sta avvenendo è un semplice prodotto di $\mathbf{P}^H$ per $\mathbf{P}^H$. Avendo mostrato che $\mathbf{P}^H\mathbf{P} = \mathbf{P}^H\mathbf{P}^H$ arriviamo a $\mathbf{P} = \mathbf{P}^H$ vedendo che

$$\mathbf{P}^H\mathbf{P} = \mathbf{P}^H\mathbf{P}^H = (\mathbf{P}\mathbf{P})^H = (\mathbf{P}^2)^H = \mathbf{P}^H$$

dove nell'ultima si è sfruttato $\mathbf{P}^2 = \mathbf{P}$. Pertanto si ha che

$$\mathbf{P}^H\mathbf{P} = \mathbf{P}^H \tag{3.36}$$

Prendendo l'H-trasposta di entrambi i membri:

$$\mathbf{P} = (\mathbf{P}^H\mathbf{P})^H = \mathbf{P}^H\mathbf{P}$$

quindi

$$\mathbf{P} = \mathbf{P}^H\mathbf{P} \tag{3.37}$$

Pertanto, considerando congiuntamente 3.36 e 3.37

$$\mathbf{P} = \mathbf{P}^H\mathbf{P} = \mathbf{P}^H$$

e come si voleva, $\mathbf{P} = \mathbf{P}^H$.
Riassumendo quanto trovato, si ha che $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^H$ ossia che $\mathbf{P}$ è una matrice di proiezione.

La prima parte del teorema da ora $U = \mathrm{C}(\mathbf{P})$. □

Lemma 3.6.2. *Sia* $\mathbf{Q}$ *una matrice* $n \times k$*. Allora* $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}_k$ *se e solo se le colonne di* $\mathbf{Q}$ *sono ortonormali.*

Dimostrazione. Se $\mathbf{Q} = [\mathbf{u}_1 \dots \mathbf{u}_k]$ allora

$$\mathbf{Q}^H = \begin{bmatrix} \mathbf{u}_1^H \\ \dots \\ \mathbf{u}_k^H \end{bmatrix}$$

per cui, calcolando il prodotto $\mathbf{Q}^H\mathbf{Q}$ a blocchi si ottiene che l'elemento di posto (i,j) di $\mathbf{Q}^H\mathbf{Q}$ è $\mathbf{u}_i^H\mathbf{u}_j = (\mathbf{u}_i|\mathbf{u}_j)$. Quindi

$$\mathbf{Q}^H\mathbf{Q} = \mathbf{I}_k \iff (\mathbf{u}_i|\mathbf{u}_j) = \begin{cases} 0 & \text{se } i \neq j \\ 1 & \text{se } i = j \end{cases} \overset{(1)}{\iff} \text{le colonne di } \mathbf{Q} \text{ sono ortonormali}$$

dove la (1) segue anche da osservazione 217. □

Osservazione 223. Il teorema che segue fornisce uno strumento per il calcolo effettivo della matrice di proiezione su di un sottospazio U di $\mathbb{C}^n$.

Teorema 3.6.3. *Una matrice complessa* $\mathbf{P}$ *quadrata di ordine* n *e rango* k *è di proiezione sul sottospazio* U *di* $\mathbb{C}^n$ *se e solo se* $\mathbf{P} = \mathbf{Q}\mathbf{Q}^H$*, dove* $\mathbf{Q}$ *è una matrice* $n \times k$ *le cui colonne sono i vettori di una base ortonormale di* U*.*

Dimostrazione. Rispettivamente:

- partiamo mostrando che se $\mathbf{P} = \mathbf{Q}\mathbf{Q}^H$, dove $\mathbf{Q}$ è una matrice con colonne ortonormali, allora

$$\mathbf{P}^H = (\mathbf{Q}\mathbf{Q}^H)^H = (\mathbf{Q}^H)^H\mathbf{Q}^H = \mathbf{Q}\mathbf{Q}^H = \mathbf{P}$$

(ossia $\mathbf{P}$ è hermitiana in quanto $\mathbf{P} = \mathbf{P}^H$) e

$$\mathbf{P}^2 = \mathbf{P}\mathbf{P} = \underbrace{\mathbf{Q}\mathbf{Q}^H}_{\mathbf{P}}\mathbf{Q}\mathbf{Q}^H \overset{(1)}{=} \mathbf{Q}\mathbf{I}_k\mathbf{Q}^H = \mathbf{Q}\mathbf{Q}^H = \mathbf{P}$$

dove (1) è in base al lemma 3.6.2. Dunque $\mathbf{P}^H = \mathbf{P} = \mathbf{P}^2$ e $\mathbf{P}$ è una matrice di proiezione.

- mostriamo invece che se $\mathbf{P}$ è matrice di proiezione, ossia tale che $\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^H$, allora si può fattorizzare in $\mathbf{Q}\mathbf{Q}^H$.
Per il teorema 3.6.1 esiste un sottospazio U di $\mathbb{C}^n$ tale che $\mathbf{P}\mathbf{v} = P_U(\mathbf{v})$ per ogni $\mathbf{v} \in \mathbb{C}^n$. Poiché inoltre $U = \mathrm{C}(\mathbf{P})$, $\dim U = \dim \mathrm{C}(\mathbf{P}) = \mathrm{rk}\,\mathbf{P} = k$ e dunque $k = \dim U$.
Dalla proposizione 3.4.4 segue che se $\mathscr{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ è una base ortonormale di U, allora la proiezione di U su V è (osservazione 217):

$$P_U(\mathbf{v}) = \sum_{i=1}^{k} (\mathbf{u}_i|\mathbf{v})\mathbf{u}_i, \forall \mathbf{v} \in \mathbb{C}^n$$

Quindi si ottiene:

$$\begin{aligned}\mathbf{P}\mathbf{v} = P_U(\mathbf{v}) &= \sum_{i=1}^{k} (\mathbf{u}_i|\mathbf{v})\mathbf{u}_i = \sum_{i=1}^{k} \mathbf{u}_i^H\mathbf{v}\mathbf{u}_i \overset{(1)}{=} \sum_{i=1}^{k} \mathbf{u}_i\mathbf{u}_i^H\mathbf{v} \\ &\overset{(2)}{=} \left(\sum_{i=1}^{k} \mathbf{u}_i\mathbf{u}_i^H\right)\mathbf{v} \overset{(3)}{=} \mathbf{Q}\mathbf{Q}^H\mathbf{v}\end{aligned}$$

dove $\mathbf{Q} = [\mathbf{u}_1 \ \ldots \ \mathbf{u}_k]$ è una matrice che ha come colonne i vettori di $\mathscr{B}$ e

  - in (1) si è solo spostato la costante $\mathbf{u}_i^H\mathbf{v}$ dopo $\mathbf{u}$ invece che prima;

– in (2) si è portato fuori da sommatoria $\mathbf{v}$ dato che non dipende dall'indice;

– in (3) si ha che $\sum_i \mathbf{u}_i\mathbf{u}_i^H = \mathbf{Q}\mathbf{Q}^H$, è un prodotto a blocchi colonne per righe (36), dove le colonne di $\mathbf{Q}$ sono $\mathbf{u}_i$ e le righe di $\mathbf{Q}^H$ sono $\mathbf{u}_i^H$.

TODO: Il resta da provare non è chiarissimo in tutti i punti ...

Resta da provare che le colonne di $\mathbf{Q}$ sono i vettori di una base ortonormale di $U = \mathrm{C}(\mathbf{P})$ ossia, dato che $U = \mathrm{C}(\mathbf{Q})$, che $\mathrm{C}(\mathbf{Q}) = \mathrm{C}(\mathbf{P})$. Ciò segue da

$$\begin{aligned}\mathrm{C}(\mathbf{P}) &= \langle \mathbf{P}\mathbf{e}_i : i = 1, \dots, n\rangle = \left\langle (\mathbf{Q}\mathbf{Q}^H)\mathbf{e}_i : i = 1, \dots, n\right\rangle \\ &= \left\langle \mathbf{Q}(\mathbf{Q}^H\mathbf{e}_i) : i = 1, \dots, n\right\rangle \subseteq \mathrm{C}(\mathbf{Q})\end{aligned}$$

dove $\mathbf{Q}^H\mathbf{e}_i$ è la i-esima colonna di $\mathbf{Q}^H$, ossia la i-esima riga di $\mathbf{Q}$, fornisce gli elementi per la combinazione lineare delle colonne di $\mathbf{Q}$ (pensandolo in termini di prodotti a blocchi colonne per righe).
Infine da

$$\dim \mathrm{C}(\mathbf{P}) = \operatorname{rk}\mathbf{P} \overset{(1)}{=} k \overset{(2)}{=} \operatorname{rk}\mathbf{Q} = \dim \mathrm{C}(\mathbf{Q})$$

dove (1) è dato dal fatto che $\mathbf{P}$ ha rango k per ipotesi e (2) dal fatto che k è il numero di colonne indipendenti essendo $\mathbf{Q}$ (di tipo $n \times k$) composta da colonne ortonormali. Si conclude con

$$\dim \mathrm{C}(\mathbf{Q}) = \dim \mathrm{C}(\mathbf{P}) \iff \mathrm{C}(\mathbf{Q}) = \mathrm{C}(\mathbf{P})$$

□

Osservazione 224. Dal teorema 3.6.3 segue che $\mathbf{P}$ dipende dal sottospazio $\mathrm{C}(\mathbf{P}) = \mathrm{C}(\mathbf{Q})$, ma è indipendente dalla particolare base ortonormale di esso che si sceglie per costruire $\mathbf{Q}$, come è illustrato dal seguente esempio.

Esempio 3.6.1. Si consideri il sottospazio $U = \left\langle \mathbf{v}_1 = [1\,1\,1]^T; \mathbf{v}_2 = [0\,1\,0]^T\right\rangle$ di $\mathbb{R}^3$. Applicando l'algoritmo di Gram-Schmidt a $\{\mathbf{v}_1, \mathbf{v}_2\}$ e normalizzando poi i vettori che si ottengono si costruisce una base ortonormale di U:

$$\mathscr{B}_1 = \left\{ \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix}-1\\2\\-1\end{bmatrix} \right\}$$

Per il teorema 3.6.3 posto

$$\mathbf{Q}_1 = \begin{bmatrix}1/\sqrt{3} & -1/\sqrt{6}\\ 1/\sqrt{3} & 2/\sqrt{6}\\ 1/\sqrt{3} & -1/\sqrt{6}\end{bmatrix}$$

la matrice di proiezione su U è:

$$\mathbf{P}_1 = \mathbf{Q}_1\mathbf{Q}_1^H = \begin{bmatrix}1/\sqrt{3} & -1/\sqrt{6}\\ 1/\sqrt{3} & 2/\sqrt{6}\\ 1/\sqrt{3} & -1/\sqrt{6}\end{bmatrix} \begin{bmatrix}1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3}\\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6}\end{bmatrix} = \begin{bmatrix}1/2 & 0 & 1/2\\ 0 & 1 & 0\\ 1/2 & 0 & 1/2\end{bmatrix}$$

Vediamo ora che si giunge alla stessa matrice di proiezione partendo da una base differente; se $\mathbf{w}_1 = \mathbf{v}_1 - \mathbf{v}_2$ si ha che $\mathbf{w}_1 \perp \mathbf{v}_2$ e $U = \langle \mathbf{w}_1, \mathbf{v}_2 \rangle$. Quindi normalizzando $\mathbf{w}_1 = [1\,0\,1]^T$ si ricava che anche

$$\mathscr{B}_2 = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

è una base ortonormale di U. Anche la matrice

$$\begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix}$$

costruita a partire dalla base ortonormale $\mathscr{B}_2$ di U è diversa da $\mathbf{Q}_1$, costruita a partire dalla base ortonormale $\mathscr{B}_1$ di U, si ha che

$$\mathbf{P}_2 = \mathbf{Q}_2\mathbf{Q}_2^H = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} = \mathbf{P}_1$$

Proposizione 3.6.4. *Dato un sottospazio U di $\mathbb{C}^n$ con matrice di proiezione $\mathbf{P}$, la matrice di proiezione sul completamento ortogonale $U^\perp$ di U è $\mathbf{I}_n - \mathbf{P}$.*

Dimostrazione. Siano U un sottospazio di $\mathbb{C}^n$ e $\mathbf{P}$ la matrice di proiezione su U. Sappiamo che essendo $\mathbb{C}^n = U \oplus U^\perp$, per la proposizione 3.3.7; per ogni $\mathbf{v} \in \mathbb{C}^n$ esistono $\mathbf{u} \in U$ e $\mathbf{w} \in U^\perp$ tali che $\mathbf{v} = \mathbf{u} + \mathbf{w}$. Dal teorema 3.6.1 segue che $\mathbf{P}\mathbf{v} = P_U(\mathbf{v}) = \mathbf{u}$, e quindi

$$(\mathbf{I}_n - \mathbf{P})\mathbf{v} = \mathbf{v} - \mathbf{P}\mathbf{v} = \mathbf{v} - \mathbf{u} = \mathbf{w} = P_{U^\perp}(\mathbf{v})$$

Pertanto, sempre per il teorema 3.6.1 si conclude. □

3.6.2 Decomposizioni QR

Osservazione 225. Mostriamo innanzitutto come ogni matrice $\mathbf{A}$ possa esser fattorizzata nel prodotto di una matrice a colonne ortogonali e una matrice unitriangolare superiore. Tale decomposizione si chiama decomposizione QR non normalizzata di $\mathbf{A}$.
Partendo da essa, costruiremo poi una fattorizzazione di $\mathbf{A}$ nel prodotto di una matrice a colonne ortonormali e una matrice a scala per righe (decomposizione QR normalizzata di $\mathbf{A}$)

3.6.2.1 Non normalizzata

Teorema 3.6.5 (Decomposizione QR non normalizzata di $\mathbf{A}$)**.** *Sia $\mathbf{A}$ una matrice $m \times n$. Allora esistono una matrice $m \times n$ $\mathbf{Q}_0$ a colonne ortogonali e una matrice unitriangolare superiore di ordine n $\mathbf{R}_0$ tali che $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$.*

Dimostrazione. Sia $\mathbf{A} = [\mathbf{v}_1\, \mathbf{v}_2\, \dots\, \mathbf{v}_n]$ una matrice (reale o complessa) $m \times n$. Applicando l'algoritmo di Gram-Schmidt alle colonne $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ di $\mathbf{A}$ si ottengono i vettori ortogonali $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ definiti nel modo seguente

$$\begin{aligned}
&\mathbf{u}_1 = \mathbf{v}_1 \\
&\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 \\
&\dots \\
&\mathbf{u}_k = \mathbf{v}_k - \alpha_{1k}\mathbf{u}_1 - \alpha_{2k}\mathbf{u}_2 - \dots - \alpha_{k-1,k}\mathbf{u}_{k-1}
\end{aligned} \tag{3.38}$$

per ogni $1 < k \leq n$ con

$$\alpha_{ij} = \begin{cases} 0 & \text{se } \mathbf{u}_i = \mathbf{0} \\ \frac{(\mathbf{u}_i|\mathbf{v}_j)}{(\mathbf{u}_i|\mathbf{u}_i)} & \text{se } \mathbf{u}_i \neq \mathbf{0} \end{cases}$$

per ogni $1 \leq i \neq j \leq n$ dove il prodotto interno è quello standard.
Siano $\mathbf{Q}_0 = [\mathbf{u}_1\, \mathbf{u}_2\, \dots\, \mathbf{u}_n]$ la matrice che ha come colonne i vettori $\mathbf{u}_1, \dots \mathbf{u}_n$ ed $\mathbf{R}_0$ la seguente matrice unitriangolare superiore costruita con i coefficienti α_{ij}

$$\mathbf{R}_0 = \begin{bmatrix} 1 & \alpha_{12} & \dots & \alpha_{1,n-1} & \alpha_{1n} \\ 0 & 1 & \dots & \alpha_{2,n-1} & \alpha_{2n} \\ \dots & & & & \\ 0 & 0 & \dots & 1 & \alpha_{n-1,n} \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

Per ogni $k = 1, \dots, n$ si ha

$$\begin{aligned}
\mathbf{Q}_0(\mathbf{R}_0\mathbf{e}_k) &\overset{(1)}{=} \mathbf{Q}_0\left(\mathbf{e}_k + \sum_{i=1}^{k-1}\alpha_{ik}\mathbf{e}_i\right) = \mathbf{Q}_0\mathbf{e}_k + \sum_{i=1}^{k-1}\alpha_{ik}\mathbf{Q}_0\mathbf{e}_i \\
&= \mathbf{u}_k + \sum_{i=1}^{k-1}\alpha_{ik}\mathbf{u}_i \overset{(2)}{=} \mathbf{v}_k \\
&= \mathbf{A}\mathbf{e}_k = k\text{-esima colonna di } \mathbf{A}
\end{aligned}$$

dove in (1) ci si è limitati a spezzare la k-esima colonna di $\mathbf{R}_0$ in due parti, un vettore che ha 1 al k-esimo posto e ha gli a_{ik} in ordine sino al posto $k-1$ e poi tutti 0) e in (2) si è sfruttata equazione 3.38.
Per cui se $\mathbf{Q}_0\mathbf{R}_0\mathbf{e}_k = \mathbf{A}\mathbf{e}_k$ allora $\mathbf{Q}_0\mathbf{R}_0 = \mathbf{A}$. □

Osservazione 226. Sia $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$ una decomposizione QR non normalizzata di una matrice $m \times n$ $\mathbf{A}$ di rango k.
Poiché $\mathbf{R}_0$ è invertibile (essendo triangolare soddisfa il punto 3 della 1.6.12 e dunque anche il punto 1), da $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$ segue che $\mathrm{C}(\mathbf{A}) = \mathrm{C}(\mathbf{Q}_0)$. Infatti si ha che $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$, ma $\mathbf{R}_0$ è invertibile, con inversa $\mathbf{R}_0^{-1}$, quindi post-moltiplico ottenendo $\mathbf{Q}_0 = \mathbf{A}\mathbf{R}_0^{-1}$. Le colonne di $\mathbf{Q}_0$ sono combinazioni lineari delle colonne di $\mathbf{A}$; infatti $\mathbf{Q}_0 = [\mathbf{A}\mathbf{R}_0^{-1}\mathbf{e}_1 \dots \mathbf{A}\mathbf{R}_0^{-1}\mathbf{e}_n]$ ma sfruttando la decomposizione a blocchi di 36 e vedendo $\mathbf{v}_i = \mathbf{R}_0^{-1}\mathbf{e}_i$ si ha sulla singola colonna che $\mathbf{A}\mathbf{v}_i = a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n$ con $\mathbf{v}_i = [v_1, \dots v_n]$ e $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_n]$.
In particolare, essendo le colonne di $\mathbf{Q}_0$ ortogonali e $k = \mathrm{rk}\,\mathbf{A} = \dim \mathrm{C}(\mathbf{A}) = \dim \mathrm{C}(\mathbf{Q}_0) = \mathrm{rk}\,\mathbf{Q}_0$, dalla proposizione 3.4.1 si ottiene che $n-k$ colonne di $\mathbf{Q}_0$ sono nulle. Inoltre se $k > 1$ per come abbiamo costruito $\mathbf{Q}_0$ e dal teorema 3.5.2 segue che $V_{k-1} = \langle \mathbf{u}_1; \dots; \mathbf{u}_{k-1} \rangle$ allora $V_{k-1} = \langle \mathbf{v}_1; \dots; \mathbf{v}_{k-1} \rangle$ (i $\mathbf{v}_i$ possono esser visti come una combinazione lineare degli $\mathbf{u}_i$ se riarrangiamo algebricamente

le equazioni di Gram-Schmidt; quindi generano lo stesso spazio) e

$$\mathbf{u}_k = \mathbf{v}_k - P_{V_{k-1}}(\mathbf{v}_k)$$

Ciò mostra che ogni vettore $\mathbf{u}_k$ ottenuto con l'algoritmo di Gram-Schmidt non è altro che la proiezione ortogonale di $\mathbf{v}_k$ sul sottospazio complemento ortogonale di V_{k-1}. Inoltre per il teorema 3.3.12 $\|\mathbf{u}_k\|_2$ è la distanza di $\mathbf{v}_k$ da V_{k-1} (si veda la definizione 3.3.5).

Esempio 3.6.2. Sia

$$\mathbf{A} = \begin{bmatrix} 1 & i & 6 & 8i \\ 2i & -2 & 0 & 8 \\ 1 & i & 6 & 8i \end{bmatrix}$$

Applicando l'algoritmo di Gram-Schmidt all'insieme $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ delle colonne di $\mathbf{A}$ si ottengono i vettori

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \mathbf{v}_2 - i\mathbf{u}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \mathbf{v}_3 - 2\mathbf{u}_1 = \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix}$$

$$\mathbf{u}_4 = \mathbf{v}_4 - \alpha_{14}\mathbf{u}_1 - \alpha_{24}\mathbf{u}_2 - \alpha_{34}\mathbf{u}_3 = \mathbf{v}_4 - 2i\mathbf{u}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Pertanto posto

$$\mathbf{Q}_0 = \begin{bmatrix} 1 & 0 & 4 & 0 \\ 2i & 0 & -4i & 0 \\ 1 & 0 & 4 & 0 \end{bmatrix}, \quad \mathbf{R}_0 = \begin{bmatrix} 1 & i & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2i \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

si ha che $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$ è una decomposizione QR non normalizzata di $\mathbf{A}$.

3.6.2.2 Normalizzata

Teorema 3.6.6 (Decomposizione QR normalizzata di $\mathbf{A}$)**.** *Sia $\mathbf{A}$ una matrice $m \times n$ di rango k. Allora esistono una matrice $m \times k$ $\mathbf{Q}$ a colonne ortonormali e una matrice a scala per righe $k \times n$ $\mathbf{R}$ tali che $\mathbf{A} = \mathbf{QR}$.*

Dimostrazione. Se $\mathbf{Q}_1$ è la matrice $m \times k$ ottenuta da $\mathbf{Q}_0$ (della non normalizzata) sopprimendone le colonne nulle ed $\mathbf{R}_1$ è la matrice $k \times n$ ottenuta da $\mathbf{R}_0$ (della non normalizzata) sopprimendone le righe di indice corrispondente a quello delle colonne nulle di $\mathbf{Q}_0$, allora $\mathbf{Q}_0\mathbf{R}_0 = \mathbf{Q}_1\mathbf{R}_1$. Infatti sono entrambe matrici $m \times n$ ed eseguendo i due prodotti, a blocchi, suddividendo $\mathbf{Q}_0$ e $\mathbf{Q}_1$ in righe ed $\mathbf{R}_0$ e $\mathbf{R}_1$ in colonne, si ha che al posto (i, j) del prodotto $\mathbf{Q}_0\mathbf{R}_0$ sarà $\mathbf{q}_i\mathbf{r}_j$ con $\mathbf{q}_i$ riga i-esima di $\mathbf{Q}_0$. Ora, se per esempio $\mathbf{q}_i$ ha coefficiente nullo al posto k (k esima

colonna di $\mathbf{Q}_0$ nulla), ossia $\mathbf{q}_i = [q_1 \dots 0 \dots q_n]$ e $\mathbf{r}_j = \begin{bmatrix} r_1 & \dots & r_k & \dots & r_n \end{bmatrix}^T$, il prodotto sarà $\mathbf{q}_i\mathbf{r}_j = r_1q_1 + \ldots + 0 \cdot r_k + \ldots r_nq_n$; che è uguale al prodotto che si ottiene in $\mathbf{Q}_1\mathbf{R}_1$, ovvero $r_1q_1 + \ldots + r_nq_n$.
Sia $\mathbf{D}$ la matrice diagonale di ordine k i cui gli elementi diagonali sono le norme euclidee delle colonne di $\mathbf{Q}_1$ e poniamo $\mathbf{QD} = \mathbf{Q}_1$; poiché le colonne di $\mathbf{Q}_1$ sono vettori non nulli, le loro norme euclidee sono diverse da 0, per cui $\mathbf{D}$ è invertibile (matrice diagonale con diagonale principale non nulla), e dunque post moltiplicando si ottiene che $\mathbf{Q} = \mathbf{Q}_1\mathbf{D}^{-1}$ (dalla quale $\mathbf{Q}$ è composto di vettori ortonormali). Posto infine $\mathbf{R} = \mathbf{DR}_1$ si ha che $\mathbf{QR} = \mathbf{Q}_1\mathbf{D}^{-1}\mathbf{DR}_1 = \mathbf{Q}_1\mathbf{R}_1$, per cui $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0 = \mathbf{Q}_1\mathbf{R}_1 = \mathbf{QR}$ □

Osservazione 227. Sia $\mathbf{A} = \mathbf{QR}$ una decomposizione QR normalizzata di una matrice $m \times n$ $\mathbf{A}$ di rango k. Poiché la i-esima colonna di $\mathbf{Q}$, per $i = 1, \ldots, k$ è la i-esima colonna di $\mathbf{Q}_1$ normalizzata (ossia moltiplicata per una costante), allora

$$\mathrm{C}(\mathbf{Q}) = \mathrm{C}(\mathbf{Q}_1) \overset{(1)}{=} \mathrm{C}(\mathbf{Q}_0) \overset{(2)}{=} \mathrm{C}(\mathbf{A})$$

(dove in (1) l'eguaglianza vale perché da $\mathbf{Q}_0$ a $\mathbf{Q}_1$ sono state tolte solo le colonne nulle, che non impattano in una combinazione lineare; in (2) per osservazione 226) e le colonne di Q formano una base ortonormale di $\mathrm{C}(\mathbf{A})$.

Osservazione 228. Si osservi inoltre che essendo $\mathbf{R}_0$ invertibile (quadrata unitriangolare superiore, per cui ciò discende da 1.6.12), le sue colonne sono linearmente indipendenti. Dunque anche le colonne di $\mathbf{R}_1$ sono linearmente indipendenti, per cui $\mathrm{rk}\,\mathbf{R}_1 = k \leq n$. Da $\mathbf{D}$ invertibile ed $\mathbf{R} = \mathbf{DR}_1$ segue infine che $\mathrm{rk}\,\mathbf{A} = k$. Pertanto la decomposizione $\mathbf{A} = \mathbf{QR}$ è una decomposizione a rango pieno di $\mathbf{A}$.

Esempio 3.6.3. Si consideri nuovamente la matrice di esempio 3.6.2

$$\mathbf{A} = \begin{bmatrix} 1 & i & 6 & 8i \\ 2i & -2 & 0 & 8 \\ 1 & i & 6 & 8i \end{bmatrix}$$

Partendo dalla decomposizione QR non normalizzata $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$, si costruiscono prima le matrici $\mathbf{Q}_1$ (da $\mathbf{Q}_0$, senza seconda e quarta colonna) e $\mathbf{R}_1$ (da $\mathbf{R}_0$ senza seconda e quarta riga):

$$\mathbf{Q}_1 = \begin{bmatrix} 1 & 4 \\ 2i & -4i \\ 1 & 4 \end{bmatrix}, \quad \mathbf{R}_1 = \begin{bmatrix} 1 & i & 2 & 0 \\ 0 & 0 & 1 & 2i \end{bmatrix}$$

Poiché

$$\|\mathbf{u}_1\|_2 = \sqrt{\begin{bmatrix} 1 & -2i & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2i \\ 1 \end{bmatrix}} = \sqrt{1+4+1} = \sqrt{6}$$

$$\|\mathbf{u}_3\|_2 = \sqrt{\begin{bmatrix} 4 & 4i & 4 \end{bmatrix} \begin{bmatrix} 4 \\ -4i \\ 4 \end{bmatrix}} = \sqrt{16+16+16} = 4\sqrt{3}$$

allora

$$\mathbf{D} = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & 4\sqrt{3} \end{bmatrix}, \quad \mathbf{D}^{-1} = \begin{bmatrix} 1/\sqrt{6} & 0 \\ 0 & 1/4\sqrt{3} \end{bmatrix}$$

Posto

$$\mathbf{Q} = \mathbf{Q}_1\mathbf{D}^{-1} = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} \\ 2i/\sqrt{6} & -i/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}, \quad \mathbf{R} = \mathbf{D}\mathbf{R}_1 = \begin{bmatrix} \sqrt{6} & \sqrt{6}i & 2\sqrt{6} & 0 \\ 0 & 0 & 4\sqrt{3} & 8\sqrt{3}i \end{bmatrix}$$

si ha che $\mathbf{A} = \mathbf{QR}$ è una decomposizione QR normalizzata di $\mathbf{A}$.

3.6.2.3 Esercizi

Esempio 3.6.4. Trovare una decomposizione QR-normalizzata di

$$\mathbf{A} = \begin{bmatrix} 1+i & i & 1 \\ 1-i & -i & 1 \end{bmatrix}$$

Partiamo dalla QR non normalizzata $\mathbf{A} = \mathbf{Q}_0\mathbf{R}_0$:

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1+i \\ 1-i \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} i \\ -i \end{bmatrix} - \frac{\begin{bmatrix} 1-i \\ 1+i \end{bmatrix}^T \begin{bmatrix} i \\ -i \end{bmatrix}}{\begin{bmatrix} 1-i \\ 1+i \end{bmatrix}^T \begin{bmatrix} 1+i \\ 1-i \end{bmatrix}} \begin{bmatrix} 1+i \\ 1-i \end{bmatrix}$$

$$= \begin{bmatrix} i \\ -i \end{bmatrix} - \frac{(1-i)i - i(1+i)}{1-i^2+1-i^2} \begin{bmatrix} 1+i \\ 1-i \end{bmatrix} = \begin{bmatrix} i \\ -i \end{bmatrix} - \frac{i+1-i+1}{4} \begin{bmatrix} 1+i \\ 1-i \end{bmatrix}$$

$$= \begin{bmatrix} i \\ -i \end{bmatrix} - \begin{bmatrix} (1+i)/2 \\ (1-i)/2 \end{bmatrix} = \begin{bmatrix} (i-1)/2 \\ (-i-1)/2 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 1-i \\ i+i \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \end{bmatrix}}{4} \begin{bmatrix} 1+i \\ 1-i \end{bmatrix} - \frac{\frac{i-1}{2} + \frac{-i-1}{2}}{\begin{bmatrix} (-i-1)/2 \\ (i-1)/2 \end{bmatrix}^T \begin{bmatrix} (i-1)/2 \\ (-i-1)/2 \end{bmatrix}} \begin{bmatrix} (i-1)/2 \\ (-i-1)/2 \end{bmatrix}$$

$$= \ldots = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1+i \\ 1-i \end{bmatrix} + 1 \begin{bmatrix} (i-1)/2 \\ (-i-1)/2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -\frac{1+i}{2} \\ -\frac{1-i}{2} \end{bmatrix} + \begin{bmatrix} (i-1)/2 \\ -(i+1)/2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Quindi

$$\mathbf{Q}_0 = \begin{bmatrix} 1+i & \frac{-1+i}{2} & 0 \\ 1-i & \frac{-1-i}{2} & 0 \end{bmatrix}, \mathbf{R}_0 = \begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

e dunque

$$\mathbf{Q}_1 = \begin{bmatrix} 1+i & \frac{-1+i}{2} \\ 1-i & \frac{-1-i}{2} \end{bmatrix}, \mathbf{R}_1 = \begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & 1 & 1 \end{bmatrix}$$

Ora per calcolare $\mathbf{D}$ determiniamo le norme dei vettori di $\mathbf{Q}_1$

$$\left\| \begin{bmatrix} 1+i \\ i-i \end{bmatrix} \right\| = \sqrt{\begin{bmatrix} 1-i \\ i+i \end{bmatrix}^T \begin{bmatrix} 1+i \\ i-i \end{bmatrix}} = \sqrt{1-i^2+1-i^2} = \sqrt{4} = 2$$

$$\left\| \begin{bmatrix} (-1+i)/2 \\ (-i-1)/2 \end{bmatrix} \right\| = \sqrt{\frac{-1+i}{2}\frac{-1-i}{2} + \frac{-1-i}{2}\frac{-1+i}{2}} = \sqrt{\frac{(-1)^2-i^2}{4} + \frac{1^2-i^2}{4}} = \sqrt{\frac{1+1}{2}} = 1$$

Pertanto

$$\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \mathbf{D}^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$$

e

$$\mathbf{Q} = \mathbf{Q}_1\mathbf{D}^{-1} = \begin{bmatrix} 1+i & \frac{-1+i}{2} \\ 1-i & \frac{-1-i}{2} \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} (1+i)/2 & (-1+1)/2 \\ (1-i)/2 & (-1-i)/2 \end{bmatrix}$$

e

$$\mathbf{R} = \mathbf{D}\mathbf{R}_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & -1/2 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

Esempio 3.6.5. Si trovi una decomposizione QR normalizzata di

$$\mathbf{A} = \begin{bmatrix} 1-i & 1 & 0 \\ 0 & 0 & 1 \\ 1+i & i & 1+i \end{bmatrix}$$

Si ha

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1-i \\ 0 \\ 1+i \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix} - \frac{\begin{bmatrix} 1+i \\ 0 \\ 1-i \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}}{\begin{bmatrix} 1+i \\ 0 \\ 1-i \end{bmatrix}^T \begin{bmatrix} 1-i \\ 0 \\ 1+i \end{bmatrix}} \begin{bmatrix} 1-i \\ 0 \\ 1+i \end{bmatrix}$$

$$= \ldots = \begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix} - \frac{1+i}{2} \begin{bmatrix} 1-i \\ 0 \\ 1+i \end{bmatrix} = \ldots = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1+i \end{bmatrix} - \frac{\begin{bmatrix} 1+i \\ 0 \\ 1-i \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 1+i \end{bmatrix}}{4} \begin{bmatrix} 1-i \\ 0 \\ 1+i \end{bmatrix} - 0\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1+i \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1-i \\ 0 \\ 1+i \end{bmatrix} =$$

Quindi

$$\mathbf{A} = \underbrace{\begin{bmatrix} 1-i & 0 & \frac{1-i}{2} \\ 0 & 0 & 1 \\ 1+i & 0 & \frac{1+i}{2} \end{bmatrix}}_{\mathbf{Q}_0} \underbrace{\begin{bmatrix} 1 & \frac{1+i}{2} & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_0} = \underbrace{\begin{bmatrix} 1-i & \frac{1-i}{2} \\ 0 & 1 \\ 1+i & \frac{1+i}{2} \end{bmatrix}}_{\mathbf{Q}_1} \underbrace{\begin{bmatrix} 1 & \frac{1+i}{2} & 1/2 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_1}$$

Per il calcolo di $\mathbf{D}$ si ha

$$\|\mathbf{u}_1\| = \sqrt{(1-i)(1+i) + (1-i)(1+i)} = \sqrt{1 - i^2 + 1 - i^2} = \sqrt{4} = 2$$

$$\|\mathbf{u}_2\| = \sqrt{\frac{(i-1)(-i-1)}{4} + 1 + \frac{(1+i)(1-i)}{4}} = \sqrt{\frac{-(i^2-1)}{4} + 1 + \frac{1-i^2}{4}} = \sqrt{1/2 + 1 + 1/2} = \sqrt{2}$$

Quindi

$$\mathbf{D} = \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix}, \mathbf{D}^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix}$$

Quindi

$$\mathbf{R} = \mathbf{D}\mathbf{R}_1 = \begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & \frac{1+i}{2} & 1/2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1+i & 1 \\ 0 & 0 & \sqrt{2} \end{bmatrix}$$

$$\mathbf{Q} = \mathbf{Q}_1\mathbf{D}^{-1} = \begin{bmatrix} 1-i & \frac{1-i}{2} \\ 0 & 1 \\ 1+i & \frac{1+i}{2} \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix} = \ldots = \begin{bmatrix} \frac{1-i}{2} & \frac{(-i+1)}{2sqrt2} \\ 0 & 1/\sqrt{2} \\ \frac{1+i}{2} & \frac{1+i}{2\sqrt{2}} \end{bmatrix}$$

Esempio 3.6.6. Si trovi una decomposizione QR-normalizzata di

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 2 \\ -i & 1 & i & -2i \\ 0 & i & 1 & -i \end{bmatrix}$$

Per $\mathbf{Q}_0$ si ha:

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 0 \\ 1 \\ i \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ i \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ i \end{bmatrix}}{\begin{bmatrix} 0 \\ i \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ i \end{bmatrix} - \frac{i}{1} \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ i \end{bmatrix} + \begin{bmatrix} 0 \\ i^2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ i \end{bmatrix}$$

$$\mathbf{u}_3 = v_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ i \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix}}{1} \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 0 \\ -i \end{bmatrix}^T \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 0 \\ -i \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ i \end{bmatrix}} \begin{bmatrix} 0 \\ 0 \\ i \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix} - \frac{-1}{1} \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix} - \frac{-i}{1} \begin{bmatrix} 0 \\ 0 \\ i \end{bmatrix} = \begin{bmatrix} 0 \\ i \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_4 = \mathbf{v}_4 - \alpha_{14}\mathbf{u}_1 - \alpha_{24}\mathbf{u}_2 - \alpha_{34}\mathbf{u}_3$$

$$= \begin{bmatrix} 2 \\ -2i \\ -i \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ i \\ 0 \end{bmatrix}^T \begin{bmatrix} 2 \\ -2i \\ -i \end{bmatrix}}{1} \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 0 \\ -i \end{bmatrix}^T \begin{bmatrix} 2 \\ -2i \\ -i \end{bmatrix}}{1} \begin{bmatrix} 0 \\ 0 \\ i \end{bmatrix} - 0\mathbf{u}_3$$

$$= \begin{bmatrix} 2 \\ -2i \\ -i \end{bmatrix} - \frac{2}{1} \begin{bmatrix} 0 \\ -i \\ 0 \end{bmatrix} - \frac{-1}{1} \begin{bmatrix} 0 \\ 0 \\ i \end{bmatrix} = \begin{bmatrix} 2 \\ -2i \\ -i \end{bmatrix} + \begin{bmatrix} 0 \\ 2i \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ i \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

Si giunge allora a

$$\mathbf{A} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 2 \\ -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{bmatrix}}_{\mathbf{Q}_0} \underbrace{\begin{bmatrix} 1 & i & -1 & 2 \\ 0 & 1 & -i & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_0} = \underbrace{\begin{bmatrix} 0 & 0 & 2 \\ -i & 0 & 0 \\ 0 & i & 0 \end{bmatrix}}_{\mathbf{Q}_1} \underbrace{\begin{bmatrix} 1 & i & -1 & 2 \\ 0 & 1 & -i & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_1}$$

Per le norme necessarie al calcolo di $\mathbf{D}$

$$\|\mathbf{u}_1\| = 1, \quad \|\mathbf{u}_2\| = 1, \quad \|\mathbf{u}_3\| = 2$$

Pertanto

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \mathbf{D}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$$

e

$$\mathbf{R} = \mathbf{D}\mathbf{R}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & i & -1 & 2 \\ 0 & 1 & -i & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & i & -1 & 2 \\ 0 & 1 & -i & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$
$$\mathbf{Q} = \mathbf{Q}_1\mathbf{D}^{-1} = \begin{bmatrix} 0 & 0 & 2 \\ -i & 0 & 0 \\ 0 & i & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ -i & 0 & 0 \\ 0 & i & 0 \end{bmatrix}$$

Esempio 3.6.7. È data la matrice reale $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$:

- trovare basi ortonormali dei 4 sottospazi fondamentali della matrice $\mathbf{A}$
- trovare una decomposizione QR normalizzata di $\mathbf{A}$
- trovare la matrice di proiezione sullo spazio delle colonne $\mathrm{C}(\mathbf{A})$

TODO: Da fare

Esempio 3.6.8. Trovare una decomposizione QR normalizzata della matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 5 & 0 \\ 1 & 2 & -1 & -1 \\ 0 & -1 & 1 & 1 \end{bmatrix}$$

e trovare la matrice di proiezione sullo spazio nullo $\mathrm{N}(\mathbf{A})$.
Per il primo punto

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} - \frac{2}{1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix}}{1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} - \frac{-1}{1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{-1}{1}\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_4 = \mathbf{v}_4 - \alpha_{14}\mathbf{u}_1 - \alpha_{24}\mathbf{u}_2 - \alpha_{34}\mathbf{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}} \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} - 0\begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} - \frac{-1}{1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{-1}{1}\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Da cui

$$\mathbf{A} = \underbrace{\begin{bmatrix} 0 & 0 & 5 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}}_{\mathbf{Q}_0} \underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_0} = \underbrace{\begin{bmatrix} 0 & 0 & 5 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}}_{\mathbf{Q}_1} \underbrace{\begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\mathbf{R}_1}$$

Si ha poi

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}, \mathbf{D}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/5 \end{bmatrix}$$

Quindi

$$\mathbf{R} = \mathbf{D}\mathbf{R}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 5 & 0 \end{bmatrix}$$

$$\mathbf{Q} = \mathbf{Q}_1\mathbf{D}^{-1} = \begin{bmatrix} 0 & 0 & 5 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$$

e facendo la verifica della moltiplicazione $\mathbf{QR}$ si ritrova $\mathbf{A}$.
Per la matrice di proiezione sullo spazio nullo devo avere una base dello spazio nullo. Facciamo la riduzione della matrice

$$\begin{bmatrix} 0 & 0 & 5 & 0 \\ 1 & 2 & -1 & -1 \\ 0 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 & -1 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Si ha dunque il sistema

$$\begin{cases} x_1 + x_2 = 0 \\ x_2 - x_3 - h = 0 \\ x_3 = 0 \\ x_4 = h \end{cases} \quad \begin{cases} x_1 + x_2 = 0 \\ x_2 = h \\ x_3 = 0 \\ x_4 = h \end{cases} \quad \begin{cases} x_1 = -x_2 \\ x_2 = h \\ x_3 = 0 \\ x_4 = h \end{cases} \quad \begin{cases} x_1 = -h \\ x_2 = h \\ x_3 = 0 \\ x_4 = h \end{cases} \quad \rightarrow h \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Dunque

$$\mathrm{N}\left(\mathbf{A}\right) = \left\langle \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle$$

Ora trovata la base, essendo composta da 1 elemento è ortogonale e per la base ortonormale calcoliamo la norma

$$\sqrt{\begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}} = \sqrt{1+1+1} = \sqrt{3}$$

Dunque la base ortonormale è

$$\mathscr{B} = \left\langle \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix} \right\rangle$$

La matrice di proiezione è dunque

$$\mathbf{P} = \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \\ 1/\sqrt{3} \end{bmatrix}^H = \begin{bmatrix} 1/3 & -1/3 & 0 & -1/3 \\ -1/3 & 1/3 & 0 & 1/3 \\ 0 & 0 & 0 & 0 \\ -1/3 & 1/3 & 0 & 1/3 \end{bmatrix}$$

Esempio 3.6.9. Sia V il sottospazio di $\mathbb{C}^3$ dotato del prodotto interno standard generato dai vettori

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

- Si calcoli la matrice di proiezione $\mathbf{P}$ di $\mathbb{C}^3$ su V;
- Si calcoli la proiezione ortogonale di $\mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix}$ su V

Per la matrice di proiezione costruiamo la base ortonormale

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ -i \\ 0 \end{bmatrix}^T \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1-i}{1+i} \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} (1-i)/2 \\ (1+i)/2 \\ 0 \end{bmatrix} = \begin{bmatrix} (1+i)/2 \\ (1-i)/2 \\ 0 \end{bmatrix}$$

I vettori hanno norma

$$\|\mathbf{u}_1\| = \sqrt{1+1} = \sqrt{2}$$

$$\|\mathbf{u}_2\| = \sqrt{\frac{1+i}{2}\frac{1-i}{2} + \frac{1-i}{2}\frac{1+i}{2}} = \sqrt{\frac{(1-i^2)}{2}} = \sqrt{1} = 1$$

La base ortonormale è

$$\begin{bmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1+i}{2} \\ \frac{1-i}{2} \\ 0 \end{bmatrix}$$

La matrice di proiezione è dunque

$$\mathbf{P} = \begin{bmatrix} 1/\sqrt{2} & \frac{1+i}{2} \\ i/\sqrt{2} & \frac{1-i}{2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -i/\sqrt{2} & 0 \\ (1-i)/2 & (1+i)/2 & 0 \end{bmatrix} = \ldots = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

La proiezione dunque risulta

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$$

3.7 Approssimazione ai minimi quadrati

Se un sistema lineare $\mathbf{Ax} = \mathbf{y}$ ha la matrice dei coefficienti $\mathbf{A}$ (quadrata) invertibile, allora il vettore $\mathbf{A}^{-1}\mathbf{y}$ è l'unica soluzione del sistema. Ricordando che, data una generica matrice $m \times n$, la sua pseudo inversa $\mathbf{A}^{+}$ generalizza per certi aspetti la matrice inversa, è naturale porsi la seguente domanda: dato il sistema lineare $\mathbf{Ax} = \mathbf{y}$ con $\mathbf{A}$ matrice qualunque, quale significato ha il vettore $\mathbf{A}^{+}\mathbf{y}$?

3.7.1 Il vettore $\mathbf{A}^{+}\mathbf{y}$

Sia $\mathbf{A}$ una matrice complessa $m \times n$ e sia $\mathbf{y} \in \mathbb{C}^m$. Ricordiamo che il sistema $\mathbf{Ax} = \mathbf{y}$ ha soluzione se e solo se $\mathbf{y} \in \mathrm{C}(\mathbf{A})$. Dal teorema 3.3.12 (punto 2) applicato a $\mathbf{v} = \mathbf{y}$ e $U = \mathrm{C}(\mathbf{A})$ sappiamo che esiste un unico vettore in $\mathrm{C}(\mathbf{A})$ che ha distanza euclidea minima da $\mathbf{y}$: esso è la proiezione ortogonale $\mathbf{y}_0$ di $\mathbf{y}$ su $\mathrm{C}(\mathbf{A})$

$$\left\| \mathbf{y} - \underbrace{P_{\mathrm{C}(\mathbf{A})}(\mathbf{y})}_{\mathbf{y}_0} \right\| < \|\mathbf{y} - \mathbf{z}\|, \quad \mathbf{z} \in \mathrm{C}(\mathbf{A}) : \mathbf{z} \neq \mathbf{y}_0$$

Sappiamo altresì che $\mathbf{y} = \mathbf{y}_0$ se e solo se $\mathbf{y} \in \mathrm{C}(\mathbf{A})$.

Definizione 3.7.1 (Sistema compatibile)**.** Dato un sistema $\mathbf{Ax} = \mathbf{y}$ è detto così il sistema

$$\mathbf{Ax} = \mathbf{y}_0 \tag{3.39}$$

dove si è sostituita la proiezione $\mathbf{y}_0$ di $\mathbf{y}$ su $\mathrm{C}(\mathbf{A})$ al posto di $\mathbf{y}$ stesso.

Definizione 3.7.2 (Sistema delle equazioni normali)**.** Dato un sistema $\mathbf{Ax} = \mathbf{y}$, è detto così il sistema

$$\mathbf{A}^H\mathbf{Ax} = \mathbf{A}^H\mathbf{y} \tag{3.40}$$

Osservazione 229. Andiamo a caratterizzare le soluzioni del sistema compatibile.

Teorema 3.7.1. *Dato il sistema lineare* $\mathbf{Ax} = \mathbf{y}$ *con* $\mathbf{A}$ *matrice* $m \times n$ *e detto* $\mathbf{y}_0$ *il vettore di proiezione di* $\mathbf{y}$ *su* $\mathrm{C}(\mathbf{A})$*, le seguenti condizioni per un vettore* $\mathbf{v} \in \mathbb{C}^n$ *sono equivalenti:*

1. $\mathbf{v}$ *è soluzione del sistema compatibile* $\mathbf{Ax} = \mathbf{y}_0$*, ossia* $\mathbf{Av} = \mathbf{y}_0$

2. *la distanza di* $\mathbf{Ax}$ *da* $\mathbf{y}$ *è minima quando* $\mathbf{x} = \mathbf{v}$*:*

$$\|\mathbf{Av} - \mathbf{y}\|_2 < \|\mathbf{Au} - \mathbf{y}\|_2, \quad \forall \mathbf{u} \in \mathbb{C}^n : \mathbf{u} \neq \mathbf{v}$$

3. *il sistema delle equazioni normali è risolubile e ha esattamente le soluzioni del sistema compatibile:*

$$\mathbf{A}^H\mathbf{Av} = \mathbf{A}^H\mathbf{y}$$

4. $\mathbf{Rv} = \mathbf{Q}^H\mathbf{y}$*, con* $\mathbf{Q}$ *ed* $\mathbf{R}$ *prese dalla decomposizione QR normalizzata* $\mathbf{A} = \mathbf{QR}$

Dimostrazione. La:

- (1) $\iff$ (2) discende dalla caratterizzazione della proiezione $\mathbf{y}_0$ di $\mathbf{y}$ su $\mathrm{C}(\mathbf{A})$ (vedi 3.3.12, punto 2) e dal fatto che un vettore appartiene a $\mathrm{C}(\mathbf{A})$ se e solo se è del tipo $\mathbf{Au}$ per un opportuno vettore $\mathbf{u} \in \mathbb{C}^n$: $\mathbf{y}_0$ appartiene a $\mathrm{C}(\mathbf{A})$ perché è proiezione di $\mathbf{y}$ su $\mathrm{C}(\mathbf{A})$. Dato questo si ha che (per 3.3.12)
$$\|\mathbf{y} - \mathbf{y}_0\| < \|\mathbf{y} - \mathbf{Au}\|$$
con $\mathbf{Au}$ come lo $\mathbf{z}$ del teorema. Per arrivare alla formula pongo $-$ sotto norma (i moduli non cambiano, solo la direzione) per cui:
$$\|\mathbf{y}_0 - \mathbf{y}\| < \|\mathbf{Au} - \mathbf{y}\|$$
da cui
$$\|\mathbf{Av} - \mathbf{y}\| < \|\mathbf{Au} - \mathbf{y}\|$$

- (3) $\iff$ (1) ricordando che $\mathrm{C}(\mathbf{A})^\perp = \mathrm{N}\left(\mathbf{A}^H\right)$ (per 3.3.4) si ha
$$\begin{aligned} \mathbf{A}^H\mathbf{Av} = \mathbf{A}^H\mathbf{y} &\iff \mathbf{A}^H\mathbf{Av} - \mathbf{A}^H\mathbf{y} = 0 \\ &\iff \mathbf{A}^H(\mathbf{Av} - \mathbf{y}) = \mathbf{0} \\ &\iff \mathbf{Av} - \mathbf{y} \in \mathrm{N}\left(\mathbf{A}^H\right) = \mathrm{C}(\mathbf{A})^\perp \end{aligned}$$
Inoltre poiché $\mathbf{y} = \mathbf{Av} + (\mathbf{y} - \mathbf{Av})$ e $\mathbf{Av} \in \mathrm{C}(\mathbf{A})$, dalla proposizione 3.3.7 segue che
$$\mathbf{Av} - \mathbf{y} \in \mathrm{C}(\mathbf{A})^\perp \iff \mathbf{Av} = \mathbf{y}_0$$
Tentativo di spiegazione di quest'ultimo passaggio: $\mathbf{y}_0 = P_{\mathrm{C}(\mathbf{A})}(\mathbf{y})$ quindi dato che $\mathbf{y}_0 \in \mathrm{C}(\mathbf{A})$, $\mathbf{y} - \mathbf{y}_0 \in \mathrm{C}(\mathbf{A})^\perp$, idem per $-(\mathbf{y} - \mathbf{y}_0) \in \mathrm{C}(\mathbf{A})^\perp$, ossia $\mathbf{y}_0 - \mathbf{y} \in \mathrm{C}(\mathbf{A})^\perp$ e quindi appunto si conclude che $\mathbf{Av} - \mathbf{y} \in \mathrm{C}(\mathbf{A})^\perp$ se e solo se $\mathbf{Av} = \mathbf{y}_0$.
In conclusione si ottiene che $\mathbf{A}^H\mathbf{Av} = \mathbf{A}^H\mathbf{y}$ se e solo se $\mathbf{Av} = \mathbf{y}_0$

- (1) $\iff$ (4): data una decomposizione QR normalizzata $\mathbf{A} = \mathbf{QR}$ abbiamo che
$$\begin{aligned} &\mathbf{Av} = \mathbf{y} \\ &\mathbf{QRv} = \mathbf{y} \end{aligned}$$
Ma dato che $\mathbf{Q}$ è composta da colonne ortonormali, sappiamo per il lemma 3.6.2 che $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$. Pre-moltiplicando allora entrambi i membri del punto in cui siamo giunti per $\mathbf{Q}^H$ ed effettuando le sostituzioni si conclude:
$$\begin{aligned} &\mathbf{Q}^H\mathbf{QRv} = \mathbf{Q}^H\mathbf{y} \\ &\mathbf{IRv} = \mathbf{Q}^H\mathbf{y} \\ &\mathbf{Rv} = \mathbf{Q}^H\mathbf{y} \end{aligned}$$

□

Osservazione 230. Possiamo ora dare la risposta alla domanda posta all'inizio di questo paragrafo.

Teorema 3.7.2. *Dato il sistema lineare* $\mathbf{Ax} = \mathbf{y}$ *e detto* $\mathbf{y}_0$ *il vettore proiezione di* $\mathbf{y}$ *su* $\mathrm{C}(\mathbf{A})$ *il vettore* $\mathbf{A}^+\mathbf{y}$ *è l'unica soluzione del sistema* $\mathbf{Ax} = \mathbf{y}_0$ *di norma euclidea minima.*

Dimostrazione. Ordinatamente:

- proviamo innanzitutto che il vettore $\mathbf{A}^+\mathbf{y}$ è soluzione del sistema $\mathbf{Ax} = \mathbf{y}_0$. Per il teorema 3.7.1 (punto 3) ciò accade se e solo se $\mathbf{A}^+\mathbf{y}$ soddisfa il sistema delle equazioni normali, ossia se è vera

$$\mathbf{A}^H\mathbf{A}\mathbf{A}^+\mathbf{y} = \mathbf{A}^H\mathbf{y}$$

  Ma per la proprietà 5 delle pseudo-inverse (1.6.21) sappiamo che

$$\mathbf{A}^H\mathbf{A}\mathbf{A}^+ = \mathbf{A}^H$$

  e quindi effettuando la sostituzione nel primo membro dell'equazione precedente effettivamente l'uguaglianza è provata e si conclude.

- proviamo ora che, se $\mathbf{u} \in \mathbb{C}^n$ è un'altra soluzione di $\mathbf{Ax} = \mathbf{y}_0$ diversa da $\mathbf{A}^+\mathbf{y}$, allora $\|\mathbf{u}\|_2 > \|\mathbf{A}^+\mathbf{y}\|_2$ (ossia $\mathbf{A}^+\mathbf{y}$ è di norma euclidea minima).
  Osserviamo preliminarmente che $\mathbf{A}^+\mathbf{y} \in \mathrm{N}(\mathbf{A})^\perp$. Infatti, se $\mathbf{w} \in \mathrm{N}(\mathbf{A})$, da cui $\mathbf{Aw} = \mathbf{0}$, risulta

$$\begin{aligned}(\mathbf{w}|\mathbf{A}^+\mathbf{y}) = \mathbf{w}^H\mathbf{A}^+\mathbf{y} &\overset{(1)}{=} \mathbf{w}^H\mathbf{A}^+\mathbf{A}\mathbf{A}^+\mathbf{y} \overset{(2)}{=} \mathbf{w}^H(\mathbf{A}^+\mathbf{A})^H\mathbf{A}^+\mathbf{y} \\ &\overset{(3)}{=} \mathbf{w}^H\mathbf{A}^H(\mathbf{A}^+)^H\mathbf{A}^+\mathbf{y} \overset{(4)}{=} (\underbrace{\mathbf{Aw}}_{=\mathbf{0}})^H(\mathbf{A}^+)^H\mathbf{A}^+\mathbf{y} \\ &= 0\end{aligned}$$

  dove in (1) si applica la seconda proprietà della definizione della pseudo-inversa, in (2) la quarta proprietà, e in (3) e (4) le proprietà dell'H-trasposizione. Pertanto dato che $(\mathbf{w}|\mathbf{A}^+\mathbf{y}) = 0$ allora $\mathbf{w} \perp \mathbf{A}^+\mathbf{y}$ e dunque $\mathbf{A}^+\mathbf{y} \in \mathrm{N}(\mathbf{A})^\perp$ (dato che $\mathbf{w} \in \mathrm{N}(\mathbf{A})$) .
  Osserviamo poi che, essendo $\mathbf{A}^+\mathbf{y}$ una soluzione del sistema $\mathbf{Ax} = \mathbf{y}_0$, ogni altra soluzione $\mathbf{u}$ è del tipo $\mathbf{u} = \mathbf{A}^+\mathbf{y} + \mathbf{n}$ (per considerazioni in osservazione 166) dove $\mathbf{n} \in \mathrm{N}(\mathbf{A})$. Ora se $\mathbf{u} \neq \mathbf{A}^+\mathbf{y}$, risulta $\mathbf{n} \neq \mathbf{0}$. Poiché $\mathbf{n} \in \mathrm{N}(\mathbf{A})$ e $\mathbf{A}^+\mathbf{y} \in \mathrm{N}(\mathbf{A})^\perp$ si ha

$$\begin{aligned}\|\mathbf{u}\|_2^2 &= \left\|\mathbf{A}^+\mathbf{y} + \mathbf{n}\right\|_2^2 = (\mathbf{A}^+\mathbf{y} + \mathbf{n})^H(\mathbf{A}^+\mathbf{y} + \mathbf{n}) \\ &= \left[(\mathbf{A}^+\mathbf{y})^H + \mathbf{n}^H\right](\mathbf{A}^+\mathbf{y} + \mathbf{n}) \\ &= (\mathbf{A}^+\mathbf{y})^H(\mathbf{A}^+\mathbf{y}) + \underbrace{(\mathbf{A}^+\mathbf{y})^H\mathbf{n}}_{=0(*)} + \underbrace{\mathbf{n}^H(\mathbf{A}^+\mathbf{y})}_{=0(*)} + \mathbf{n}^H\mathbf{n} \\ &= \left\|\mathbf{A}^+\mathbf{y}\right\|_2^2 + \|\mathbf{n}\|_2^2 > \left\|\mathbf{A}^+\mathbf{y}\right\|_2^2\end{aligned}$$

  dove i termini $(*)$ scompaiono poiché si tratta del prodotto interno di vettori tra loro perpendicolari.

□

Osservazione 231. In virtù di quanto provato nel teorema 3.7.2 e poiché la norma euclidea di un vettore si ottiene dalla somma dei quadrati dei moduli delle sue coordinate (estraendone la radice quadrata), il vettore $\mathbf{A}^+\mathbf{y}$ si chiama *soluzione ai minimi quadrati* del sistema $\mathbf{Ax} = \mathbf{y}$.

Esempio 3.7.1. Si consideri la matrice dell'esempio 1.6.4 del capitolo 1

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}$$

che ha come pseudo-inversa la matrice $\mathbf{A}^+ = \frac{1}{70}\mathbf{A}^T$. Se $\mathbf{y} = [1\ 1]^T$, il sistema $\mathbf{Ax} = \mathbf{y}$ non ha soluzioni perché:

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \notin \mathrm{C}(\mathbf{A}) = \left\langle \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\rangle$$

Il vettore proiezione di $[1\,1]^T$ su $\mathrm{C}(\mathbf{A})$ è il vettore

$$\mathbf{y}_0 = \frac{\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}}{\begin{bmatrix} 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6/5 \\ 3/5 \end{bmatrix}$$

La soluzione ai minimi quadrati del sistema $\mathbf{Ax} = \mathbf{y}$ è il vettore

$$\mathbf{A}^+\mathbf{y} = \frac{1}{70}\mathbf{A}^T \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{70} \begin{bmatrix} 6 \\ 0 \\ 3 \\ 9 \end{bmatrix} = \frac{3}{70} \begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix}$$

la cui norma euclidea

$$\left\| \mathbf{A}^+\mathbf{y} \right\|_2 = \frac{3}{70}\sqrt{14}$$

è la minima tra le norme euclidee delle soluzioni del sistema $\mathbf{Ax} = \mathbf{y}_0$.

Osservazione 232. Data la matrice $m \times n$ $\mathbf{A}$, qualora la sua pseudo inversa $\mathbf{A}^+$ sia nota, è facile calcolare la proiezione di un vettore di $\mathbb{C}^m$ sullo spazio delle colonne $\mathrm{C}(\mathbf{A})$ e quella di un vettore di $\mathbb{C}^n$ sullo spazio delle righe $\mathrm{C}\left(\mathbf{A}^H\right)$ di $\mathbf{A}$

Corollario 3.7.3. *Data la matrice $m \times n$ $\mathbf{A}$ si ha che:*

1. *la proiezione di $\mathbf{y} \in \mathbb{C}^m$ su $\mathrm{C}(\mathbf{A})$ è il vettore $\mathbf{AA}^+\mathbf{y}$*
2. *la proiezione di $\mathbf{v} \in \mathbb{C}^n$ su $\mathrm{C}\left(\mathbf{A}^H\right)$ è il vettore $\mathbf{A}^+\mathbf{Av}$*

Dimostrazione. Rispettivamente:

1. sia $\mathbf{y}_0$ la proiezione di $\mathbf{y}$ su $\mathrm{C}(\mathbf{A})$; dal teorema 3.7.2 sappiamo che $\mathbf{A}^+\mathbf{y}$ è soluzione di $\mathbf{Ax} = \mathbf{y}_0$, dunque che vale l'uguaglianza $\mathbf{AA}^+\mathbf{y} = \mathbf{y}_0$. Ma quindi per ottenere la proiezione $\mathbf{y}_0$ basta pre-moltiplicare $\mathbf{y}$ per $\mathbf{AA}^+$ (che assume il ruolo di matrice di proiezione?)

2. per quanto mostrato al punto (1) la proiezione $\mathbf{v}_0$ di $\mathbf{v}$ su $\mathrm{C}\left(\mathbf{A}^H\right)$ è $\mathbf{v}_0 = \mathbf{A}^H(\mathbf{A}^H)^+\mathbf{v}$ (per mera sostituzione). Ma risulta:

$$\mathbf{v}_0 = \mathbf{A}^H(\mathbf{A}^H)^+\mathbf{v} \overset{(1)}{=} \mathbf{A}^H(\mathbf{A}^+)^H\mathbf{v} \overset{(2)}{=} (\mathbf{A}^+\mathbf{A})^H\mathbf{v} \overset{(3)}{=} \mathbf{A}^+\mathbf{A}\mathbf{v}$$

con (1) per la seconda proprietà della definizione delle matrici pseudo-inverse, in (3) la quarta e in (2) proprietà dell'H-trasposizione (dunque qui la matrice $\mathbf{A}^+\mathbf{A}$ diviene matrice di proiezione?)

□

3.7.2 Metodo dei minimi quadrati

La seconda parte è dedicata ad illustrare il metodo dei minimi quadrati, che è il modo più usato per approssimare dati sperimentali. Applicheremo questo metodo a un fenomeno descritto da una equazione polinomiale del tipo

$$Y = a_0 + a_1X + a_2X^2 + \ldots + a_nX^n \tag{3.41}$$

dove sono da determinarsi i coefficienti a_i dell'equazione: effettuando N misurazioni del fenomeno si ottengono N valori distinti (ipotizziamo) $x_1, \ldots, x_N$ della variabile indipendente X ed altrettanti valori $y_1, \ldots, y_N$ della variabile dipendente Y:

- se $N \leq n$ (sistema indeterminato), esistono infiniti polinomi $a_0 + a_1X + \ldots + a_nX^n$ che assumono negli N punti $x_1, \ldots, x_N$ rispettivamente i valori $y_1, \ldots, y_N$ (si pensi alle infinite rette per un punto, o alle infinite parabole per due punti)
- se $N = n + 1$ (sistema determinato) esiste esattamente un polinomio soddisfacente tali requisiti (si pensi alla retta per due punti, o alla parabola per tre punto; vedi esercizio 3.53).

TODO: fixme

- se $N > n + 1$, ossia nel caso usuale in cui si effettuano molte più misurazioni rispetto al grado dell'equazione polinomiale, non esiste in generale il polinomio cercato (sistema impossibile). Ci si chiede allora: quale è il polinomio che meglio approssima i dati ottenuti?

Una risposta convincente alla domanda di quest'ultimo punto è che il polinomio $a_0 + a_1X + a_2X^2 + \ldots + a_nX^n$ che meglio approssima i dati ottenuti è quello che minimizza la quantità

$$\sum_{i=1}^{N} \left(y_i - (a_0 + a_1x_i + a_2x_i^2 + \ldots + a_nx_i^n)\right)^2 \tag{3.42}$$

cioè la somma dei quadrati degli errori commessi nelle misurazioni rispetto ai valori ipotizzati come reali. Posto $\mathbf{y}$ vettore dei dati misurati, $\mathbf{x}$ vettore incognito e $\mathbf{A}$ matrice $N \times (n + 1)$

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_N & x_N^2 & \ldots & x_N^n \end{bmatrix}$$

la quantità da minimizzare in 3.42 riscritta in termini matriciali corrisponde a

$$\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2$$

Ossia dobbiamo cercare la soluzione $\mathbf{v}$ che sostituita ad $\mathbf{x}$ minimizzi quest'ultima). Per il teorema 3.7.1 (punto 2[1]) i vettori $\mathbf{v}$ per cui la norma $\|\mathbf{y} - \mathbf{A}\mathbf{v}\|_2$ è minima sono tutte e sole le soluzioni del sistema delle equazioni normali $\mathbf{A}^H\mathbf{A}\mathbf{x} = \mathbf{A}^H\mathbf{y}$. Ci serve ora sapere che:

Lemma 3.7.4. *Nel caso in esame ($N \geq n+1$) la matrice $\mathbf{A}^H\mathbf{A}$ ha rango $n+1$ (ed è invertibile).*

Dimostrazione. Per provarlo è sufficiente dimostrare che la matrice quadrata seguente ($\mathbf{V}$ è la parte in alto di $\mathbf{A}$), di ordine $n+1$, è invertibile

$$\mathbf{V} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n+1} & x_{n+1}^2 & \dots & x_{n+1}^n \end{bmatrix}$$

Se invertibile, infatti, le colonne della matrice $\mathbf{A}$ risultano linearmente indipendenti, il che comporta che $\operatorname{rk}\mathbf{A} = n+1$. Dunque dato che $\mathbf{A}$ ha rango pari al numero di colonne, per la caratterizzazione del teorema 1.6.10 la matrice $\mathbf{A}^H\mathbf{A}$ è invertibile, ed essendo quest'ultima quadrata di ordine $n+1$ (ed invertibile) si conclude che $\operatorname{rk}\mathbf{A}^H\mathbf{A} = n+1$.
Tutto questo, come detto, se la matrice $\mathbf{V}$ (detta matrice di Vandermonde ottenuta dagli $n+1$ valori *distinti* $x_1, x_2, \dots, x_{n+1}$) è effettivamente invertibile: come si vedrà nel capitolo sul determinante , ciò è verificato. □

TODO: Negli esercizi sul libro, a fine capitolo; da approfondire

Proposizione 3.7.5. *Nel caso in esame ($N \geq n+1$), $\mathbf{A}^+\mathbf{y}$ è l'unica soluzione del sistema $\mathbf{A}\mathbf{x} = \mathbf{y}_0$ (con $\mathbf{y}_0$ proiezione di $\mathbf{y}$ su $\mathrm{C}(\mathbf{A})$); pertanto è la soluzione che minimizza $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2$*

Dimostrazione. Considerando che:

- come detto $\mathbf{A}^H\mathbf{A}$ è quadrata di ordine $(n+1)$ e ha rango $n+1$; è dunque invertibile (per la caratterizzazione delle matrici quadrate), con inversa $(\mathbf{A}^H\mathbf{A})^{-1}$;
- $\mathbf{A}$ ha rango $n+1$ (pari al numero delle colonne) e dunque ammette inversa sinistra per il teorema 1.6.10; alla luce della proposizione 1.6.19, essendo non quadrata (è verticale dato che con $N > n+1$) e con inversa sinistra ammette pseudo-inversa pari a

$$\mathbf{A}^+ = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$$

pre-moltiplicando entrambi i membri di $\mathbf{A}^H\mathbf{A}\mathbf{x} = \mathbf{A}^H\mathbf{y}$ per $(\mathbf{A}^H\mathbf{A})^{-1}$ ed effettuando le sostituzioni del caso si ha:

TODO: più o meno ci siamo con questo tentativo di chiarificazione

$$(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{A}\mathbf{x} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H\mathbf{y}$$
$$\mathbf{x} = \mathbf{A}^+\mathbf{y}$$

[1]Nel teorema vi è un segno meno sotto, ma il modulo/norma rimane la stessa.

si ha che $\mathbf{A}^{+}\mathbf{y}$ è unica soluzione del sistema delle equazioni normali e dunque, per il teorema 3.7.1, del sistema compatibile $\mathbf{Ax} = \mathbf{y}_0$ (dove $\mathbf{y}_0$ è la proiezione di $\mathbf{y}$ su $\mathrm{C}(\mathbf{A})$); pertanto è l'unico vettore che minimizza $\|\mathbf{y} - \mathbf{Ax}\|_2$. $\square$

Esempio 3.7.2. Si vuole determinare la retta di equazione $Y = a + bX$ che meglio approssima i quattro punti (X, Y) di $\mathbb{R}^2$:

$$P_1 = (1,1) \quad P_2 = (3,2) \quad P_3 = (0,1) \quad P_4 = (2,3)$$

In termini matriciali possiamo riscrivere il tutto come $\mathbf{y} = \mathbf{Ax}$ con

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix}$$

Vogliamo risolvere il sistema, determinando la soluzione che sostituita $\mathbf{x}$ fornisce una migliore approssimazione. Dato che $\mathbf{A}$ è una matrice reale con $N = 4 > n + 1 = 2$ la soluzione è il vettore

$$\mathbf{A}^{+}\mathbf{y} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}$$

Si ha che

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}$$

Dunque essendo una matrice 2×2 l'inversa è immediatamente determinata come:

$$(\mathbf{A}^T\mathbf{A})^{-1} = \frac{1}{4 \cdot 14 - 6 \cdot 6} \begin{bmatrix} 14 & -6 \\ -6 & 4 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 14 & -6 \\ -6 & 4 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 7 & -3 \\ -3 & 2 \end{bmatrix}$$

Pertanto

$$\mathbf{A}^{+} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T = \frac{1}{10} \begin{bmatrix} 7 & -3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 0 & 2 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 4 & -2 & 7 & 1 \\ -1 & 3 & -3 & 1 \end{bmatrix}$$

e in conclusione

$$\mathbf{A}^{+}\mathbf{y} = \frac{1}{10} \begin{bmatrix} 4 & -2 & 7 & 1 \\ -1 & 3 & -3 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \\ 3 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 4 - 4 + 7 + 3 \\ -1 + 6 - 3 + 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$$

Dunque la retta cercata ha pertanto equazione $Y = 1 + \frac{1}{2}X$. Volendo verificare con R applicando la formula $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}$:

```
A <- rmatrix(c(1,1,
               1,3,
               1,0,
               1,2), ncol = 2)
y <- c(1,2,1,3)
(beta <- solve(t(A) %*% A) %*% t(A) %*% y)

##      [,1]
## [1,]  1.0
## [2,]  0.5
```

Esempio 3.7.3. Dato il sistema di equazioni lineari $\mathbf{Ax} = \mathbf{b}$, dove

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 3 & 4 \\ 2 & 5 & 3 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$$

- Si trovi una decomposizione QR-normalizzata di $\mathbf{A}$
- Si trovi la proiezione del vettore $\mathbf{b}$ sullo spazio delle colonne $\mathrm{C}(\mathbf{A})$ di $\mathbf{A}$, e la si denoti con $\mathbf{b}_0$
- Si mostri che l'insieme delle soluzioni del sistema $\mathbf{Ax} = \mathbf{b}_0$ coincide con l'insieme delle soluzioni del sistema $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$

Per il primo punto

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}^T \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} - \frac{15}{6} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 - 15/6 \\ 3 - 15/6 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}^T \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}}{6} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \frac{\begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix}^T \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}}{\begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix}^T \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix}} \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix} - \frac{9}{6} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - 5 \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix} = \ldots = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Quindi

$$\mathbf{A} = \underbrace{\begin{bmatrix} 1 & -1/2 & 0 \\ 1 & 1/2 & 0 \\ 2 & 0 & 0 \end{bmatrix}}_{\mathbf{Q}_0} \underbrace{\begin{bmatrix} 1 & 15/6 & 3/2 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{R}_0} = \underbrace{\begin{bmatrix} 1 & -1/2 \\ 1 & 1/2 \\ 2 & 0 \end{bmatrix}}_{\mathbf{Q}_1} \underbrace{\begin{bmatrix} 1 & 15/6 & 3/2 \\ 0 & 1 & 5 \end{bmatrix}}_{\mathbf{R}_1}$$

Si ha che

$$\mathbf{D} = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{2}/2 \end{bmatrix}, \mathbf{D}^{-1} = \begin{bmatrix} 1/\sqrt{6} & 0 \\ 0 & \sqrt{2} \end{bmatrix}$$

e

$$\mathbf{R} = \mathbf{DR}_1 = \begin{bmatrix} \sqrt{6} & 0 \\ 0 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} 1 & 15/6 & 3/2 \\ 0 & 1 & 5 \end{bmatrix} = \begin{bmatrix} \sqrt{6} & 15/\sqrt{6} & 3\sqrt{6}/2 \\ 0 & \sqrt{2}/2 & 5\sqrt{2}/2 \end{bmatrix}$$

$$\mathbf{Q} = \mathbf{Q}_1\mathbf{D}^{-1} = \begin{bmatrix} 1 & -1/2 \\ 1 & 1/2 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 0 \\ 0 & \sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} & -\sqrt{2}/2 \\ 1/\sqrt{6} & \sqrt{2}/2 \\ 2/\sqrt{6} & 0 \end{bmatrix}$$

Si può verificare mediante prodotto che la scomposizione $\mathbf{A} = \mathbf{QR}$ è corretta (ossia moltiplicando $\mathbf{Q}$ per $\mathbf{R}$ si ottiene effettivamente $\mathbf{A}$).
Per il secondo punto debbo costruire la matrice di proiezione su $\mathrm{C}(\mathbf{A})$ e per farlo parto dal determinarne una base, mediante la riduzione

$$\begin{bmatrix} 1 & 2 & -1 \\ 1 & 3 & 4 \\ 2 & 5 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 1 & 3 & 4 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \end{bmatrix}$$

Quindi le prime due colonne formano una base; ora applicando Gram Schmidt per ottenere una base ortogonale si giunge a quanto già calcolato

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1/2 & 1/2 & 0 \end{bmatrix}$$

e le rispettive norme sono

$$\|\mathbf{u}_1\| = \sqrt{1^2 + 1^2 + 2^2} = \sqrt{6}$$
$$\|\mathbf{u}_2\| = \sqrt{\frac{1}{4} + \frac{1}{4}} = \sqrt{\frac{1}{2}} = \frac{1}{\sqrt{2}}$$

e la base ortonormale è dunque

$$\mathscr{B} = \left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \sqrt{2} \begin{bmatrix} -1/2 \\ 1/2 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} \begin{bmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \\ 0 \end{bmatrix} \right\}$$

La matrice di proiezione è dunque

$$\mathbf{P} = \mathbf{QQ}^H = \begin{bmatrix} 1/\sqrt{6} & -\sqrt{2}/2 \\ 1/\sqrt{6} & \sqrt{2}/2 \\ 2/\sqrt{6} & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \end{bmatrix} = \ldots = \begin{bmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{bmatrix}$$

La proiezione $\mathbf{b}_0$ è

$$\mathbf{b}_0 = \begin{bmatrix} 2/3 & -1/3 & 1/3 \\ -1/3 & 2/3 & 1/3 \\ 1/3 & 1/3 & 2/3 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/3 \\ -1/3 \\ 4/3 \end{bmatrix}$$

Il terzo punto se/quando ne si avrà voglia.

Capitolo 4

Il determinante

Contents

4.1 Funzioni determinanti

Osservazione 233. Ogni matrice $\mathbf{A} \in M_2(\mathbb{R})$ può essere considerata come una trasformazione f del piano in se: identificando un punto come un vettore colonna $P = [a\, b]^T \in \mathbb{R}^2$ a esso possiamo associare il punto

$$f(\mathbf{P}) = \mathbf{A} \begin{bmatrix} a \\ b \end{bmatrix}$$

La composizione di due trasformazioni di questo tipo corrisponde al prodotto delle matrici associate. Alcuni esempi sulle matrici elementari seguono.

Esempio 4.1.1. La matrice

$$\mathbf{A} = \mathbf{E}_{12} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

attraverso la pre-moltiplicazione implementa la trasformazione

$$f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} b \\ a \end{bmatrix}$$

che coincide con la riflessione attorno alla bisettrice del primo e terzo quadrante (inverte di x e y).

Geometricamente in una trasformazione del genere, i percorsi cambiano senso di percorrenza: se immaginiamo di percorrere un triangolo in senso antiorario, il triangolo trasformato verrò percorso in senso orario.

Esempio 4.1.2. Se
$$\mathbf{A} = \mathbf{E}_1(\mathbf{c}) = \begin{bmatrix} c & 0 \\ 0 & 1 \end{bmatrix}$$
avremo
$$f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} ca \\ b \end{bmatrix}$$
Geometricamente l'area del triangolo trasformato (nei propri vertici), è c volte quella di partenza prima: ad esempio il triangolo di vertici $[0\,0]^T$, $[1\,0]^T$, $[1\,1]^T$ ha area $1/2$, mentre quello con i vertici trasformati in $[0\,0]^T$, $[c\,0]^T$, $[c\,1]^T$ ha area $c/2$.

Esempio 4.1.3. Nel caso
$$\mathbf{A} = \mathbf{E}_{21}(d) = \begin{bmatrix} 1 & 0 \\ d & 1 \end{bmatrix}$$
si ha
$$f\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a \\ ad + b \end{bmatrix}$$
il medesimo triangolo di prima si trasforma in quello di vertici $[0\,0]^T$, $[1\,d]^T$, $[1\,d+1]^T$ che ha ancora area $1/2$.

Osservazione 234. Alcune considerazioni geometriche sulla matrice di trasformazione che stanno alla base della definizione e idea di determinante:

- le $\mathbf{E}_i(c)$ moltiplicano le aree (o i volumi, nello spazio) per c;
- le $\mathbf{E}_{ij}(d)$ non modificano le aree (o i volumi);
- le trasformazioni $\mathbf{E}_{ij}$, cambiano orientazione/rendono speculari;
- nelle trasformazioni implementate da matrici non invertibili, l'immagine dell'applicazione in uno spazio è tutta contenuta in un piano[1]: in tal caso i volumi si annullano.

Definizione 4.1.1 (Funzione determinante)**.** Un'applicazione $\varphi : M_n(\mathbb{C}) \to \mathbb{C}$ si dice *una* funzione determinante se, per ogni $\mathbf{A} \in M_n(\mathbb{C})$ ha le seguenti proprietà:

$$\varphi(\mathbf{A}\mathbf{E}_i(c)) = c\varphi(\mathbf{A}) \tag{4.1}$$
$$\varphi(\mathbf{A}\mathbf{E}_{ij}(d)) = \varphi(\mathbf{A}) \tag{4.2}$$
$$\varphi(\mathbf{A}\mathbf{E}_{ij}) = -\varphi(\mathbf{A}) \tag{4.3}$$

Osservazione 235. Sebbene vi possano essere più di una funzione determinante. dimostreremo che una tale applicazione è *essenzialmente unica.*

[1] O addirittura in una retta e, se la matrice $\mathbf{A}$ è nulla, in un punto

Osservazione 236. L'idea della definizione è che il comportamento della funzione φ è "buono" rispetto alle trasformazioni sulle colonne.
Notiamo che avremmo potuto analogamente considerare la pre-moltiplicazione per matrici elementari, esaminando il comportamento di φ rispetto alle trasformazioni elementari sulle righe. Vedremo comunque che $\varphi(\mathbf{A}) = \varphi(\mathbf{A}^T)$

Lemma 4.1.1 (Determinante nullo 1)**.** *Se $\varphi : M_n(\mathbb{C}) \to \mathbb{C}$ è una funzione determinante e $\mathbf{A} \in M_n(\mathbb{C})$ ha una colonna nulla, allora $\varphi(\mathbf{A}) = 0$.*

Dimostrazione. Supponiamo che la colonna nulla sia la i-esima. Allora $\mathbf{A} = \mathbf{A}\mathbf{E}_i(2)$ (la matrice coincide con quella avente la i-esima colonna moltiplicata per 2) e per la prima proprietà del determinante $\varphi(\mathbf{A}) = \varphi(\mathbf{A}\mathbf{E}_i(2)) = 2\varphi(\mathbf{A})$, da cui la tesi (ossia dato che $\varphi(\mathbf{A})$ è una costante c, $c = 2c \iff c = 0$). □

Lemma 4.1.2 (Determinante nullo 2)**.** *Se $\varphi : M_n(\mathbb{C}) \to \mathbb{C}$ è una funzione determinante e $\mathbf{A} \in M_n(\mathbb{C})$ non è invertibile, allora $\varphi(\mathbf{A}) = 0$.*

Dimostrazione. Siccome $\mathbf{A}$ non è invertibile (ed è quadrata) una colonna è combinazione lineare della altre. Non è restrittivo supporre che sia la prima; al massimo si può utilizzare la post-moltiplicazione per una $\mathbf{E}_{1j}$ opportuna per far in modo che sia così; per la terza proprietà del determinante questo può cambiare di segno il determinante, ma dovendo provare che $\varphi(\mathbf{A}) = 0$ ciò non ha rilevanza. Abbiamo allora, dato $\mathbf{A} = [\mathbf{a}_1\, \mathbf{a}_2\, \dots\, \mathbf{a}_n]$

$$\mathbf{a}_1 = \alpha_2\mathbf{a}_2 + \ldots + \alpha_n\mathbf{a}_n$$

Definiamo $\mathbf{B}$ applicando ad $\mathbf{A}$, in post-moltiplicazione, le trasformazioni elementari $\mathbf{E}_{21}(-\alpha_2), \ldots, \mathbf{E}_{n1}(-\alpha_n)$:

$$\mathbf{B} = \mathbf{A}\mathbf{E}_{21}(-\alpha_2)\ldots\mathbf{E}_{n1}(-\alpha_n)$$

(ricordando che con $\mathbf{E}_{21}(-\alpha_2)$, ad esempio, si toglie alla prima colonna di $\mathbf{A}$ la seconda moltiplicata per α_2). Si ha che:

- essendo tutte matrici del tipo $\mathbf{E}_{ij}(d)$, per la seconda proprietà del determinante non cambiano lo stesso quindi $\varphi(\mathbf{B}) = \varphi(\mathbf{A})$;
- la matrice $\mathbf{B}$ che otteniamo ha la prima colonna nulla, quindi per il lemma 4.1.1, $\varphi(\mathbf{B}) = 0$.

Pertanto si conclude che $\varphi(\mathbf{A}) = 0$ □

Teorema 4.1.3. *Dato $a \in \mathbb{C}$, esiste al più una funzione determinante $\varphi : M_n(\mathbb{C}) \to \mathbb{C}$ tale che $\varphi(\mathbf{I}_n) = a$*

Dimostrazione. L'applicazione φ è determinata una volta che ne conosciamo il valore sulle matrici invertibili, per il lemma 4.1.2 (su quelle non invertibili è sempre 0). Ora, una matrice invertibile $\mathbf{A}$ è prodotto di matrici elementari (proposizione 1.8.11) quindi per le proprietà imposte alla funzione determinante φ possiamo calcolare $\varphi(\mathbf{A})$ una volta che conosciamo il valore di φ sulle matrici elementari:

TODO: Non eccessivamente chiara fino in fondo

$$\varphi(\mathbf{E}_i(c)) = \varphi(\mathbf{I}\mathbf{E}_i(c)) = c\varphi(\mathbf{I}) = ca \tag{4.4}$$
$$\varphi(\mathbf{E}_{ij}(d)) = \varphi(I\mathbf{E}_{ij}(d)) = \varphi(\mathbf{I}) = a \tag{4.5}$$
$$\varphi(\mathbf{E}_{ij}) = \varphi(\mathbf{I}\mathbf{E}_{ij}) = -\varphi(\mathbf{I}) = -a \tag{4.6}$$

L'idea delle equivalenze di sopra è (presumo) che se si hanno 2 funzioni accomunate da queste caratteristiche, di fatto sono uguali e quindi ve ne è solo una. $\square$

Prima di dimostrare l'esistenza di una funzione determinante, vediamo due proprietà importanti.

Teorema 4.1.4 (Binet)**.** *Se* $\varphi : M_n(\mathbb{C}) \to \mathbb{C}$ *è una funzione determinate e* $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$*, allora* $\varphi(\mathbf{AB}) = \varphi(\mathbf{A})\varphi(\mathbf{B})$

Dimostrazione. Fissiamo una matrice $\mathbf{C} \in M_n(\mathbb{C})$ e definiamo un'altra funzione oltre a φ, ossia $\psi : M_n(\mathbb{C}) \to \mathbb{C}$, con

$$\psi(\mathbf{X}) = \varphi(\mathbf{CX})$$

ovvero che data una matrice restituisce il determinante della pre-moltiplicata per $\mathbf{C}$. Allora:

$$\begin{aligned}
\psi(\mathbf{XE}_i(c)) &= \varphi(\mathbf{CXE}_i(c)) = c\varphi(\mathbf{CX}) = c\psi(\mathbf{X}) \\
\psi(\mathbf{XE}_{ij}(d)) &= \varphi(\mathbf{CXE}_{ij}(d)) = \varphi(\mathbf{CX}) = \psi(\mathbf{X}) \\
\psi(\mathbf{XE}_{ij}) &= \varphi(\mathbf{CXE}_{ij}) = -\varphi(\mathbf{CX}) = -\psi(\mathbf{X})
\end{aligned}$$

ossia si vede che i pattern rispettano i dettami della definizione di determinante; dunque ψ è una funzione determinante e si ha anche che $\psi(\mathbf{I}) = \varphi(\mathbf{CI}) = \varphi(\mathbf{C})$. Se consideriamo la funzione $\psi' : M_n(\mathbb{C}) \to \mathbb{C}$ definita da

$$\psi'(\mathbf{X}) = \varphi(\mathbf{C})\varphi(\mathbf{X})$$

con verifiche del tutto analoghe

$$\begin{aligned}
\psi'(\mathbf{XE}_i(c)) &= \varphi(\mathbf{C})\varphi(\mathbf{XE}_i(c)) = c\varphi(\mathbf{C})\varphi(\mathbf{X}) = c\psi'(\mathbf{X}) \\
\psi'(\mathbf{XE}_{ij}(d)) &= \varphi(\mathbf{C})\varphi(\mathbf{XE}_{ij}(d)) = \varphi(\mathbf{C})\varphi(\mathbf{X}) = \psi'(\mathbf{X}) \\
\psi'(\mathbf{XE}_{ij}) &= \varphi(\mathbf{C})\varphi(\mathbf{XE}_{ij}) = \varphi(\mathbf{C})(-\varphi(\mathbf{X})) = -\psi'(\mathbf{X})
\end{aligned}$$

troviamo che anche ψ' è una funzione determinante e che per le peculiarità delle funzioni determinanti le due funzioni coincidono, ossia $\psi(\mathbf{X}) = \psi'(\mathbf{X})$, da cui:

$$\varphi(\mathbf{CX}) = \psi(\mathbf{X}) = \psi'(\mathbf{X}) = \varphi(\mathbf{C})\varphi(\mathbf{X})$$

Calcolando per $\mathbf{C} = \mathbf{A}$ e $\mathbf{X} = \mathbf{B}$

$$\varphi(\mathbf{AB}) = \varphi(\mathbf{A})\varphi(\mathbf{B})$$

abbiamo la tesi. $\square$

Teorema 4.1.5. *Se* $\varphi : M_n(\mathbb{C}) \to \mathbb{C}$ *è una funzione determinante e* $\mathbf{A} \in M_n(\mathbb{C})$*, allora*

$$\varphi(\mathbf{A}^T) = \varphi(\mathbf{A})$$

Dimostrazione. La tesi è ovvia se $\mathbf{A}$ non è invertibile, perché in tal caso lo è anche $\mathbf{A}^T$ (per proposizione 1.6.13) e in entrambi i casi il determinante è nullo. Supponiamo allora $\mathbf{A}$ invertibile e scriviamola come prodotto di matrici elementari (per la 1.8.11) $\mathbf{A} = \mathbf{E}_1\mathbf{E}_2 \ldots \mathbf{E}_r$; la sua trasposta sarà

$$\mathbf{A}^T = \mathbf{E}_r^T \ldots \mathbf{E}_2^T\mathbf{E}_1^T$$

Per il teorema 4.1.4 si ha che

$$\varphi(\mathbf{A}) = \varphi(\mathbf{E}_1) \dots \varphi(\mathbf{E}_r)$$

e al contempo

$$\varphi(\mathbf{A}^T) = \varphi(\mathbf{E}_r^T) \dots \varphi(\mathbf{E}_1^T)$$

ma essendo le matrici elementari tali che $\mathbf{E}_i(c)^T = \mathbf{E}_i(c)$, $\mathbf{E}_{ij}(d)^T = \mathbf{E}_{ji}(d)$ e $\mathbf{E}_{ij}^T = \mathbf{E}_{ij}$ si ha la tesi, ossia $\varphi(\mathbf{A}) = \varphi(\mathbf{A}^T)$. Infatti le matrici $\mathbf{E}_i(c)$ ed $\mathbf{E}_{ij}$ coincidono con la loro trasposta (quindi avranno lo stesso determinante), mentre la moltiplicazione per un $\mathbf{E}_{ij}(d)$ a piacere (o anche la sua trasposta, che rimane del tipo) non modifica il determinante del prodotto, per definizione. □

Osservazione 237. Con la stessa tecnica si dimostra anche il risultato seguente.

Teorema 4.1.6. *Supponiamo esista una funzione determinante $\varphi^{(n)} : M_n(\mathbb{C}) \to \mathbb{C}$, per le matrici di ordine n, tale che $\varphi^{(n)}(\mathbf{I}) = 1$. Allora per ogni $a \in \mathbb{C}$, l'unica funzione determinante $\varphi_a : M_n(\mathbb{C}) \to \mathbb{C}$ tale che $\varphi_a(\mathbf{I}) = a$ si ottiene ponendo:*

$$\varphi_a(\mathbf{A}) = a\varphi^{(n)}(\mathbf{A}) \tag{4.7}$$

per $\mathbf{A} \in M_n(\mathbb{C})$.

Dimostrazione. da fare □ **TODO**: fixme

Osservazione 238. Nella sezione successiva dimostreremo che una funzione determinante come quella del teorema 4.1.6 *esiste* e quindi saremo autorizzati a chiamarla *la* funzione determinante.

Osservazione 239. Per notazione: se $\mathbf{A} \in M_n(\mathbb{C})$, invece di scrivere $\varphi^{(n)}(\mathbf{A})$ scriveremo $\det \mathbf{A}$; questo numero sarà chiamato determinante di $\mathbf{A}$.

4.2 Esistenza di funzioni determinanti

Osservazione 240. Quanto abbiamo visto finora dice che le funzioni determinanti hanno alcune proprietà importanti, purché esistano. Ci sono vari modi di dimostrarne l'esistenza; quello che scegliamo ci permetterà di dare anche un metodo per il calcolo.

Osservazione 241. Quanto abbiamo visto prima riguardo all'unicità delle funzioni determinanti potrebbe apparire una dimostrazione anche dell'esistenza: si scrive una matrice invertibile $\mathbf{A}$ come prodotto di matrici elementari e le proprietà richieste permetterebbero di calcolarne il determinante.
Il fatto è che la scrittura di $\mathbf{A}$ come prodotto di matrici elementari non è affatto unica, quindi occorrerebbe verificare che da due scritture del genere si ricava lo stesso determinante, cosa non facile. Pertanto si preferisce seguire una strada diversa.

Teorema 4.2.1. *Per ogni ordine di matrice $n \geq 1$, esiste una funzione determinante $\varphi^{(n)} : M_n(\mathbb{C}) \to \mathbb{C}$ tale che $\varphi^{(n)}(\mathbf{I}) = 1$*

Dimostrazione. Per induzione; per $n = 1$ l'applicazione identità $M_1(\mathbb{C}) \to \mathbb{C}$ definita da

$$\varphi([a]) = a$$

soddisfa le richieste non mi tornano secondo e terzo requirement **TODO**: fixme

$$\begin{aligned}\varphi([a]\mathbf{E}_1(c)) &= \varphi([a][c]) = \varphi([ac]) = ac = c\cdot\varphi([a])\\ \varphi([a]\mathbf{E}_{11}(d)) &= \varphi([a][d]) = \varphi([ad]) = ad \neq a??\\ \varphi([a]\mathbf{E}_{11}) &= \varphi([a][1]) = \varphi([a]) = a \neq -a??\end{aligned}$$

e quindi la proposizione è verificata.
Supponiamo allora $n > 1$ e di avere già a disposizione la funzione determinante $\varphi(n-1)$ calcolabile su ogni matrice che si ottiene da una data matrice $n \times n$ cancellandone una riga e una colonna.
Fissiamo alcune notazioni: se $\mathbf{A} \in M_n(\mathbb{C})$ denoteremo con a_{ij} il suo coefficiente di posto (i,j); con $\mathbf{A}_{ij}$ la matrice che si ottiene da $\mathbf{A}$ cancellandone la i-esima riga e la j-esima colonna.
Definiamo allora, per $\mathbf{A} \in M_n(\mathbb{C})$

$$\varphi(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j}\varphi^{(n-1)}(\mathbf{A}_{1j}) \tag{4.8}$$

e verifichiamo che questa è una funzione determinante, verificando le tre proprietà della definizione:

1. consideriamo $\mathbf{B} = \mathbf{A}\mathbf{E}_k(c)$ (moltiplichiamo la k-esima colonna per c); abbiamo

$$b_{1j} = \begin{cases} a_{1j} & \text{se } j \neq k \\ ca_{1k} & \text{se } j \overset{(1)}{=} k \end{cases}, \quad \mathbf{B}_{1j} = \begin{cases} \mathbf{A}_{1j}\mathbf{E}_k(c) & \text{se } j \neq k \\ \mathbf{A}_{1k} & \text{se } j = k \end{cases}$$

dove in (1) non vi è bisogno di moltiplicare perché la colonna è stata soppressa. Perciò applicando la definizione:

$$\begin{aligned}\varphi(\mathbf{B}) &= \sum_{j=1}^{n} (-1)^{1+j} b_{1j}\varphi^{(n-1)}(B_{1j})\\ &\overset{(1)}{=} (-1)^{1+k} ca_{1k}\varphi^{(n-1)}(\mathbf{A}_{1k}) + \sum_{j\neq k}^{n} (-1)^{j+1} a_{1j} \underbrace{\varphi^{(n-1)}(\mathbf{A}_{1j}\mathbf{E}_k(c))}_{\overset{(2)}{=} c\varphi(\mathbf{A}_{ij})}\\ &\overset{(3)}{=} c\varphi(\mathbf{A})\end{aligned}$$

dove in (1) si è spezzato nel k-esimo termine e i rimanenti, in (2) applicato l'ipotesi induttiva per $\varphi^{(n-1)}$, in (3) raccolto/portato fuori c.

2. consideriamo $\mathbf{B} = \mathbf{A}\mathbf{E}_{12}$ (qui facciamo mediante un esempio, invertendo prima e seconda colonna, nel caso generale si aggiungono solo complicazioni tecniche). Abbiamo:

$$b_{ij} = \begin{cases} a_{12} & \text{se } j = 1 \\ a_{11} & \text{se } j = 2 \\ a_{1j} & \text{altrimenti} \end{cases} \quad \mathbf{B}_{ij} = \begin{cases} \mathbf{A}_{12} & \text{se } j = 1 \\ \mathbf{A}_{11} & \text{se } j = 2 \\ \mathbf{A}_{1j}\mathbf{E}_{12} & \text{altrimenti} \end{cases}$$

Perciò spezzando in tre parti la sommatoria ed effettuando le sostituzioni

$$\begin{aligned}
\varphi(\mathbf{B}) &= \sum_{j=1}^{n}(-1)^{1+j}b_{1j}\varphi^{n-1}(\mathbf{B}_{1j}) \\
&= (-1)^{1+1}b_{11}\varphi^{(n-1)}(\mathbf{B}_{11}) + (-1)^{1+2}b_{12}\varphi^{n-1}(\mathbf{B}_{12}) + \sum_{j=3}^{n}(-1)^{1+j}b_{1j}\varphi^{(n-1)}(\mathbf{B}_{ij}) \\
&= a_{12}\varphi^{(n-1)}(\mathbf{A}_{12}) - a_{11}\varphi^{(n-1)}(\mathbf{A}_{11}) + \sum_{j=3}^{n}(-1)^{1+j}a_{1j}\underbrace{\varphi^{(n-1)}(\mathbf{A}_{1j}\mathbf{E}_{12})}_{\overset{(1)}{=}-\varphi^{(n-1)}(\mathbf{A}_{ij})} \\
&= -\Big(a_{11}\varphi^{(n-1)}(\mathbf{A}_{11}) - a_{12}\varphi^{(n-1)}(\mathbf{A}_{12})\Big) + (-1)\sum_{j=3}^{n}(-1)^{1+j}a_{ij}\varphi^{(n-1)}(\mathbf{A}_{1j}) \\
&= -\varphi(\mathbf{A})
\end{aligned}$$

applicando ancora una volta l'ipotesi induttiva in (1) e portando successivamente fuori il segno negativo.

3. venendo all'ultima richiesta della definizione, esaminiamo come caso particolare $\mathbf{B} = \mathbf{AE}_{21}(d)$ (aggiungere alla prima colonna la seconda moltiplicata per d) per evitare le complicazioni tecniche. Abbiamo:

$$b_{ij} = \begin{cases} a_{11} + da_{12} & sej = 1 \\ a_{1j} & sej > 1 \end{cases} \qquad \mathbf{B}_{ij} = \begin{cases} \mathbf{A}_1 1 & \text{se } j = 1 \\ \mathbf{X} & sej = 2 \\ \mathbf{A}_{1j}\mathbf{E}_{21}(d) & \text{se } j > 2 \end{cases}$$

dove $\mathbf{X}$ rispetto ad $\mathbf{A}$ ha la prima colonna modificata ($\mathbf{a}_1 + d\mathbf{a}_2$) e tutto il resto uguale (ad eccezione della seconda colonna, mancante). Perciò:

$$\begin{aligned}
\varphi(\mathbf{B}) &= \sum_{j=1}^{n}(-1)^{1+j}b_{1j}\varphi^{(n-1)}(\mathbf{B}_{ij}) \\
&= (-1)^{1+1}\underbrace{(a_{11} + da_{12})}_{=b_{11}}\varphi^{(n-1)}(\mathbf{A}_{11}) + (-1)^{1+2}\underbrace{a_{12}}_{=b_{12}}\varphi^{(n-1)}(\mathbf{X}) + \sum_{j=3}^{n}(-1)^{1+j}a_{1j}\varphi^{(n-1)}(\mathbf{A}_{ij}\mathbf{E}_{21}(d)) \\
&\overset{(1)}{=} (-1)^{1+1}(a_{11} + da_{12})\varphi^{(n-1)}(\mathbf{A}_{11}) + (-1)^{1+2}a_{12}\varphi^{(n-1)}(\mathbf{X}) + \sum_{j=3}^{n}(-1)^{1+j}a_{1j}\varphi^{(n-1)}(\mathbf{A}_{1j}) \\
&\overset{(2)}{=} \varphi(\mathbf{A}) + da_{12}\varphi^{(n-1)}(\mathbf{A}_{11}) + (-1)^{1+2}a_{12}\varphi^{(n-1)}(\mathbf{X}) - (-1)^{1+2}a_{12}\varphi^{(n-1)}(\mathbf{A}_{12})
\end{aligned}$$

dove in (1) abbiamo sostituito al terzo addendo secondo l'ipotesi induttiva e in (2) abbiamo aggiunto e tolto $(-1)^{1+2}a_{12}\varphi^{(n-1)}(\mathbf{A}_{12})$. A questo punto la tesi ($\varphi(\mathbf{B}) = \varphi(\mathbf{A})$) sarà dimostrata se proviamo che:

$$da_{12}\varphi^{(n-1)}(\mathbf{A}_{11}) + (-1)^{1+2}a_{12}\varphi^{(n-1)}(\mathbf{X}) - (-1)^{1+2}a_{12}\varphi^{(n-1)}(\mathbf{A}_{12}) = 0$$

da cui

$$\varphi^{(n-1)}(\mathbf{X}) = \varphi^{(n-1)}(\mathbf{A}_{12}) + d\varphi^{n-1}(\mathbf{A}_{11})$$

Sviluppiamo questo terzo punto nel prosieguo.

In merito a quest'ultimo punto, più in generale, dimostriamo che posto $m = n-1$, $\varphi^{(m)}$ è lineare nella prima colonna (tutti i vettori considerati nel prosieguo sono in $\mathbb{C}^m$) ossia, per prodotto e somma si ha nell'ordine:

- per il prodotto per costante sappiamo già che, per la proprietà 1 della definizione di determinante (post-moltiplicando per $\mathbf{E}_1(\alpha)$), se:

  $$\begin{aligned}\mathbf{Y} &= [\mathbf{y}, \mathbf{y}_2\, \mathbf{y}_3\, \ldots\, \mathbf{y}_m] \\ \mathbf{Y}' &= [\alpha\mathbf{y}, \mathbf{y}_2\, \mathbf{y}_3\, \ldots\, \mathbf{y}_m]\end{aligned}$$

  allora

  $$\varphi^{(m)}(\mathbf{Y}') = \alpha\varphi^{(m)}(\mathbf{Y})$$

- per la somma di due vettori vogliamo dimostrare che se:

  $$\begin{aligned}\mathbf{X}' &= [\mathbf{x}'\, \mathbf{x}_2\, \mathbf{x}_3\, \ldots\, \mathbf{x}_m] \\ \mathbf{X}'' &= [\mathbf{x}''\, \mathbf{x}_2\, \mathbf{x}_3\, \ldots\, \mathbf{x}_m] \\ \mathbf{X} &= [\mathbf{x}' + \mathbf{x}''\, \mathbf{x}_2\, \mathbf{x}_3\, \ldots\, \mathbf{x}_m]\end{aligned}$$

  allora

  $$\varphi^{(m)}(\mathbf{X}) = \varphi^{(m)}(\mathbf{X}') + \varphi^{(m)}(\mathbf{X}'') \tag{4.9}$$

  Approfondiamo questo in quanto segue.

Procedendo per induzione su m (sull'ordine della matrice quadrata), possiamo senz'altro dare l'asserto di equazione 4.9 per valido quando $m = 1$: la funzione $\varphi^{(1)}$ è l'identità su $\mathbb{C}$

$$\begin{aligned}\varphi([a+b]) &= \varphi([a]) + \varphi([b]) \\ a+b &= a+b\end{aligned}$$

Ipotizzando l'asserto vero per $\varphi^{(m-1)}$ dimostriamolo per φ^m; per la notazione:

- siano x_1', x_1'' e x_1 i coefficienti sulla prima riga di $\mathbf{x}', \mathbf{x}''$ e $\mathbf{x}$ rispettivamente
- $\mathbf{y}'$, $\mathbf{y}''$, $\mathbf{y}$ e $\mathbf{y}_j$ ($j = 2, \ldots, m$) i vettori di $\mathbb{C}^{m-1}$ che si ottengono da quelli dati cancellando il coefficiente sulla prima riga

Quando applichiamo la definizione di φ di equazione 4.8 a $\mathbf{X}$

$$\varphi(\mathbf{X}) = \sum_{j=1}^{n} (-1)^{1+j} x_{1j} \varphi^{(n-1)}(\mathbf{X}_{1j})$$

dobbiamo cancellare via via una colonna, oltre alla prima riga; se cancelliamo la prima colonna, rimane la matrice $[\mathbf{y}_2\, \mathbf{y}_3\, \ldots\, \mathbf{y}_m]$, mentre se cancelliamo la seconda colonna otteniamo la matrice $[\mathbf{y} + \mathbf{y}''\, \mathbf{y}_3\, \ldots\, \mathbf{y}_m]$ e così anche per le successive, alle quali potremmo dunque applicare l'ipotesi induttiva. Con facili calcoli si ottiene allora la tesi. da sviluppare □

TODO: fixme

4.3 Come calcolare il determinante

Possiamo riassumere quanto visto nelle sezioni precedenti nel modo seguente.

Teorema 4.3.1. *Per ogni $n > 0$ esiste un'unica funzione determinante $\varphi^{(n)} : M_n(\mathbb{C}) \to \mathbb{C}$ tale che $\varphi^{(n)}(\mathbf{I}_n) = 1$*

Osservazione 242. Per semplicità di notazione si scrive $\varphi^{(n)}(\mathbf{A}) = \det \mathbf{A}$; il numero $\det \mathbf{A}$ si chiama determinante di $\mathbf{A}$.

Osservazione 243. Possiamo allora, usando le notazioni precedenti, trascrivere i risultati ottenuti nel modo seguente.

Teorema 4.3.2 (Binet)**.** *Se $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$, allora*

$$\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B}) \tag{4.10}$$

Teorema 4.3.3. *Una matrice $\mathbf{A} \in M_n(\mathbb{C})$ non è invertibile se solo se* $\det \mathbf{A} = 0$

Teorema 4.3.4. *Per ogni $\mathbf{A} \in M_n(\mathbb{C})$,* $\det \mathbf{A} = \det(\mathbf{A}^T)$

Osservazione 244. Esistono almeno due modi per calcolare il determinante di una matrice. Uno viene dalla definizione di funzione determinante, l'altro dalla dimostrazione di esistenza

4.3.1 Primo metodo di calcolo

Proposizione 4.3.5. *Sia $\mathbf{A}$ una matrice $n \times n$ invertibile. Sia s il numero di scambi di riga e siano $c_1, c_2, \dots, c_n$ i pivot per ottenere da $\mathbf{A}$ una matrice in forma ridotta. Allora*

$$\det \mathbf{A} = (-1)^s c_1 c_2 \dots c_n \tag{4.11}$$

Dimostrazione. Sfruttiamo la fattorizzazione 1.8.8, scrivendo $\mathbf{A} = \mathbf{P}^{-1}\mathbf{LU} = \mathbf{P}^T\mathbf{LU}$ con $\mathbf{U}$ unitriangolare superiore (forma ridotta) dato che $\mathbf{A}$ è quadrata e invertibile; allora

- $\mathbf{P}$ è il prodotto di s matrici del tipo $\mathbf{E}_{ij}$ (scambi di riga);
- $\mathbf{L}$ è prodotto di matrici elementari dove compaiono le matrici $\mathbf{E}_1(c_1), \dots, \mathbf{E}_n(c_n)$ (prodotti di riga) oltre a matrici del tipo $\mathbf{E}_{ij}(d)$ (sostituzione riga con combinazione lineare)
- $\mathbf{U}$ è prodotto di matrici elementari del tipo $\mathbf{E}_{ij}(d)$

Siccome det è una funzione determinante che vale 1 sulla matrice identità, il teorema di Binet permette di calcolare $\det A$ come prodotto dei determinanti delle matrici elementari suddette. Si ha, applicando la definizione per $\mathbf{A} = \mathbf{I}$

$$\begin{aligned}
&\varphi(\mathbf{IE}_i(c)) = c\varphi(\mathbf{I}) = c \cdot 1 = c \\
&\varphi(\mathbf{IE}_{ij}(d)) = \varphi(\mathbf{I}) = 1 \\
&\varphi(\mathbf{IE}_{ij}) = -\varphi(\mathbf{I}) = -1
\end{aligned}$$

Le matrici $\mathbf{E}_{ij}(d)$ hanno determinante 1 (quindi ignorabili ai fini del calcolo), quelle di tipo $\mathbf{E}_i(c)$ hanno determinante c e quelle del tipo E_{ij} hanno determinante -1 (per cui nella formula 4.11, $(-1)^s$ è per gli scambi di riga se si hanno s scambi di riga), da cui si ha subito la conclusione. □

Osservazione 245. Vediamo alcune conseguenze del metodo di calcolo

Corollario 4.3.6. *Se* $\mathbf{A}$ *è una matrice quadrata a coefficienti reali, allora* $\det \mathbf{A}$ *è un numero reale*

Dimostrazione. Siccome l'eliminazione su $\mathbf{A}$ si può eseguire interamente con numeri reali, i pivot saranno reali è così il loro prodotto □

Corollario 4.3.7. *Il determinante di una matrice triangolare è il prodotto degli elementi sulla diagonale*

Dimostrazione. Una matrice triangolare non è invertibile se e solo se uno dei coefficienti sulla diagonale è nullo (e in tal caso il determinante è nullo).
Se la matrice è invertibile e triangolare superiore, i pivot dell'eliminazione, che si può eseguire senza scambi di righe, son esattamente i coefficienti sulla diagonale. Per le triangolari inferiori, basta applicare la trasposizione. □

Corollario 4.3.8. *Se* $\mathbf{A} \in M_n(\mathbb{C})$ *allora* $\det \overline{\mathbf{A}} = \overline{\det \mathbf{A}}$

Dimostrazione. Se $\mathbf{A}$ è non invertibile, non c'è niente da dimostrare: il coniugato di zero è zero. Supponiamo allora che $\mathbf{A}$ sia invertibile e scriviamola nella forma $\mathbf{A} = \mathbf{P}^T\mathbf{L}\mathbf{U}$ con $\mathbf{L}$ triangolare inferiore e $\mathbf{U}$ unitriangolare superiore. Allora abbiamo $\overline{\mathbf{A}} = \mathbf{P}^T\overline{\mathbf{L}\mathbf{U}}$ (qui $\mathbf{P}^T$ è reale quindi coincide con il suo coniugato); per il teorema di Binet

$$\det(\overline{\mathbf{A}}) = \det(\mathbf{P}^T)\det(\overline{\mathbf{L}})\det(\overline{\mathbf{U}})$$

e per il corollario precedente $\det \overline{\mathbf{U}} = 1$ ($\mathbf{U}$ è unitriangolare superiore quindi ha diagonale composta da tutti 1). Inoltre, siccome $\det \overline{\mathbf{L}}$ è il prodotto dei coefficienti sulla diagonale, e dato che il coniugato di un prodotto di complessi coincide con il prodotto dei coniugati, vale $\det \mathbf{L} = \overline{\det \mathbf{L}}$.
Quindi si ha che

$$\det \overline{\mathbf{A}} = \det \mathbf{P}^T \cdot \overline{\det \mathbf{L}}$$

ma si ha anche che

$$\det \mathbf{A} = \det \mathbf{P}^T \cdot \det \mathbf{L}$$

Prendendo il coniugato di quest'ultima:

$$\begin{aligned}\overline{\det \mathbf{A}} = \overline{\det \mathbf{P}^T \cdot \det \mathbf{L}} &\overset{(1)}{=} \overline{\det \mathbf{P}^T} \cdot \overline{\det \mathbf{L}} \\ &\overset{(2)}{=} \det \mathbf{P}^T \cdot \overline{\det \mathbf{L}} \\ &= \det \overline{\mathbf{A}}\end{aligned}$$

con (1) per il fatto che coniugato del prodotto è prodotto dei coniugati e (2) poiché $\det \mathbf{P}^T \in \mathbb{R}$ e quindi coincide con il suo coniugato. □

Corollario 4.3.9. *Se* $\mathbf{A} \in M_n(\mathbb{C})$*, allora* $\det(\mathbf{A}^H) = \overline{\det \mathbf{A}}$

Dimostrazione. Si ha:

$$\det \mathbf{A}^H \overset{(1)}{=} \det \overline{\mathbf{A}}^T \overset{(2)}{=} \det \overline{\mathbf{A}} \overset{(3)}{=} \overline{\det \mathbf{A}}$$

con (1) dato che $\mathbf{A}^H = (\overline{\mathbf{A}})^T$, (2) per teorema 4.1.5, (3) per il corollario 4.3.8, □

4.3.2 Secondo metodo di calcolo

Osservazione 246. Il secondo metodo di calcolo lo ricaviamo dalla dimostrazione di esistenza. Ricordiamo che:

- denotiamo con $\mathbf{A}_{ij}$ la matrice che si ottiene da $\mathbf{A}$ cancellandone la i-esima riga e la j-esima colonna (supponendo che $\mathbf{A}$ sia $n \times n$ e che $n > 1$);
- il determinante di una matrice 1×1, cioè un numero, è il numero stesso.

Proposizione 4.3.10 (Formule di Laplace)**.** *Sia* $A \in M_n(\mathbb{C}), n > 1$ *. Allora*

$$\det \mathbf{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ij} \quad \textit{sviluppo secondo la } i\textit{-esima riga} \tag{4.12}$$

$$\det \mathbf{A} = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ij} \quad \textit{sviluppo secondo la } j\textit{-esima colonna} \tag{4.13}$$

Dimostrazione. La prima delle due formule, nel caso $i = 1$ è esattamente la formula (in via più generale) usata per dimostrare l'esistenza delle funzioni determinanti in equazione 4.8.

Prima di dimostrarla nel caso generale, proviamo a calcolare quanti scambi di righe occorrono per portare la i-esima riga sulla prima e far scalare le altre. Per esempio supponiamo $n = 5$ (matrice 5×5) e di voler portare la terza riga al posto della prima, la prima al posto della seconda e la seconda al posto della terza lasciando ferme la quarta e la quinta. Se la matrice è, scritta per righe,

$$\begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_3 \\ \mathbf{R}_4 \\ \mathbf{R}_5 \end{bmatrix}$$

la vogliamo far diventare

$$\begin{bmatrix} \mathbf{R}_3 \\ \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_4 \\ \mathbf{R}_5 \end{bmatrix}$$

Occorrerà impiegare la $\mathbf{E}_{13}$, poi la $\mathbf{E}_{23}$.
Supponiamo invece di voler portare la quarta riga in alto ottenendo la matrice

$$\begin{bmatrix} \mathbf{R}_4 \\ \mathbf{R}_1 \\ \mathbf{R}_2 \\ \mathbf{R}_3 \\ \mathbf{R}_5 \end{bmatrix}$$

Occorrerà la $\mathbf{E}_{14}$, poi la $\mathbf{E}_{34}$ e infine la $\mathbf{E}_{23}$. Ci si accorge che, in ogni caso, il numero di scambi di riga necessari per portare in alto la i-esima riga è $i - 1$.
Se $\mathbf{B}$ è la matrice ottenuta portando la i-esima riga in alto e facendo scalare le altre, avremo che:

$$\mathbf{B} = \underbrace{\mathbf{E}_{i_1 i_2} \cdot \ldots \cdot \mathbf{E}_{i_m i_n}}_{1-i \text{ termini}} \mathbf{A} \iff \det \mathbf{B} = \det \left(\mathbf{E}_{i_1 i_2} \cdot \ldots \cdot \mathbf{E}_{i_m i_n} \mathbf{A} \right)$$

da cui $\det \mathbf{B} = (-1)^{i-1} \det \mathbf{A}$, perché abbiamo applicato $i-1$ scambi di righe. Inoltre $\mathbf{B}_{1j} = \mathbf{A}_{ij}$ (togliere la prima riga a $\mathbf{B}$ o la i-esima ad $\mathbf{A}$ si ottiene la stessa matrice) quindi

$$\det \mathbf{A} = (-1)^{i-1} \det \mathbf{B} \overset{(1)}{=} \sum_{j=1}^{n} (-1)^{i-1}(-1)^{1+j} b_{1j} \det \mathbf{B}_{1j} \overset{(2)}{=} \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ij}$$

dove in (1) si è sostituita la formula per $\det \mathbf{B}$ di 4.8 e si è portato dentro sommatoria il termine $(-1)^{i-1}$ non dipendente da j, in (2) si è sfruttata la sostituzione derivante dall'equivalenza $b_{1j} \det \mathbf{B}_{1j} = a_{ij} \det \mathbf{A}_{ij}$ (alla luce degli scambi di riga effettuati).
La formula dello sviluppo per colonne segue dal fatto che $\det \mathbf{A} = \det(\mathbf{A}^T)$. □

Osservazione 247 (Efficienza 1: utilità del metodo)*.* Questo metodo di calcolo è utile quando la matrice $\mathbf{A}$ ha molti zeri, ma non è molto efficiente nel caso generale. Infatti si richiedono n prodotti per il primo sviluppo; ciascuno dipende dal determinante di una matrice $(n-1) \times (n-1)$ e così via.

Osservazione 248 (Efficienza 2: comparazione con il primo metodo)*.* Lo sviluppo completo per righe richiede un gran numero ($n!$) di moltiplicazioni. Il metodo basato sull'eliminazione di Gauss è certamente più efficiente, perché richiede n^2 moltiplicazioni, $(n-1)^2$ per il secondo e così via. In totale il numero di moltiplicazioni è uguale $s_n = n(n+1)(2n+1)/6$

Esempio 4.3.1. Abbiamo $s_4 = 30$ e $4! = 24$; invece $s_5 = 55$ e $5! = 120$, $s_6 = 91$ e $6! = 720$. Come si vede il metodo con eliminazione è di gran lunga più efficiente.

Osservazione 249*.* Alcuni risultati utili che possiamo giustificare ora seguono

Teorema 4.3.11.

$$\det \mathbf{I} = 1$$

Dimostrazione. Si vede facilmente applicando equivalentemente la formula delle righe o delle colonne del determinante □

Teorema 4.3.12. *Se* $\mathbf{A}$ *è invertibile*

$$\det(\mathbf{A}^{-1}) = \frac{1}{\det \mathbf{A}}$$

Dimostrazione. Infatti se

$$\begin{aligned}\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} &\iff \det(\mathbf{A}\mathbf{A}^{-1}) = \det(\mathbf{I}) \iff \det(\mathbf{A})\det(\mathbf{A}^{-1}) = 1 \\ &\iff \det(\mathbf{A}^{-1}) = \frac{1}{\det \mathbf{A}}\end{aligned}$$

□

4.3.3 Esercizi

Esempio 4.3.2. Calcolare il determinante di

$$\mathbf{A} = \begin{bmatrix} 2-i & 1 & 0 \\ 2 & 1+i & 3 \\ i & 1 & 1 \end{bmatrix}$$

Si ha

$$\begin{aligned}\det \mathbf{A} &= (-1)^{1+1}(2-i)\det \begin{bmatrix} 1+i & 3 \\ 1 & 1 \end{bmatrix} + (-1)^{1+2}\det \begin{bmatrix} 2 & 3 \\ i & 1 \end{bmatrix} \\ &= (2-i)\left[(1+i)1-3\right] - (2\cdot 1 - 3i) = (2-i)(-2+i) - 2 + 3i \\ &= -(2-i)^2 - 2 + 3i = -(4-1-4i) - 2 + 3i \\ &= -3 + 4i - 2 + 3i = -5 + 7i\end{aligned}$$

Esempio 4.3.3. Calcolare il determinante di

$$\mathbf{B} = \begin{bmatrix} i & 1 & 1+i \\ -1 & 1 & 2 \\ i & i & 2 \end{bmatrix}$$

Si ha

$$\begin{aligned}\det \mathbf{B} &= (-1)^{2+1}(-1)\det \begin{bmatrix} 1 & 1+i \\ i & 1 \end{bmatrix} + 1(-1)^{2+2}\det \begin{bmatrix} i & 1+i \\ i & 1 \end{bmatrix} + 2(-1)^{2+3}\det \begin{bmatrix} i & 1 \\ i & i \end{bmatrix} \\ &= [1\cdot 1 - i(1+i)] + [i - i(1+i)] - 2[-1-i] \\ &= (1-i+1) + (i-i+1) + 2 + 2i \\ &= (2-i) + 1 + 2 + 2i = 5 + i\end{aligned}$$

Esempio 4.3.4. Calcolare il determinante di

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1+i \end{bmatrix}$$

Si ha

$$\det \mathbb{C} = 1\cdot(-1)^{1+2}\det \begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1+i \end{bmatrix} + \det \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1+i \end{bmatrix} + (-1)^{1+4}\det \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Sviluppiamone i pezzi

$$\begin{aligned}\det \begin{bmatrix} 2 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1+i \end{bmatrix} &= 1(-1)^{2+1}\det \begin{bmatrix} 2 & 1 \\ 0 & 1+i \end{bmatrix} + 1(-1)^{2+2}\det \begin{bmatrix} 2 & 1 \\ 1 & 1+i \end{bmatrix} \\ &= -(2(1+i)) + 2(1+i) - 1 = -2 - 2i +\end{aligned}$$

Inoltre

$$\begin{aligned}\det \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1+i \end{bmatrix} &= (-1)^{2+1}\det \begin{bmatrix} 1 & 1 \\ 1 & 1+i \end{bmatrix} + (-1)^{2+2}\det \begin{bmatrix} 2 & 1 \\ 1 & 1+i \end{bmatrix} \\ &= -1((1+i)-1) + 2 + 2i - 1 = 1 + i\end{aligned}$$

e infine

$$\begin{aligned}\det \begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} &= (-1)^{3+1}\det \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} + (-1)^{3+2}\det \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \\ &= 1 - 2 + (-1)\cdot 0 = -1\end{aligned}$$

Concludendo

$$\det \mathbf{C} = (-1)(-1) + (1+i) + 1 = 1 + 1 + i + 1 = 3 + i$$

4.3.4 Altre formule utili del determinante

Proposizione 4.3.13 (Formula del determinante a blocchi). *Se la matrice* $\mathbf{M}$ *in forma bordata è*

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^T & d \end{bmatrix}$$

con $\mathbf{A}$ *quadrata invertibile, si ha che*

$$\det \mathbf{M} = d \cdot \det \mathbf{A} - \mathbf{c}^T (\operatorname{adj} \mathbf{A}) \mathbf{b} \tag{4.14}$$

o dato che $\mathbf{A}$ *è quadrata ed invertibile, possiamo rielaborare quest'ultima alla luce del fatto che, come si vedrà per proposizione 4.4.1 si ha*

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A}$$

Allora

$$\begin{aligned} \det \mathbf{M} &= d \cdot \det \mathbf{A} - \mathbf{c}^T (\det \mathbf{A}) \mathbf{A}^{-1} \mathbf{b} \\ &= \det \mathbf{A} \left[d - \mathbf{c}^T \mathbf{A}^{-1} \mathbf{b} \right] \end{aligned}$$

e in sintesi

$$\det \mathbf{M} = \det \mathbf{A} \left[d - \mathbf{c}^T \mathbf{A}^{-1} \mathbf{b} \right] \tag{4.15}$$

Dimostrazione. Da fare (esercizio 4.13 pag 182): se ouò essere utile, dato che $\mathbf{A}$ è quadrata invertibile,

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A}$$

per proposizione 3.9 pag 176. □

4.3.5 L'uso di R e di maxima

La funzione in R è `det`, da applicarsi a matrici reali.

```
## il determinante di matrice triangolare è il prodotto degli elementi
## sulla diagonale principale
m <- rmatrix(c(1, 3, 5,
               0, 4, 4,
               0, 0, 2), nrow = 3)
det(m)

## [1] 8
```

Se la matrice è complessa bisogna passare a maxima e usare la funzione `determinant`

```
##
## M:matrix([0,1,1,1],[2,1,2,1],[1,1,1,0],[1,1,0,1+%i])
##                               [ 0  1  1    1    ]
##                               [                 ]
##                               [ 2  1  2    1    ]
##                               [                 ]
##                               [ 1  1  1    0    ]
##                               [                 ]
##                               [ 1  1  0  %i + 1 ]
## determinant(M)
##                               2 (%i + 1) - %i + 1
```

4.4 Determinanti e sistemi lineari

TODO: Rivedere meglio qui.

Osservazione 250. I determinanti hanno un uso anche nella risoluzione di sistemi lineari, con un metodo più teorico che pratico, visto che richiederebbe il calcolo di $n+1$ determinanti.

Osservazione 251. Cominciamo con una considerazione: se supponiamo j, k colonne di una matrice quadrata con $j \neq k$, $1 \leq j \leq n$ e $1 \leq k \leq n$, il determinante (sviluppato rispetto alla j-esima colonna) della matrice che si ottiene da $\mathbf{A}$ sostituendo la k-esima al posto della j-esima colonna e lasciando invariate le altre è

$$\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ik} = 0 \tag{4.16}$$

Prima dell'uguaglianza, rispetto alla formula 4.13 l'unica cosa che cambia è $\mathbf{A}_{ik}$ al posto di $\mathbf{A}_{ij}$. Tale determinante è nullo perché la matrice ha due colonne uguali e pertanto non è invertibile.
Possiamo scrivere una sorta di generalizzazione che racchiuda equazione 4.13 e 4.16 come casi particolari:

$$\sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \mathbf{A}_{ik} = (\det \mathbf{A})\delta_{jk}, \quad \delta_{jk} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases} \tag{4.17}$$

(sono cioè i coefficienti della matrice identità). Questo permetterà di scrivere in modo esplicito l'inversa di una matrice invertibile $\mathbf{A}$ di ordine $n > 1$.

Definizione 4.4.1 (Matrice aggiunta)**.** Si definisce come $\operatorname{adj} \mathbf{A} = [a_{ij}]$ dove

$$a_{ji} = (-1)^{i+j} \det \mathbf{A}_{ij} \tag{4.18}$$

Osservazione 252. In altre parole il coefficiente di posto (j, i) della matrice $\operatorname{adj} \mathbf{A}$ si calcola moltiplicando per $(-1)^{i+j}$ il determinante della matrice ottenuta cancellando la i-esima riga e la ej-esima colonna. Si faccia attenzione all'inversione degli indici.

Proposizione 4.4.1. *Se* $\mathbf{A}$ *è una matrice quadrata allora* $(\operatorname{adj} \mathbf{A})\mathbf{A} = (\det \mathbf{A})\mathbf{I}$. *In particolare, se* $\mathbf{A}$ *è invertibile*

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj} \mathbf{A}$$

Dimostrazione. Basta eseguire il calcolo di $(\operatorname{adj}\mathbf{A})\mathbf{A}$; il coefficiente di posto (k,j) in questo prodotto è (mediante i effettuiamo il ciclo sulle righe di $\operatorname{adj}\mathbf{A}$ e sule colonne di $\mathbf{A}$)

$$\sum_{i=1}^{n} \alpha_{ki} a_{ij} \overset{(1)}{=} \sum_{i=1}^{n} a_{ij}(-1)^{i+k} \det \mathbf{A}_{ik} \overset{(2)}{=} (\det \mathbf{A})\delta_{jk}$$

dove in (1) abbiamo sostituito α_{ki} in base alla definizione della matrice aggiunta e in (2) si è sfruttata 4.17. Pertanto:

$$(\operatorname{adj}\mathbf{A})\mathbf{A} = (\det \mathbf{A})\mathbf{I}_n$$

dalla quale si conclude. □

Esempio 4.4.1. L'unico caso in cui questo metodo sia migliore del metodo dell'eliminazione per il calcolo dell'inversa è quando $n = 2$. Infatti in questo caso i determinanti per calcolare l'aggiunta sono semplici numeri. Se

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ -2 & 4 \end{bmatrix}$$

allora $\det \mathbf{A} = 10$ e quindi, per il calcolo dell'aggiunta:

- $(-1)^{1+1} \cdot 4 = 4$ va al posto $(1,1)$
- $(-1)^{1+2} \cdot (-2) = 2$ va al posto $(2,1)$
- $(-1)^{2+1} \cdot 3 = 2$ va al posto $(1,2)$
- $(-1)^{2+2} \cdot 1 = 1$ va al posto $(2,2)$

e si ha (come già verificato)

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj}\mathbf{A} = \frac{1}{10}\begin{bmatrix} 4 & -3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4/10 & -3/10 \\ 2/10 & 1/10 \end{bmatrix}$$

cioè: si scambiano gli elementi della diagonale e gli altri vengono cambiati di segno. Poi tutti i coefficienti vanno divisi per il determinante.

Definizione 4.4.2. Se $\mathbf{A}$ è una matrice $n \times n$ e $\mathbf{b} \in \mathbb{C}^n$ indichiamo con $\mathbf{A}(\mathbf{b}, j)$ la matrice che si ottiene da $\mathbf{A}$ sostituendo la j-esima colonna con $\mathbf{b}$.

Osservazione 253. Possiamo analizzare il sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ tenendo conto della teoria appena sviluppata e ottenere un classico risultato dovuto a Cramer.

Teorema 4.4.2 (Cramer)**.** *Siano* $\mathbf{A}$ *una matrice* $n \times n$ *invertibile e* $\mathbf{b} \in \mathbb{C}^n$*. Se la soluzione del sistema* $\mathbf{A}\mathbf{x} = \mathbf{b}$ *è* $\mathbf{x} = [x_1\, x_2\, \dots\, x_n]^T$ *allora*

$$x_k = \frac{\det \mathbf{A}(\mathbf{b}, k)}{\det \mathbf{A}}$$

Dimostrazione. Scriviamo come prima $\operatorname{adj}\mathbf{A} = [a_{ij}]$ e poniamo $\mathbf{b} = [b_1\, b_2\, \dots\, n]^T$. Sappiamo che $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ dunque post-moltiplicando $(\det \mathbf{A})\mathbf{I} = (\operatorname{adj}\mathbf{A})\mathbf{A}$ per $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$

$$(\det \mathbf{A})\mathbf{I}\mathbf{x} = (\operatorname{adj}\mathbf{A})\mathbf{A}\mathbf{A}^{-1}\mathbf{b}$$

per cui

$$(\det \mathbf{A})\mathbf{x} = (\operatorname{adj} \mathbf{A})\mathbf{b} \tag{4.19}$$

e dunque guardando il prodotto $(\det \mathbf{A})\mathbf{x}$ (vettore) al k-esimo elemento

$$(\det \mathbf{A})x_k \overset{(1)}{=} \sum_{j=1}^{n} \alpha_{kj} b_j = \sum_{j=1}^{n} (-1)^{j+k} b_j \det \mathbf{A}_{jk} = \det \mathbf{A}(\mathbf{b}, k)$$

dove in (1) α_{kj} proviene da $\operatorname{adj} \mathbf{A}$ e b_j da $\mathbf{b}$, alla luce di 4.19. Da quest'ultima si conclude. $\square$

Esempio 4.4.2. Il sistema lineare

$$\begin{cases} ax + by = l \\ cx + dy = m \end{cases}$$

ha una ed una sola soluzione se e solo se $ad - bc \neq 0$ e in tal caso

$$x = \frac{\det \begin{bmatrix} a & l \\ c & m \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \frac{am - cl}{ad - bc}, \quad y = \frac{\det \begin{bmatrix} l & b \\ m & d \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \frac{dl - bm}{ad - bc}$$

È questa la nota regola di Cramer. Anche in questo caso il metodo è efficiente solo per sistemi 2×2

Capitolo 5

Autovalori e autovettori

Contents

5.1 Un esempio di modello discreto

5.1.1 Il modello

A un certo istante iniziale t_0 sono presenti in un territorio un numero P_0 di animali predatori e un numero p_0 di prede:

- in assenza di prede, i predatori tenderebbero a estinguersi velocemente. Supponiamo che all'istante t_{i+1} il numero di predatori P_{i+1} sia una frazione del numero P_i all'istante t_i per esempio

$$P_{i+1} = \frac{3}{5}P_i$$

Ma in presenza di prede il numero P_{i+1} aumenta proporzionalmente al numero di prede p_i disponibili al tempo precedente; supponiamo per esempio che

$$P_{i+1} = \frac{3}{5}P_i + \frac{1}{2}p_i$$

- simmetricamente, in assenza di predatori le prede tenderebbero ad aumentare nel tempo; supponiamo che all'istante t_{i+1} il numero di prede p_{i+1} sia aumentato rispetto all'istante precedente p_i seguendo:

$$p_{i+1} = \frac{6}{5}p_i$$

Ma in presenza di predatori, il numero p_{t+1} decresce esattamente del numero di prede uccise dai predatori nell'intervallo di tempo $[t_i, t_{i+1}]$. Se k denota il numero medio di prede uccise in un intervallo temporale da ciascun predatore risulta

$$p_{i+1} = \frac{6}{5}p_i - kP_i$$

Si suppone che il numero k sia costante nel tempo.

Si chiama vettore della popolazione all'istante t_i la coppia ordinata di numeri

$$\mathbf{v}_i = \begin{bmatrix} P_i \\ p_i \end{bmatrix}$$

Il modello preda-predatore è quindi completamente descritto dalle due equazioni

$$\begin{aligned} P_{i+1} &= \frac{3}{5}P_i + \frac{1}{2}p_i \\ p_{i+1} &= -kP_i + \frac{6}{5}p_i \end{aligned}$$

nonché dal vettore della popolazione iniziale $\mathbf{v}_0 = [P_0\, p_0]^T$. Possiamo riscrivere le equazioni del modello in forma matriciale:

$$\mathbf{v}_{i+1} = \mathbf{A}(k)\mathbf{v}_i, \quad (i = 0, 1, 2, \ldots)$$

con

$$\mathbf{A}(k) = \begin{bmatrix} 3/5 & 1/2 \\ -k & 6/5 \end{bmatrix}$$

Per ogni $i > 0$ si ha dunque

$$\mathbf{v}_i = \mathbf{A}(k)\mathbf{v}_{i-1} = \mathbf{A}(k)^2\mathbf{v}_{i-2} = \mathbf{A}(k)^3\mathbf{v}_{i-3} = \ldots = [\mathbf{A}(k)]^i\mathbf{v}_0 \qquad (5.1)$$

quindi l'andamento del vettore della popolazione nel tempo dipende dalle potenze della matrice $\mathbf{A}(k)$. Scopo del modello è predire cosa avverrà nel tempo del vettore della popolazione.

	t_0	t_1	t_2	t_4	t_{15}	t_{29}	t_{99}	
P_i	100	560	931	1244	1523	5107	19409	15328199
p_i	1000	1190	1372	1553	1739	5111	19409	15328199

Tabella 5.1: Simulazione 1: $k = 10/100$

	t_0	t_1	t_2	t_4	t_{15}	t_{29}	t_{59}	t_{79}	t_{99}	
P_i	100	560	927	1214	1434 1654	713	43	3	0	
p_i	1000	1182	1317	1413	1477	1177	459	25	2	0

Tabella 5.2: Simulazione 2: $k = 18/100$

5.1.2 Simulazioni della dinamica

Facciamo tre simulazioni per il parametro k, numero medio di prede uccise in un intervallo temporale da ciascun predatore e vediamo cosa succede del vettore della popolazione, partendo da una situazione iniziale rappresentata dal vettore:

$$\mathbf{v}_0 = \begin{bmatrix} 100 \\ 10000 \end{bmatrix}$$

Verranno rappresentati in tre tabelle i numeri di predatori e prede in diversi istanti t_i. Naturalmente tali numeri saranno arrotondati all'unità:

- nella simulazione di tabella 5.1 ($k = 10/100$) le due popolazioni tendono a crescere, quella dei predatori più velocemente di quella delle prede, diventando da un certo istante in poi uguali. Al tendere del tempo all'infinito le due popolazioni tendono a crescere indefinitamente. Sinteticamente: i predatori predano troppo poco. Il parametro k deve aumentare perché le due popolazioni non si espandano troppo;

- in tabella 5.2 ($k = 18/100$) dopo una buona espansione iniziale le due popolazioni delle prede e dei predatori diminuiscono entrambi gradualmente fino ad azzerarsi; qui i predatori predano troppo;

- in tabella 5.3 infine ($k = 16/100$) le due popolazioni all'inizio tendono a crescere in una misura che è a metà strada tra i due casi precedenti e che per valori grandi del tempo si stabilizzano sulle quantità $P = 480$ e $p = 1920$. Si è così raggiunto l'equilibrio.

È naturale chiedersi da che cosa dipende essenzialmente il diverso comportamento del modello nei tre casi simulati, e se i valori dati dal vettore della popolazione all'istante iniziale sono un fattore influente oppure no. Come si vedrà:

- il comportamento del modello dipende da due numeri individuati dalla matrice $\mathbf{A}(k)$, che si chiamano gli *autovalori* della matrice e, più preci-

	t_0	t_1	t_2	t_4	t_{32}	t_{64}	t_{99}	
P_i	100	560	928	1222	1457	478	480	480
p_i	1000	1184	1331	1448	1542	1919	1920	1920

Tabella 5.3: Simulazione 3: $k = 16/100$

samente, dal fatto che il più grande degli autovalori sia minore, uguale o maggiore di 1;

- il vettore della popolazione iniziale influenza il comportamento del modello nel tempo, contrariamente a quanto accade in altri modelli rappresentati da matrici a coefficienti non negativi. Per pervenire alla risposta occorre "diagonalizzare" la matrice $\mathbf{A}(k)$ o, se ciò non è possibile, "triangolarizzarla".

5.1.3 Diagonalizzazione e triangolarizzazione di $\mathbf{A}(k)$

La matrice $\mathbf{A}(k)$ può essere fattorizzata in modo diverso nei tre casi considerati.

5.1.3.1 Caso 1 ($k = 10/100$)

Se $k = 10/100 = 1/10$ la matrice $\mathbf{A}(1/10)$ si può fattorizzare nel prodotto di tre matrici nel modo seguente:

$$\mathbf{A}(1/10) = \overbrace{\begin{bmatrix} 1 & 5 \\ 1 & 1 \end{bmatrix}}^{\mathbf{S}} \overbrace{\begin{bmatrix} 11/10 & 0 \\ 0 & 7/10 \end{bmatrix}}^{\mathbf{D}} \overbrace{\begin{bmatrix} -1/4 & 5/4 \\ 1/4 & -1/4 \end{bmatrix}}^{\mathbf{S}^{-1}}$$

Il primo fattore, chiamiamolo $\mathbf{S}$ è la matrice inversa del terzo fattore, che denotiamo quindi con $\mathbf{S}^{-1}$. Il secondo fattore, chiamiamolo $\mathbf{D}$ è una matrice diagonale. Si dice che la matrice $\mathbf{A}(1/10)$ è stata *diagonalizzata*.
Posto $\mathbf{A} = \mathbf{A}(1/10)$ l'uguaglianza $\mathbf{A} = \mathbf{SDS}^{-1}$ è equivalente a $\mathbf{AS} = \mathbf{SD}$; leggendo la prima e la seconda colonna separatamente dell'uguaglianza abbiamo

$$\mathbf{A} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{11}{10} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$\mathbf{A} \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \frac{7}{10} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

Ciò si esprime dicendo che i due vettori

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$$

sono *autovettori* della matrice relativi rispettivamente agli *autovalori* $\lambda_1 = 11/10$ e $\lambda_2 = 7/10$ (che compaiono sulla diagonale di $\mathbf{D}$), dato che valgono le uguaglianze $\mathbf{Av}_i = \lambda_i \mathbf{v}_i$ $(i = 1, 2)$.
Vedremo a cosa serve tutto questo e da dove saltano fuori i numeri 11/10 e 7/10 (che risultano cruciali per il modello). Prima osserviamo che la diagonalizzazione di $\mathbf{A}$ è utile per calcolare facilmente le potenze della matrice $\mathbf{A}$ (e ciò torna comodo per il calcolo della dinamica del modello, dato che $\mathbf{v}_i = \mathbf{A}(k)^i \mathbf{v}_0$).

Proposizione 5.1.1. *Per ogni intero positivo n si ha*

$$\mathbf{A}^n = \mathbf{SD}^n\mathbf{S}^{-1}$$

Dimostrazione. Proviamola per induzione. Per il passo base si adotta $n = 0$, si ha:

$$\mathbf{A}^0 = \mathbf{S}\mathbf{D}^0\mathbf{S}^{-1} \iff \mathbf{I} = \mathbf{S}\mathbf{I}\mathbf{S}^{-1} \iff \mathbf{I} = \mathbf{I}$$

Supponendo allora che $\mathbf{A}^n = \mathbf{S}\mathbf{D}^n\mathbf{S}^{-1}$ valga per $n \geq 1$ verifichiamolo che vale per $n+1$; moltiplicando il generico $\mathbf{A}^n$ per $\mathbf{A}$ si ha

$$\mathbf{A}^n\mathbf{A} = \mathbf{S}\mathbf{D}^n\mathbf{S}^{-1}\mathbf{S}\mathbf{D}\mathbf{S}^{-1} \iff \mathbf{A}^{n+1} = \mathbf{S}\mathbf{D}^{n+1}\mathbf{S}^{-1}$$

$\square$

Nello specifico caso dato che $\mathbf{D}$ è diagonale sarà

$$\mathbf{D}^n = \begin{bmatrix} (11/10)^n & 0 \\ 0 & (7/10)^n \end{bmatrix}$$

Si può allora concludere che, fissato il vettore della popolazione iniziale $\mathbf{v}_0 = [P_0\, p_0]^T$, sfruttando equazione 5.1, per ogni n risulta:

$$\begin{aligned}
\begin{bmatrix} P_n \\ p_n \end{bmatrix} = \mathbf{A}^n \begin{bmatrix} P_0 \\ p_0 \end{bmatrix} &= \begin{bmatrix} 1 & 5 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (11/10)^n & 0 \\ 0 & (7/10)^n \end{bmatrix} \begin{bmatrix} -1/4 & 5/4 \\ 1/4 & -1/4 \end{bmatrix} \begin{bmatrix} P_0 \\ p_0 \end{bmatrix} \\
&= \begin{bmatrix} (11/10)^n & 5(7/10)^n \\ (11/10)^n & (7/10)^n \end{bmatrix} \begin{bmatrix} (-1/4)P_0 & (5/4)p_0 \\ (1/4)P_0 & (-1/4)p_0 \end{bmatrix} \\
&= \begin{bmatrix} (11/10)^n(-(1/4)P_0 + (5/4)p_0) + (7/10)^n((5/4)P_0 - (5/4)p_0) \\ (11/10)^n(-(1/4)P_0 + (5/4)p_0) + (7/10)^n(p_0 - (1/4)p_0) \end{bmatrix}
\end{aligned}$$

Si noti che nell'ultima matrice la prima parte degli addendo è comune sia a preda che a predatori, mentre la seconda è specifica per ciascuno dei due (per $n \to +\infty$ tende a 0 comunque e possono essere ignorate).
Evidentemente sia P_n che p_n tendono all'infinito al tendere di n all'infinito se e solo se il termine tra parentesi dei primi addendi è positivo ossia:

$$\frac{5}{4}p_0 - \frac{1}{4}P_0 > 0$$

in funzione dal fatto che $\lambda_1 = 11/10 > 1$; l'autovalore (elevato a potenza) incrementa qualsiasi valore positivo facendolo tendere a $+\infty$. Se invece risulta:

$$\frac{5}{4}p_0 - \frac{1}{4}P_0 \leq 0$$

allora sia P_n che p_n si azzerano a un certo istante t_n; infatti se $= 0$ è ovvio, se < 0 si può sviluppare il limite. Ad esempio ponendo per comodità di notazione $P_0 = \alpha$ e $p_0 = \beta$ si ha

$$\begin{aligned}
&\lim_{n\to+\infty} \left(\frac{11}{10}\right)^n \left(-\frac{1}{4}\alpha + \frac{5}{4}\beta\right) + \left(\frac{7}{10}\right)^n \left(\frac{5}{4}\alpha - \frac{5}{4}\beta\right) \\
&\lim_{n\to+\infty} \left(\frac{11}{10}\right)^n \left[\left(-\frac{1}{4}\alpha + \frac{5}{4}\beta\right) + \left(\frac{(7/10)^n}{(11/10)^n}\right)\left(\frac{5}{4}\alpha - \frac{5}{4}\beta\right)\right] \\
&\lim_{n\to+\infty} \underbrace{\left(\frac{11}{10}\right)^n}_{+\infty} \underbrace{\left(-\frac{1}{4}\alpha + \frac{5}{4}\beta\right)}_{<0} = -\infty
\end{aligned}$$

quindi il limite tende a $-\infty$ ma essendo vincolato a valori positivi finirà a 0.
Il lettore può verificare che, posto per esempio $P_0 = 50$ e $p_0 = 10$ le prede si azzerano all'istante t_8 e i predatori all'istante t_{10}.
Nel caso della simulazione 1, i valori $P_0 = 100$ e $p_0 = 1000$ comportano che

$$\frac{5}{4}1000 - \frac{1}{4}100 = 1225 > 0$$

e quindi la popolazione tende a espandersi indefinitamente

5.1.3.2 Caso 2 ($k = 18/100 = 9/50$)

Se $k = 18/100 = 9/50$ la matrice $\mathbf{A}(9/50)$ si può fattorizzare nel prodotto di tre matrici nel modo seguente

$$\mathbf{A}(9/50) = \overbrace{\begin{bmatrix} 5/\sqrt{34} & -3/\sqrt{34} \\ 3/\sqrt{34} & 5/\sqrt{34} \end{bmatrix}}^{\mathbf{Q}} \overbrace{\begin{bmatrix} 9/10 & 17/25 \\ 0 & 9/10 \end{bmatrix}}^{\mathbf{T}} \overbrace{\begin{bmatrix} 5/\sqrt{34} & 3/\sqrt{34} \\ -3/\sqrt{34} & 5/\sqrt{34} \end{bmatrix}}^{\mathbf{Q}^{-1}}$$

Il primo fattore, chiamiamolo $\mathbf{Q}$ è la matrice inversa del terzo fattore, $\mathbf{Q}^{-1}$; inoltre si ha che $\mathbf{Q}^{-1} = \mathbf{Q}^T$ (ossia $\mathbf{Q}$ è una *matrice ortogonale*, come si vedrà in definizione 5.6.3). Il secondo fattore, chiamiamolo $\mathbf{T}$, è una matrice triangolare superiore.
Si dice che la matrice $\mathbf{A}(9/50)$ è stata *triangolarizzata ortogonalmente*; sulla diagonale di $\mathbf{T}$ compaiono ancora gli autovalori di $\mathbf{A}$: $\lambda_1 = \lambda_2 = 9/10$ è autovalore doppio di $\mathbf{A}$.
La triangolarizzazione di $\mathbf{A}$ è ancora utile per calcolare le potenze della matrice $\mathbf{A}$ anche se c'è qualche problema in più rispetto a quando la matrice è diagonalizzabile. Infatti è immediato verificare induttivamente in modo analogo a quanto fatto in proposizione 5.1.1 che per ogni intero positivo n si ha sempre

$$\mathbf{A}^n = \mathbf{Q}\mathbf{T}^n\mathbf{Q}^{-1}$$

che nel caso presente si traduce[1] per $\mathbf{T}^n$ (triangolare superiore) in:

$$\mathbf{T}^n = \begin{bmatrix} (9/10)^n & n(9/10)^{n-1}(17/25) \\ 0 & (9/10)^n \end{bmatrix}$$

[1]Come si può mostrare per induzione

$$\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}^n = \begin{bmatrix} a^n & na^{n-1}b \\ 0 & a^n \end{bmatrix}$$

Infatti per il passo base con $n = 1$

$$\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}^1 = \begin{bmatrix} a^1 & 1 \cdot a^0 \cdot b \\ 0 & a^1 \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$$

Ipotizzando poi che valga per n mostriamo che vale anche per $n+1$, moltiplicando entrambi i membri dell'uguaglianza stessa per $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$. Si ha

$$\begin{bmatrix} a & b \\ 0 & a \end{bmatrix}^n \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} = \begin{bmatrix} a^n & na^{n-1}b \\ 0 & a^n \end{bmatrix} \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \iff \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}^{n+1} = \begin{bmatrix} a^{n+1} & a^n b + n \cdot a^n \cdot b \\ 0 & a^{n+1} \end{bmatrix} = \begin{bmatrix} a^{n+1} & (n+1)a^n b \\ 0 & a^{n+1} \end{bmatrix}$$

$\mathbf{T}^n$ tende alla matrice nulla al tendere di n all'infinito; questo è chiaro per gli elementi della diagonale; per quello in posizione $(1, 2)$ si ha (ponendo $\alpha = P_0$ e $\beta = p_0$):

$$\lim_{n\to+\infty} n \cdot \left(\frac{9}{10}\right)^{n-1} \cdot \frac{17}{25} \iff \lim_{n\to+\infty} \frac{n \cdot \frac{17}{25}}{\left(\frac{10}{9}\right)^{n-1}}$$

che è una forma di indeterminazione ∞/∞; applicando De l'Hopital si ha:

$$\lim_{n\to+\infty} \frac{17/25}{\left(\frac{10}{9}\right)^{n-1}\log(10/9)} = \frac{c}{\infty} = 0$$

$\mathbf{T}^n$ tende a $\mathbb{O}$ dunque perché l'unico autovalore $9/10 < 1$.
Ne consegue che il vettore della popolazione $[P_n\, p_n]^T$ tende ad annullarsi qualunque sia il vettore della popolazione $[P_0\, p_0]^T$ all'istante iniziale.

5.1.3.3 Caso 3 ($k = 16/100$)

Se $k = 16/100$ la matrice $\mathbf{A}(4/25)$ si può fattorizzare nel prodotto di tre matrici nel modo seguente:

$$\mathbf{A}(4/25) = \overbrace{\begin{bmatrix} 5 & 5 \\ 4 & 2 \end{bmatrix}}^{\mathbf{U}} \overbrace{\begin{bmatrix} 1 & 0 \\ 0 & 4/5 \end{bmatrix}}^{\mathbf{\Lambda}} \overbrace{\begin{bmatrix} -1/5 & 1/2 \\ 2/5 & -1/2 \end{bmatrix}}^{\mathbf{U}^{-1}}$$

il primo fattore, chiamiamolo $\mathbf{U}$ è ancora la matrice inversa del terzo fattore, che chiamiamo $\mathbf{U}^{-1}$; il secondo, $\mathbf{\Lambda}$, è una matrice diagonale. La matrice $\mathbf{A}(4/25)$ è stata diagonalizzata e i suoi autovalori sono $\lambda_1 = 1$ e $\lambda_2 = 4/5$.
La diagonalizzazione di $\mathbf{A}$ ci permette, come nel primo caso, di calcolare facilmente le potenze della matrice $\mathbf{A} = \mathbf{A}(4/25)$. Si ha per ogni $n \geq 1$:

$$\mathbf{A}^n = \mathbf{U}\mathbf{\Lambda}^n\mathbf{U}^{-1}$$

con in questo caso

$$\mathbf{\Lambda}^n = \begin{bmatrix} 1 & 0 \\ 0 & (4/5)^n \end{bmatrix}$$

Si può allora concludere, fissato il vettore della popolazione $[P_0\, p_0]^T$ all'istante t_0 per ogni n risulta:

$$\begin{aligned}
\begin{bmatrix} P_n \\ p_n \end{bmatrix} = \mathbf{A}^n \begin{bmatrix} P_0 \\ p_0 \end{bmatrix} &= \begin{bmatrix} 5 & 5 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (4/5)^n \end{bmatrix} \begin{bmatrix} -1/5 & 1/2 \\ 2/5 & -1/2 \end{bmatrix} \begin{bmatrix} P_0 \\ p_0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 5(4/5)^n \\ 4 & 2(4/5)^n \end{bmatrix} \begin{bmatrix} -(1/5)P_0 + (1/2)p_0 \\ (2/5)P_0 - (1/2)p_0 \end{bmatrix} \\
&= \begin{bmatrix} -(1/5)P_0 + (1/2)p_0 + (4/5)^n(2P_0 - (5/2)p_0) \\ -(4/5)P_0 + 2p_0 + (4/5)^n((4/5)P_0 - p_0) \end{bmatrix}
\end{aligned}$$

Evidentemente per n che tende all'infinito il vettore della popolazione $[P_n\, p_n]^T$ tende al vettore

$$\begin{bmatrix} -(1/5)P_0 + (1/2)p_0 \\ -(4/5)P_0 + 2p_0 \end{bmatrix}$$

per l'annullarsi dei secondi addendi $(4/5)^n(\dots)$ di entrambi gli elementi.
Perché ciò abbia significato occorre che $5p_0 > 2P_0$, perché in caso contrario la popolazione giunge ad azzerarsi per qualche valore del tempo t_n (sempre se $\alpha = P_0$, $\beta = p_0$):

$$\begin{cases} -\frac{1}{5}\alpha + \frac{1}{2}\beta > 0 \\ -\frac{4}{5}\alpha + 2\beta > 0 \end{cases} \quad \begin{cases} \frac{1}{2}\beta > \frac{1}{5}\alpha \\ 2\beta > \frac{4}{5}\alpha \end{cases} \quad \begin{cases} 5p_0 > 2P_0 \\ 10p_0 > 4P_0 \end{cases} \quad \to 5p_0 > 2P_0$$

Nel caso della simulazione 3 i valori $P_0 = 100$ e $p_0 = 1000$ comportano che $5 \cdot 1000 > 2 \cdot 100$ e quindi la popolazione tende al valore limite

$$\begin{bmatrix} -(1/5)100 + (1/2)1000 \\ -(4/5)100 + 21000 \end{bmatrix} = \begin{bmatrix} 480 \\ 1920 \end{bmatrix}$$

5.2 Diagonalizzazione e triangolarizzazione di una matrice

Osservazione 254. In questa sezione vediamo nel complesso il percorso da intraprendere, a fini pratici, con riferimento alla matrice introdotta nel prrecedente modello discreto (posto che il procedimento è il medesimo anche per altre); nel seguito esso verrà giustificato nelle sue radici teoriche.

5.2.1 Calcolo di autovalori e autovettori

Vediamo il procedimento per trovare autovalori/autovettori della matrice $\mathbf{A}(k)$, rinviando al seguito la spiegazione di come vi si perviene:

1. trovare il polinomio caratteristico di $\mathbf{A}(k)$, che si ottiene calcolando:

   $$\det(\mathbf{A}(k) - X\mathbf{I})$$

   Esempio 5.2.1.

   $$\mathbf{A}(k) - X\mathbf{I} = \begin{bmatrix} 3/5 & 1/2 \\ -k & 6/5 \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix} = \begin{bmatrix} 3/5 - X & 1/2 \\ -k & 6/5 - X \end{bmatrix}$$

   Nel caso di matrice 2×2 per ricavare il polinomio caratteristico possiamo procedere nel calcolo del determinante agevolmente come:

   $$\begin{aligned} p(X) = \det(\mathbf{A}(k) - X\mathbf{I}) &= \left(\frac{3}{5} - X\right)\left(\frac{6}{5} - X\right) + \frac{1}{2}k \\ &= \frac{18}{25} - \frac{3}{5}X - \frac{6}{5}X + X^2 + \frac{1}{2}k \\ &= X^2 - \frac{9}{5}X + \frac{18}{25} + \frac{1}{2}k \end{aligned}$$

2. gli autovalori di $\mathbf{A}(k)$ non sono che le radici di questo polinomio.
   Noi ci limiteremo a considerare autovalori reali, perché con autovalori complessi la diagonalizzazione perde di significato concreto, pur restando utile per il calcolo delle potenze della matrice.

Esempio 5.2.2.

$$\lambda_1 = \frac{9+\sqrt{9-50k}}{10}, \quad \lambda_2 = \frac{9-\sqrt{9-50k}}{10}$$

Si hanno:

- due radici reali distinte per $\Delta = 9 - 50k > 0$ ossia $k < 9/50$;
- due radici coincidenti con $\lambda = 9/10$ per $k = 9/50$ (cioè quanto previsto nella seconda simulazione);
- non consideriamo gli autovalori complessi che risultino qualora $9 - 50k < 0$

Per $k \leq 9/50$ il più grande autovalore λ_1 è maggiore, uguale o minore di 1 a seconda che risulta $k < 8/50$, $k = 8/50$ o $k > 8/50$:

$$\frac{9+\sqrt{9-50k}}{10} > 1 \iff \frac{9+\sqrt{9-50k}-10}{10} > 0 \iff 1 < \sqrt{9-50k}$$
$$\iff 1 < 9 - 50k \iff k < \frac{8}{50}$$

l'ipotesi $k = 8/50$ è quanto previsto nella terza simulazione.

3. una volta trovati gli autovalori λ_i si cercano i corrispondenti autovettori, ossia dei vettori non nulli $\mathbf{v}_i$ tali che

$$\mathbf{A}(k)\mathbf{v}_i = \lambda_i \mathbf{v}_i$$

o equivalentemente

$$\begin{aligned} \mathbf{A}(k)\mathbf{v}_i &= \lambda_i \mathbf{v}_i \\ \mathbf{A}(k)\mathbf{v}_i &= \lambda_i \mathbf{I}\mathbf{v}_i \\ \mathbf{A}(k)\mathbf{v}_i - \lambda_i \mathbf{I}\mathbf{v}_i &= \mathbf{0} \\ (\mathbf{A}(k) - \lambda_i \mathbf{I})\mathbf{v}_i &= \mathbf{0} \end{aligned}$$

Per cui

$$\mathbf{v}_i \in \mathrm{N}\left(\mathbf{A}(k) - \lambda_i \mathbf{I}\right)$$

Esempio 5.2.3. Completiamo il primo caso studiato, con $k = 10/100 = 1/10$. I due autovalori, come visto, sono

$$\begin{aligned} \lambda_1 &= \frac{9+\sqrt{9-50\frac{1}{10}}}{10} = \frac{9+\sqrt{4}}{10} = \frac{11}{10} \\ \lambda_2 &= \frac{9-2}{10} = \frac{7}{10} \end{aligned}$$

Per giungere agli autovettori calcoliamo le matrici $\mathbf{A}(k) - \lambda_i \mathbf{I}$; per λ_1, λ_2 rispettivamente abbiamo:

$$\begin{bmatrix} 3/5 & 1/2 \\ -1/10 & 6/5 \end{bmatrix} - \begin{bmatrix} 11/10 & 0 \\ 0 & 11/10 \end{bmatrix} = \begin{bmatrix} -1/2 & 1/2 \\ -1/10 & 1/10 \end{bmatrix}$$
$$\begin{bmatrix} 3/5 & 1/2 \\ -1/10 & 6/5 \end{bmatrix} - \begin{bmatrix} 7/10 & 0 \\ 0 & 7/10 \end{bmatrix} = \begin{bmatrix} -1/10 & 1/2 \\ -1/10 & 1/2 \end{bmatrix}$$

Determiniamo ora gli autovettori nel primo caso trovando gli elementi di $\mathrm{N}\left(\mathbf{A}(k)-\lambda_i\mathbf{I}\right)$; per il primo autovalore la riduzione porta a

$$\begin{bmatrix}-1/2 & 1/2\\ -1/10 & 1/10\end{bmatrix}\begin{bmatrix}-1/2 & 1/2\\ -1/2 & 1/2\end{bmatrix}\begin{bmatrix}-1/2 & 1/2\\ 0 & 0\end{bmatrix}\begin{bmatrix}1/2 & -1/2\\ 0 & 0\end{bmatrix}\to\begin{cases}x_2=h\\ \frac{1}{2}x_1-\frac{1}{2}h=0\end{cases}\quad\begin{cases}x_1=h\\ x_2=h\end{cases}\to\begin{bmatrix}h\\ h\end{bmatrix}$$

Quindi vettori del tipo $\begin{bmatrix}h\\ h\end{bmatrix}$, con $h\neq 0$ (altrimenti si ha il vettore nullo), sono autovettori per l'autovalore $\lambda_1=\frac{11}{10}$. Quindi non solo il $\begin{bmatrix}1\\ 1\end{bmatrix}$ già mostrato negli esempi precedenti costituisce un autovettore, ma anche, ad esempio $\begin{bmatrix}2\\ 2\end{bmatrix}$. Infatti la seguente uguaglianza è verificata

$$\begin{bmatrix}3/5 & 1/2\\ -1/10 & 6/5\end{bmatrix}\begin{bmatrix}2\\ 2\end{bmatrix}=\frac{11}{10}\begin{bmatrix}2\\ 2\end{bmatrix}$$

dato che

$$\begin{bmatrix}3/5 & 1/2\\ -1/10 & 6/5\end{bmatrix}\begin{bmatrix}2\\ 2\end{bmatrix}=\begin{bmatrix}\frac{11}{5}\\ \frac{11}{5}\end{bmatrix}$$

Infine per quanto concerne gli autovettori dell'autovalore $\lambda_2=\frac{7}{10}$ si ha

$$\begin{bmatrix}-1/10 & 1/2\\ -1/10 & 1/2\end{bmatrix}\begin{bmatrix}-1/10 & 1/2\\ 0 & 0\end{bmatrix}\begin{bmatrix}1/10 & -1/2\\ 0 & 0\end{bmatrix}\to\begin{cases}x_2=h\\ \frac{1}{10}x_1-\frac{1}{2}h=0\end{cases}\quad\begin{cases}x_2=h\\ \frac{1}{10}x_1=\frac{1}{2}h\end{cases}\quad\begin{cases}x_1=5h\\ x_2=h\end{cases}\to$$

Quindi vettori del tipo $\begin{bmatrix}5h\\ h\end{bmatrix}$ con $h\neq 0$ costituiscono autovettore per λ_2. Tra questi, come si è già visto ad esempio, anche $\begin{bmatrix}5\\ 1\end{bmatrix}$ come si può facilmente verificare seguendo i passaggi svolti ad esempio per il precedente autovalore.

Esempio 5.2.4. In merito al caso con $k=9/50$, si ha $\lambda_1=\lambda_2=\lambda=9/10$. Per gli autovettori procediamo a determinare $\mathrm{N}\left(\mathbf{A}(9/50)-\mathbf{I}9/10\right)$

$$\mathbf{A}(9/50)-\mathbf{I}9/10=\begin{bmatrix}3/5 & 1/2\\ -9/50 & 6/5\end{bmatrix}-\begin{bmatrix}9/10 & 0\\ 0 & 9/10\end{bmatrix}=\begin{bmatrix}-3/10 & 1/2\\ -9/50 & 3/10\end{bmatrix}$$

Si ha

$$\begin{bmatrix}-3/10 & 1/2\\ -9/50 & 3/10\end{bmatrix}\begin{bmatrix}-3/50 & 1/10\\ -9/50 & 3/10\end{bmatrix}\begin{bmatrix}-9/50 & 3/10\\ -9/50 & 3/10\end{bmatrix}\begin{bmatrix}9/50 & -3/10\\ 0 & 0\end{bmatrix}\to\begin{cases}x_2=h\\ \frac{9}{50}x_1-\frac{3}{10}h=0\end{cases}$$

$$\begin{cases}x_1=\frac{5}{3}h\\ x_2=h\end{cases}\to h\begin{bmatrix}5/3\\ 1\end{bmatrix}$$

Pertanto $\mathrm{N}\left(\mathbf{A}(9/50)-\mathbf{I}9/10\right)=\left\langle\begin{bmatrix}5/3\\ 1\end{bmatrix}\right\rangle$

5.2.2 L'utilizzo di R e maxima

In:

- R per il calcolo di autovalori e autovettori si impiega la funzione `eigen`. Esso restituisce una lista con autovalori (values) e una matrice (in `vectors`) composta da autovettori.
- maxima:
  - il comando `eigenvalues` di una matrice quadrata restituisce gli autovalori sotto forma di due liste, la prima contiene gli autovalori mentre la secondo le relative molteplicità (algebriche) nell'ordine;
  - il comando `eigenvectors` restituisce due liste: la prima è la lista degli autovalori della matrice e delle relative molteplicità, la seconda è una lista di autovettori (vi è una lista di autovettori per ciascun autovalore, in ogni lista vi potrebbero essere più di un autovettore)

L'esempio in R con $k = 10/100$:

```
## ----------------
## caso k = 10/100
## ----------------
m <- rmatrix(c(3/5, 1/2,
               -10/100, 6/5), nrow = 2)
(res <- eigen(m))

## eigen() decomposition
## $values
## [1] 1.1 0.7
##
## $vectors
##            [,1]       [,2]
## [1,] -0.7071068 -0.9805807
## [2,] -0.7071068 -0.1961161

## gli autovettori sono due soluzioni particolari, ma comunque i
## rapporti tra i primi elementi e i secondi debbono essere 1 e 5,
## come si verifica facilmente
res$vectors[1,]/res$vectors[2,]

## [1] 1 5
```

E in maxima:

```
##
## M:matrix([3/5,1/2],[(-10)/100,6/5])
## eigenvalues(M)
##                                    7   11
##                                  [[--, --], [1, 1]]
##                                    10  10
## eigenvectors(M)
```

```
##                       7   11                    1
##                    [[[--, --], [1, 1]], [[[1, -]], [[1, 1]]]]
##                       10  10                    5
```

L'esempio di $k = 18/100 = 9/50$ in R:

```
## ----------------------
## caso k = 18/100 = 9/50
## ----------------------
m <- rmatrix(c(3/5, 1/2,
               -9/50, 6/5), nrow = 2)
(res <- eigen(m))

## eigen() decomposition
## $values
## [1] 0.9 0.9
##
## $vectors
##           [,1]       [,2]
## [1,] 0.8574929 -0.8574929
## [2,] 0.5144958 -0.5144957

## qui si nota che numericamente R da comunque due autovalori (coincidenti).
## i corrispondenti autovettori seguono
## rapporti tra i primi elementi e i secondi debbono essere 5/3 = 1.667
res$vectors[1,]/res$vectors[2,]

## [1] 1.666667 1.666667
```

E in maxima:

```
##
## M:matrix([3/5,1/2],[(-9)/50,6/5])
## eigenvalues(M)
##                                      9
##                                    [[--], [2]]
##                                      10
## eigenvectors(M)
##                                9                3
##                             [[[--], [2]], [[[1, -]]]]
##                                10               5
```

5.2.3 Diagonalizzabilità e triangolarizzabilità

Sotto quali condizioni una matrice è diagonalizzabile o triangolarizzabile?[2]
Le uguaglianze $\mathbf{A}(k)\mathbf{v}_i = \lambda_i \mathbf{v}_i$ $(i = 1, 2)$ considerate congiuntamente (non come

[2]Qui ci riferiamo ancora al caso riferimento al caso di matrici 2×2 (a titolo di esempio), posto che ovviamente il tutto è generalizzabile a matrici di altre dimensioni.

vettori colonna separati) sono equivalenti, se si pone $\mathbf{S} = [\mathbf{v}_1\ \mathbf{v}_2]$, all'uguaglianza:

$$\mathbf{A}(k)\mathbf{S} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \mathbf{S} \stackrel{(1)}{=} \mathbf{S} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

dove (1) è permessa dal fatto che il risultato non cambia se si pre o post-moltiplica per una matrice diagonale.
A questo punto $\mathbf{A}(k)$ è diagonalizzabile se e solo se $\mathbf{S}$ è invertibile; nel caso lo sia, infatti, si può post-moltiplicare primo e ultimo membro dell'ultima uguaglianza per $\mathbf{S}^{-1}$, ottenendo la fattorizzazione desiderata:

$$\mathbf{A}(k)\mathbf{S} = \mathbf{S} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \iff \mathbf{A}(k) = \mathbf{S} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \mathbf{S}^{-1}$$

Punto fondamentale è che $\mathbf{S}$ è invertibile se e solo se i due vettori $\mathbf{v}_1, \mathbf{v}_2$ sono linearmente indipendenti:

- ciò sicuramente accade se $\lambda_1 \neq \lambda_2$ (per un risultato che vedremo in seguito, il corollario 5.5.5); se $\lambda_1 \neq \lambda_2$ si può quindi sempre diagonalizzare la matrice $\mathbf{A}(k)$ come si è fatto nel primo e nel terzo caso. Per farlo basta, una volta determinata una opportuna $\mathbf{S}$ (scegliendo degli autovalori, ciascuno in riferimento ad un dato autovettore) basta calcolare $\mathbf{S}^{-1}$ e procedere con la post-moltiplicazione

- se invece $\lambda_1 = \lambda_2$ la possibilità di diagonalizzare $\mathbf{A}(k)$ dipende dal fatto di poter trovare due autovettori $\mathbf{v}_1$ e $\mathbf{v}_2$ che siano linearmente indipendenti. Se ciò non è possibile (ad esempio caso di $k = 9/50$ non è possibile trovare autovettori linearmente indipendenti poiché provengono entrambi da $\left\langle \begin{bmatrix} 5/3 \\ 1 \end{bmatrix} \right\rangle$ come visto in esempio 5.2.4), la matrice non si può diagonalizzare.

Tuttavia anche qualora la matrice non sia diagonalizzabile, come vedremo (teorema di triangolarizzazione di Schur e in particolar modo il suo corollario 5.6.6), essa è comunque unitariamente triangolarizzabile, ossia si può fattorizzare come

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1}$$

con $\mathbf{U}^{-1} = \mathbf{U}^H$, da cui

$$\mathbf{T} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \mathbf{U}^H\mathbf{A}\mathbf{U}$$

L'algoritmo per la triangolarizzazione di una matrice quadrata $\mathbf{A}$ di ordine n è il seguente:

1. trovare un suo valore ed un suo autovettore;
2. completare, usando l'autovettore, una base ortogonale di $\mathbb{C}^n$;
3. normalizzare la base trovata e con i vettori normalizzati $\mathbf{u}_i$ costruire la matrice $\mathbf{U} = [\mathbf{u}_1 \ldots \mathbf{u}_n]$
4. calcolare $\mathbf{U}^{-1} = \mathbf{U}^H$
5. calcolare $\mathbf{T} = \mathbf{U}^H\mathbf{A}\mathbf{U}$

6. verificare la triangolarizzazione mediante, appunto,

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1} = \mathbf{U}\mathbf{T}\mathbf{U}^{H}$$

Se infine, in aggiunta, la matrice è reale con autovalori reali, essa è triangolarizzabile con un matrice reale ortogonale.

Esempio 5.2.5. Per il caso $k = 10/100$ la matrice è diagonalizzabile perché $\lambda_1 = 11/10 \neq \lambda_2 = 7/10$. In esempio 5.2.3 abbiamo visto che a $\lambda_1 = 11/10$ sono associati gli autovettori ottenibili da $\left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle$ mentre a $\lambda_2 = 7/10$ quelli ottenibili da da $\left\langle \begin{bmatrix} 5 \\ 1 \end{bmatrix} \right\rangle$. Nella diagonalizzazione iniziale si sono scelti proprio $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ e $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$ andando a formare $\mathbf{S} = \begin{bmatrix} 1 & 5 \\ 1 & 1 \end{bmatrix}$, da cui $\mathbf{S}^{-1} = \begin{bmatrix} -1/4 & 5/4 \\ 1/4 & -1/4 \end{bmatrix}$ Ciò non toglie che si possa diagonalizzare la matrice alternativamente:

- scegliendo altri autovettori

```r
## ad esempio scegliendo [2 2]^T e [5 1]^T come autovettori e tenendo il
## medesimo ordine di autovalori nella diagonale
S <- rmatrix(c(2,5,
               2,1), ncol = 2)
D <- diag(c(11/10, 7/10))
inv_S <- solve(S)
## e vediamo che il prodotto restituisce A(1/10)
S %*% D %*% inv_S

##      [,1] [,2]
## [1,]  0.6  0.5
## [2,] -0.1  1.2
```

- invertendone l'ordine nel formare $\mathbf{S}$, posto che l'ordine dei rispettivi autovalori sulla matrice diagonale venga aggiustato di conseguenza

```r
## ad esempio scegliendo [1 1]^T e [5 1]^T come autovettori ma invertendone
## l'ordine
S <- rmatrix(c(5,1,
               1,1), ncol = 2)
D <- diag(c(7/10, 11/10)) # qui inverto l'ordine anche della diagonale
inv_S <- solve(S)
## il prodotto restituisce sempre A(1/10)
S %*% D %*% inv_S

##      [,1] [,2]
## [1,]  0.6  0.5
## [2,] -0.1  1.2
```

- un combinato disposto delle strategie precedenti

```
S <- rmatrix(c(5,2,
               1,2), ncol = 2)
D <- diag(c(7/10, 11/10))
inv_S <- solve(S)
S %*% D %*% inv_S

##       [,1] [,2]
## [1,]  0.6  0.5
## [2,] -0.1  1.2
```

Questo per dire che le diagonalizzazioni possibili di una matrice diagonalizzabile sono molteplici.

Esempio 5.2.6. Come vedremo anche le triangolarizzazioni possibili sono molteplici. In precedenza si è visto che nel caso di

$$\mathbf{A}(9/50) = \begin{bmatrix} 3/5 & 1/2 \\ -9/50 & 6/5 \end{bmatrix}$$

la matrice ammette due autovalori coincidenti $\lambda_1 = \lambda_2 = \lambda = \frac{9}{10}$ e si era trovato che gli autovalori appartenevano a $\left\langle \begin{bmatrix} 5/3 \\ 1 \end{bmatrix} \right\rangle$. Procediamo innanzitutto a riprodurre la triangolarizzazione proposta nel secondo caso; per comodità di calcolo si era scelto come primo vettore $3 \begin{bmatrix} 5/3 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$. Per completare tale vettore ad una base ortogonale, in questo caso (un unico vettore) è semplice: basta trovare un vettore perpendicolare, ossia tale che

$$\begin{bmatrix} 5 \\ 3 \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix} = 0 \rightarrow 5x + 3y = 0$$

La scelta può ricadere su

$$\begin{bmatrix} -3 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ -5 \end{bmatrix}$$

Nell'esempio è stato scelto $\begin{bmatrix} -3 \\ 5 \end{bmatrix}$,. Andando a normalizzare

$$\left\{ \begin{bmatrix} 5 \\ 3 \end{bmatrix}^T, \begin{bmatrix} -3 \\ 5 \end{bmatrix}, \right\}$$

Per entrambi la norma è $\sqrt{25+9} = \sqrt{34}$ quindi si giunge alla matrice

$$\mathbf{U} = \frac{1}{\sqrt{34}} \begin{bmatrix} 5 & -3 \\ 3 & 5 \end{bmatrix}$$

Calcoliamo l'inversa di $\mathbf{U}$: si ha che $\det \mathbf{U} = \frac{25}{34} + \frac{9}{34} = 1$ per cui

$$\mathbf{U}^{-1} = \frac{1}{1} \cdot \frac{1}{\sqrt{34}} \begin{bmatrix} 5 & 3 \\ -35 & \end{bmatrix} = \mathbf{U}^H = \mathbf{U}^T$$

pertanto notiamo che la matrice è unitaria. Calcoliamo ora **T** mediante maxima (dopo aver svolto i passaggi effettuati) e verifichiamo che l'uguaglianza della matrice di partenza con la triangolarizzazione (`expand` è posto per sicurezza ma qui non servirebbe)

```
##
## A:matrix([3/5,1/2],[(-9)/50,6/5])
##                                   [  3    1 ]
##                                   [  -    - ]
##                                   [  5    2 ]
##                                   [         ]
##                                   [   9   6 ]
##                                   [ - --  - ]
##                                   [   50  5 ]
## U:(1/sqrt(34)) . matrix([5,-3],[3,5])
## determinant(U)
##                                        1
## Uinv:invert(U)
##                            [     5          3     ]
##                            [  --------   -------- ]
##                            [  sqrt(34)   sqrt(34) ]
##                            [                      ]
##                            [      3         5     ]
##                            [ - --------  -------- ]
##                            [   sqrt(34)  sqrt(34) ]
## T:Uinv . A . U
##                                   [ 9   17 ]
##                                   [ --  -- ]
##                                   [ 10  25 ]
##                                   [        ]
##                                   [     9  ]
##                                   [ 0   -- ]
##                                   [     10 ]
## is(equal(A,expand(U . T . Uinv)))
##                                      true
```

Pertanto in questo caso la triangolarizzazione è unitaria in quanto $\mathbf{U}^{-1} = \mathbf{U}^H$. Se viceversa si fosse scelto $\begin{bmatrix} 3 \\ -5 \end{bmatrix}$ come vettore col quale completare **U** si avrebbe avuto la seguente triangolarizzazione, sempre unitaria

```
##
## A:matrix([3/5,1/2],[(-9)/50,6/5])
##                                   [  3    1 ]
##                                   [  -    - ]
##                                   [  5    2 ]
##                                   [         ]
##                                   [   9   6 ]
##                                   [ - --  - ]
##                                   [   50  5 ]
```

```
## U:(1/sqrt(34)) . matrix([5,3],[3,-5])
## determinant(U)
##                                        - 1
## Uinv:invert(U)
##                            [    5          3      ]
##                            [ --------   --------  ]
##                            [ sqrt(34)   sqrt(34)  ]
##                            [                      ]
##                            [    3          5      ]
##                            [ --------  - -------- ]
##                            [ sqrt(34)    sqrt(34) ]
## T:Uinv . A . U
##                                  [ 9     17 ]
##                                  [ --  - -- ]
##                                  [ 10    25 ]
##                                  [          ]
##                                  [      9   ]
##                                  [ 0    --  ]
##                                  [      10  ]
## is(equal(A,expand(U . T . Uinv)))
##                                       true
```

Infine se avessimo sviluppato la base come si fa comunemente:

$$\begin{bmatrix} 5 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 15 & 3 & 0 \\ 15 & 0 & 5 \end{bmatrix} \begin{bmatrix} 15 & 3 & 0 \\ 0 & -3 & 5 \end{bmatrix}$$

quindi i vettori dominanti sono i primi 2 quindi aggiungiamo $\mathbf{e}_1$ (già normalizzato) all'autovettore per generare

$$\mathbf{U} = \begin{bmatrix} 5/\sqrt{34} & 1 \\ 3/\sqrt{34} & 0 \end{bmatrix}$$

Ora per il calcolo dell'inversa

$$\mathbf{U}^{-1} = -\frac{\sqrt{34}}{3} \begin{bmatrix} 0 & -1 \\ -3/\sqrt{34} & 5/\sqrt{34} \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{34}/3 \\ +1 & -5/3 \end{bmatrix}$$

quindi si nota che, il determinante non è ± 1, ne non vi è quel minimo di simmetria (almeno a livello di valori assoluti) necessario nella matrice di partenza (sulla diagonale principale e secondari), e non vale il fatto $\mathbf{U}^{-1} = \mathbf{U}^H$, pertanto la matrice $\mathbf{U}$ non è unitaria. Provimo comunque a calcolare $\mathbf{T}$ (mediante maxima) e a verificare se vale la triangolarizzazione

```
##
## A:matrix([3/5,1/2],[(-9)/50,6/5])
##                                   [  3    1 ]
##                                   [  -    - ]
##                                   [  5    2 ]
##                                   [         ]
##                                   [   9   6 ]
```

```
##                                  [ - --  - ]
##                                  [   50  5 ]
## U:matrix([5/sqrt(34),1],[3/sqrt(34),0])
## determinant(U)
##                                       3
##                                  - --------
##                                    sqrt(34)
## Uinv:invert(U)
##                                [    sqrt(34) ]
##                                [ 0  -------- ]
##                                [       3     ]
##                                [             ]
##                                [       5     ]
##                                [ 1    - -    ]
##                                [       3     ]
## is(equal(Uinv,conjugate(transpose(U))))
##                                      false
## T:Uinv . A . U
##                             [ 9     3 sqrt(34) ]
##                             [ --  - ---------- ]
##                             [ 10        50     ]
##                             [                  ]
##                             [         9        ]
##                             [ 0       --       ]
##                             [         10       ]
## is(equal(A,expand(U . T . Uinv)))
##                                     true
```

Quindi si giunge comunque a una triangolarizzazione con $\mathbf{T} = \begin{bmatrix} 9/10 & -\frac{3\sqrt{34}}{50} \\ 0 & 9/10 \end{bmatrix}$, anche se non unitaria.

Osservazione 255. In generale in tutti gli esempi di triangolarizzazione forniti, il prof per trovare la base ortonormale inverte gli elementi dell'autovettore e aggiunge un segno meno in uno dei due elementi (ai fini di pervenire ad una triangolarizzazione unitaria).

5.3 Generalità su autovalori e autospazi

5.3.1 Introduzione

Osservazione 256. Se $\mathbf{A}$ è una matrice $n \times n$ a coefficienti in $\mathbb{K}$ ($\mathbb{R}$ o $\mathbb{C}$) e $f_{\mathbf{A}} : \mathbb{K}^n \to \mathbb{K}^n$ l'applicazione lineare indotta da $\mathbf{A}$: $f_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{v}$ per ogni $\mathbf{v} \in \mathbb{K}^n$, è chiaro che maggiore è il numero dei coefficienti nulli di $\mathbf{A}$, minore è la difficoltà di calcolo di $f_{\mathbf{A}}(\mathbf{v})$. In particolare:

- se $\mathbf{A} = \mathbf{D} = \mathbf{Diag}(d_1, \ldots, d_n)$ allora

$$f_{\mathbf{D}}\left(\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}\right) = \begin{bmatrix} d_1 v_1 \\ \vdots \\ d_n v_n \end{bmatrix}$$

per ogni $\mathbf{v} \in \mathbb{K}^n$

- se $\mathbf{A} = \lambda\mathbf{I}_n$ allora $f_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{v} = \lambda\mathbf{I}_n\mathbf{v} = \lambda\mathbf{v}$;
- anche nel caso in cui $\mathbf{A}$ non sia scalare (es alla $\lambda\mathbf{I}_n$), come vedremo, oltre a $\mathbf{0}$ possono esistere (ed esistono certamente se $\mathbb{K} = \mathbb{C}$) vettori $\mathbf{v}_i \neq \mathbf{0}$ di $\mathbb{K}^n$ (gli autovettori) tali che $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$ con λ_i un opportuno scalare (gli autovalori).

Osservazione 257. Svilupperemo una teoria che ci permetterà di trovare le condizioni che $\mathbf{A}$ deve avere affinché esista una base $\mathscr{B}$ di $\mathbb{K}^n$ costituita da vettori di quest'ultimo tipo (autovettori).

Osservazione 258. Nel caso in cui ciò avvenga, la matrice associata ad $f_{\mathbf{A}}$ rispetto a $\mathscr{B}$ è una matrice diagonale $\mathbf{D}$, per cui $f_{\mathbf{A}}$ ammette una fattorizzazione del tipo $f_{\mathbf{A}} = s \circ f_{\mathbf{D}} \circ s^{-1}$ per un opportuno isomorfismo s di $\mathbb{K}^n$ in se, e dato che

$$\mathbf{A}\mathbf{v} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}\mathbf{v} \iff \mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$$

la matrice $\mathbf{A}$ ammette una fattorizzazione del tipo $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$ dove $\mathbf{S}$ è una matrice invertibile.

Osservazione 259. Cominciamo con il risolvere i due seguenti problemi:

1. stabilire per quali $\lambda \in \mathbb{K}$ esistono sottospazi non nulli U di $\mathbb{K}^n$ tali che la restrizione di $f_{\mathbf{A}}$ a U coincide con la moltiplicazione per λ
2. in corrispondenza di ciascun di tali $\lambda \in \mathbb{K}$ trovare il massimo sottospazio di $\mathbb{K}^n$ su cui $f_{\mathbf{A}}$ coincide con la moltiplicazione per λ

Osservazione 260. Osserviamo innanzitutto che se U è un sottospazio di $\mathbb{K}^n$ su cui la restrizione di $f_{\mathbf{A}}$ coincide con la moltiplicazione per $\lambda \in \mathbb{K}$, allora

$$f_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\mathbf{u} = \lambda\mathbf{u} \in U, \quad \forall \mathbf{u} \in U$$

ossia il risultato appartiene ancora a U (stiamo moltiplicando per una costante un elemento di uno spazio vettoriale, per cui il risultato appartiene ancora allo spazio vettoriale dato per definizione).
Più in generale si pone la seguente definizione.

Definizione 5.3.1 (Sottospazio $\mathbf{A}$-invariante)**.** Sia $\mathbf{A}$ una matrice $n \times n$ a coefficienti in $\mathbb{K}$ e U un sottospazio di $\mathbb{K}^n$; U si dice $\mathbf{A}$-invariante (o anche $f_{\mathbf{A}}$-invariante) se

$$f_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\mathbf{u} \in U, \quad \mathbf{u} \in U$$

Esempio 5.3.1. $\{\mathbf{0}\}$ e $\mathbb{K}^n$ sono sottospazi $\mathbf{A}$-invarianti

Esempio 5.3.2. Lo spazio nullo $\mathrm{N}(\mathbf{A})$ e lo spazio delle colonne $\mathrm{C}(\mathbf{A})$ sono sottospazi $\mathbf{A}$ invarianti.

Esempio 5.3.3. La somma e l'intersezione di sottospazi $\mathbf{A}$-invarianti di $\mathbb{K}^n$ sono sottospazi $\mathbf{A}$ invarianti. Infatti (tentativo di dimostrazione) se U, V sono spazi $\mathbf{A}$ invarianti, ossia

$$f_{\mathbf{A}}(\mathbf{u}) = \mathbf{A}\mathbf{u} \in U, \quad \forall \mathbf{u} \in U$$
$$f_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{v} \in V, \quad \forall \mathbf{v} \in V$$

allora si ha che per $\mathbf{u}+\mathbf{v} \in U+V$

$$f_{\mathbf{A}}(\mathbf{u}+\mathbf{v}) \overset{(1)}{=} f_{\mathbf{A}}(\mathbf{u})+f_{\mathbf{A}}(\mathbf{v}) = \mathbf{A}\mathbf{u}+\mathbf{A}\mathbf{v} \in U+V, \quad \forall \mathbf{u}+\mathbf{v} \in U+V$$

dove (1) deriva dalla linearità della funzione. Per l'intersezione se $\mathbf{x} \in U \cap V$ con U, V $\mathbf{A}$-invarianti si ha che

$$f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} \in U$$
$$f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} \in V$$

ma allora

$$f_{\mathbf{A}}(\mathbf{x}) = \mathbf{A}\mathbf{x} \in U \cap V, \quad \forall \mathbf{x} \in U \cap V$$

Osservazione 261. Si noti come uno scalare $\lambda \in \mathbb{K}$ è soluzione del problema (1) se e solo se esiste un sottospazio *unidimensionale* U di $\mathbb{K}^n$ tale che la restrizione di $f_{\mathbf{A}}$ a U coincida con la moltiplicazione per λ, e quindi se e solo se esiste $\mathbf{u} \in \mathbb{K}^n$, $\mathbf{u} \neq \mathbf{0}$ tale che $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$. Si pone allora la seguente definizione

5.3.2 Autovalori

Definizione 5.3.2 (Autovalore della matrice $\mathbf{A}$)**.** Scalare $\lambda \in \mathbb{K}$ per il quale esista un vettore $\mathbf{v} \in \mathbb{K}^n$, $\mathbf{v} \neq \mathbf{0}$ tale che $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

Osservazione 262. È fondamentale ricordare che il vettore $\mathbf{v}$ di cui si richiede l'esistenza nella precedente definizione deve essere non nullo: infatti l'uguaglianza $\mathbf{A}\mathbf{0} = \lambda\mathbf{0}$ è verificata per ogni $\lambda \in \mathbb{K}$

Osservazione 263. Dunque gli autovalori della matrice $\mathbf{A}$ sono la soluzione del problema (1).

Definizione 5.3.3 (Spettro)**.** L'insieme degli autovalori distinti di $\mathbf{A}$ si dice lo spettro di $\mathbf{A}$.

Esempio 5.3.4 (Autovalori di una matrice di proiezione)**.** Sia $\mathbf{P}$ una matrice di proiezione ($\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^H$) complessa (per cui $\mathbb{K} = \mathbb{C}$) di ordine n. Nella dimostrazione del teorema 3.6.1 abbiamo visto che $\mathbb{C}^n = \mathrm{N}(\mathbf{P}) \oplus \mathrm{C}(\mathbf{P})$ e che la restrizione di $f_{\mathbf{P}}$ a $\mathrm{C}(\mathbf{P})$ è l'identità (se pre-moltiplichiamo per $\mathbf{P}$ un elemento dello spazio delle colonne di $\mathbf{P}$ otteniamo come risultato l'elemento stesso). Vediamo alcuni casi:

1. se $\mathbf{P} = \mathbb{O}$ allora
$$\mathbf{P}\mathbf{v} = \mathbf{0}, \quad \forall \mathbf{v} \in \mathbb{C}^n$$
ossia $\mathbb{C}^n = \mathrm{N}(\mathbf{P})$; in questo caso
$$\mathbf{P}\mathbf{v} = \mathbf{0} = \lambda\mathbf{v} \iff \lambda = 0, \quad \forall \mathbf{v} \in \mathbb{C}^n$$
l'unico autovalore di $\mathbf{P}$ è $\lambda = 0$. Pertanto lo spettro di $\mathbf{P}$ è $\{0\}$.

2. se $\mathbf{P}$ è invertibile allora $n = \operatorname{rk}\mathbf{P} = \dim \mathrm{C}(\mathbf{P}) = \dim \mathbb{C}^n = n$ da cui $\mathbb{C}^n = \mathrm{C}(\mathbf{P})$ e l'unico autovalore di $\mathbf{P}$ è $\lambda = 1$; infatti essendo $\mathbf{P}$ di proiezione, dalla dimostrazione del 3.6.1 si ha che la pre-moltiplicazione per $\mathbf{P}$ se ristretta a $\mathrm{C}(\mathbf{P}) = \mathbb{C}^n$ è l'identità, ossia
$$\mathbf{P}\mathbf{v} = \mathbf{v}, \quad \mathbf{v} \in \mathbb{C}^n$$

per cui
$$\mathbf{P}\mathbf{v} = \mathbf{v} = \lambda\mathbf{v} \iff \lambda = 1$$
pertanto lo spettro di $\mathbf{P}$ è $\{1\}$.

3. nei rimanenti casi si ha che $\mathrm{N}(\mathbf{P}) \neq \{\mathbf{0}\}$ e $\mathrm{C}(\mathbf{P}) \neq \{\mathbf{0}\}$ e sia 0 che 1 sono autovalori di $\mathbf{P}$: 0 lo è perché se pre-moltiplichiamo per $\mathbf{P}$ un elemento $\mathbf{v} \in \mathrm{N}(\mathbf{P})$
$$\mathbf{P}\mathbf{v} = \mathbf{0} = 0\mathbf{v} \implies \lambda = 0$$
e specularmente avviene se pre-moltiplichiamo per un elemento di $\mathrm{C}(\mathbf{P})$. Mostriamo che $\mathbf{P}$ non ha altri autovalori, ossia che lo spettro di $\mathbf{P}$ è $\{0, 1\}$: siano $\lambda \in \mathbb{C}$ un autovalore di $\mathbf{P}$ e $\mathbf{0} \neq \mathbf{v} \in \mathbb{C}^n$ un vettore tale che $\mathbf{P}\mathbf{v} = \lambda\mathbf{v}$. Da $\mathbf{P} = \mathbf{P}^2$ segue che
$$\lambda\mathbf{v} = \mathbf{P}\mathbf{v} = \mathbf{P}^2\mathbf{v} = \mathbf{P}\mathbf{P}\mathbf{v} = \mathbf{P}(\lambda\mathbf{v}) = \lambda\mathbf{P}\mathbf{v} = \lambda^2\mathbf{v}$$
per cui $(\lambda - \lambda^2)\mathbf{v} = \mathbf{0}$. Dato che $\mathbf{v} \neq \mathbf{0}$ allora deve essere $\lambda - \lambda^2 = 0$ ossia $\lambda \in \{0, 1\}$.

Osservazione 264. Il seguente esempio mostra che il problema (1) può non avere soluzioni se $\mathbb{K} = \mathbb{R}$.

Esempio 5.3.5. Si consideri la matrice reale 2×2
$$\mathbf{A} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$
e si osservi che $\mathbf{A}^2 = -\mathbf{I}_2$. Se $\mathbf{A}$ avesse un autovalore reale λ, esisterebbe un vettore $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^2$ tale che $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$. Pre-moltiplicando per $\mathbf{A}$ si otterrebbe
$$\mathbf{A}\mathbf{A}\mathbf{v} = \mathbf{A}\lambda\mathbf{v} \iff -\mathbf{I}\mathbf{v} = \mathbf{A}\mathbf{v}\lambda \iff -\mathbf{v} = \lambda\mathbf{v}\lambda \iff -\mathbf{v} = \lambda^2\mathbf{v}$$
e quindi, essendo $\mathbf{v} \neq \mathbf{0}$, l'uguaglianza sarebbe verificata solo se $\lambda^2 = -1$ che non è possibile per alcun $\lambda \in \mathbb{R}$

5.3.3 Autovettori e autospazi

Definizione 5.3.4 (Autovettore di $\mathbf{A}$ relativo a λ)**.** Dato $\lambda \in \mathbb{K}$ autovalore di $\mathbf{A}$, ogni vettore $\mathbf{0} \neq \mathbf{v} \in \mathbb{K}^n$ per cui $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

Definizione 5.3.5 (Autospazio di $\mathbf{A}$ relativo a λ, $E_{\mathbf{A}}(\lambda)$)**.** Dato $\lambda \in \mathbb{K}$ autovalore di $\mathbf{A}$, il massimo sottospazio di $\mathbb{K}^n$ su cui $f_{\mathbf{A}}$ equivale alla moltiplicazione per λ:
$$E_{\mathbf{A}}(\lambda) = \{\mathbf{v} \in \mathbb{K}^n : \mathbf{A}\mathbf{v} = \lambda\mathbf{v}\}$$

Osservazione 265. Per ogni autovalore λ di $\mathbf{A}$, l'autospazio $E_{\mathbf{A}}(\lambda)$ è costituito dal vettore nullo e da tutti gli autovettori di $\mathbf{A}$ relativi a λ.

Proposizione 5.3.1. *$E_{\mathbf{A}}(\lambda)$ è effettivamente un sottospazio di $\mathbb{K}^n$*

Dimostrazione. Innanzitutto $\mathbf{0} \in E_{\mathbf{A}}(\lambda)$; nell'ipotesi poi che $\mathbf{v}, \mathbf{w} \in E_{\mathbf{A}}(\lambda)$ andiamo a verificare che $c\mathbf{v}, \mathbf{v} + \mathbf{w} \in E_A(\lambda)$ con $c \in \mathbb{K}$.
Partiamo dal verificare che $\mathbf{v} + \mathbf{w} \in \mathbb{K}^n$; dato che $\mathbf{v}, \mathbf{w} \in E_{\mathbf{A}}(\lambda)$ si avrà che
$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$
$$\mathbf{A}\mathbf{w} = \lambda\mathbf{w}$$

Si conclude che $\mathbf{v}+\mathbf{w} \in E_{\mathbf{A}}(\lambda)$, se (applicando la definizione)

$$\mathbf{A}(\mathbf{v}+\mathbf{w}) = \lambda(\mathbf{v}+\mathbf{w}) \iff \mathbf{A}\mathbf{v}+\mathbf{A}\mathbf{w} = \lambda\mathbf{v}+\lambda\mathbf{w}$$

e l'uguaglianza è verificata se, appunto, $\mathbf{v}, \mathbf{w} \in E_{\mathbf{A}}(\lambda)$.
Per quanto riguarda $c\mathbf{v}$ invece, si avrà che $c\mathbf{v} \in E_{\mathbf{A}}(\lambda)$ se

$$\mathbf{A}(c\mathbf{v}) = \lambda(c\mathbf{v}) \iff c\mathbf{A}\mathbf{v} = c\lambda\mathbf{v} \iff \mathbf{A}\mathbf{v} = \lambda\mathbf{v}$$

e l'equazione è verificata dato che $\mathbf{v} \in E_{\mathbf{A}}(\lambda)$. □

Osservazione 266. Gli autospazi di $\mathbf{A}$ sono le soluzioni del problema (2).

Definizione 5.3.6 (Molteplicità geometrica di λ, $d(\lambda)$)**.** Dimensione dello spazio vettoriale $E_{\mathbf{A}}(\lambda)$

Proposizione 5.3.2. *La molteplicità geometrica $d(\lambda)$ coincide con la nullità (dimensione dello spazio nullo) della matrice $\mathbf{A}-\lambda\mathbf{I}$*

Dimostrazione. Infatti osserviamo che

$$\begin{aligned} E_{\mathbf{A}}(\lambda) &= \{\mathbf{v} \in \mathbb{K}^n : \mathbf{A}\mathbf{v} = \lambda\mathbf{v}\} = \{\mathbf{v} \in \mathbb{K}^n : \mathbf{A}\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}\} \\ &= \{\mathbf{v} \in \mathbb{K}^n : \mathbf{A}\mathbf{v} - \lambda\mathbf{I}\mathbf{v} = \mathbf{0}\} = \{\mathbf{v} \in \mathbb{K}^n : (\mathbf{A}-\lambda\mathbf{I})\mathbf{v} = \mathbf{0}\} \\ &= \mathrm{N}\left(\mathbf{A}-\lambda\mathbf{I}\right) \end{aligned} \tag{5.2}$$

dalla quale si conclude prendendo la dimensione. □

Definizione 5.3.7 (Autosistema di $\mathbf{A}$)**.** L'insieme degli autovalori di $\mathbf{A}$ e dei corrispondenti autospazi

Esempio 5.3.6. Tornando all'esempio 5.3.4 si ha che:

- se $\mathbf{P} = \mathbb{O}$, ogni vettore di $\mathbb{C}^n \setminus \{\mathbf{0}\}$ è un autovettore di $\mathbf{P}$ relativo all'autovalore 0 in quanto
$$\mathbb{O}\mathbf{v} = 0\mathbf{v}, \quad \forall \mathbf{v} \in \mathbb{C}^n$$
(da $\mathbb{C}^n$ escludiamo lo $\mathbf{0}$ perché per definizione un autovettore non può essere nullo); si ha che l'autospazio
$$E_{\mathbf{P}}(0) \overset{(1)}{=} \mathrm{N}\left(\mathbf{P}-0\mathbf{I}\right) = \mathrm{N}\left(\mathbf{P}\right) \overset{(2)}{=} \mathbb{C}^n$$
con (1) giustificato dalla 5.2 e (2) da quanto visto nell'esempio 5.3.4;

- se $\mathbf{P}$ è invertibile, ogni vettore di $\mathbb{C}^n \setminus \{\mathbf{0}\}$ è un autovettore di $\mathbf{P}$ relativo all'autovalore 1 proprio perché
$$\mathbf{P}\mathbf{v} = 1\mathbf{v}, \quad \forall \mathbf{v} \in \mathbb{C}^n$$
ed
$$E_{\mathbf{P}}(1) \overset{(1)}{=} (\mathbb{C}^n \setminus \{\mathbf{0}\}) \cup \{\mathbf{0}\} = \mathbb{C}^n$$
con (1) per l'osservazione 265

- nei rimanenti casi $\mathrm{N}(\mathbf{P}) \neq \{\mathbf{0}\} \neq \mathrm{C}(\mathbf{P})$, ogni vettore di $\mathrm{N}(\mathbf{P}) \setminus \{\mathbf{0}\}$ è un autovettore di $\mathbf{P}$ relativo all'autovalore 0 e ogni vettore di $\mathrm{C}(\mathbf{P}) \setminus \{\mathbf{0}\}$ è un autovettore di $\mathbf{P}$ relativo all'autovalore 1. Pertanto:

$$E_{\mathbf{P}}(0) = \{\mathrm{N}(\mathbf{P}) \setminus \{\mathbf{0}\}\} \cup \{\mathbf{0}\} = \mathrm{N}(\mathbf{P})$$
$$E_{\mathbf{P}}(1) = \{\mathrm{C}(\mathbf{P}) \setminus \{\mathbf{0}\}\} \cup \{\mathbf{0}\} = \mathrm{C}(\mathbf{P})$$

Infine, si noti che se 1 è autovalore di $\mathbf{P}$, la sua molteplicità geometrica è uguale a $\operatorname{rk}\mathbf{P}$: infatti sia che $\mathbf{P}$ sia invertibile o meno, in ogni caso si ha che

$$E_{\mathbf{P}}(1) = \mathrm{C}(\mathbf{P})$$

da cui prendendo la dimensione

$$\dim E_{\mathbf{P}}(1) = \dim \mathrm{C}(\mathbf{P}) = \operatorname{rk}\mathbf{P}$$

Osservazione 267. Abbiamo visto che se $\mathbf{P}$ è matrice di permutazione (idempotente ed hermitiana) di ordine n allora $\mathbb{C}^n$ si decompone in somma diretta di autospazi di $\mathbf{P}$:

$$\mathbb{C}^n = \mathrm{N}(\mathbf{P}) \oplus \mathrm{C}(\mathbf{P}) = E_{\mathbf{P}}(0) \oplus E_{\mathbf{P}}(1)$$

Il seguente esempio mostra che esistono matrici $n \times n$ (non di permutazione) per cui $\mathbb{C}^n$ non è somma diretta di loro autospazi (quindi non è una caratteristica valida per tutte le matrici).

Esempio 5.3.7. Sia

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Se λ è un autovalore di $\mathbf{B}$ e $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ un autovettore di $\mathbf{B}$ relativo a λ, allora deve essere

$$\overbrace{\begin{bmatrix} x+y \\ y \end{bmatrix}}^{\mathbf{Bv}} = \overbrace{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}}^{\mathbf{B}} \overbrace{\begin{bmatrix} x \\ y \end{bmatrix}}^{\mathbf{v}} = \lambda \overbrace{\begin{bmatrix} x \\ y \end{bmatrix}}^{\mathbf{v}} = \overbrace{\begin{bmatrix} \lambda x \\ \lambda y \end{bmatrix}}^{\lambda\mathbf{v}}$$

Ora affinché sia

$$\overbrace{\begin{bmatrix} x+y \\ y \end{bmatrix}}^{\mathbf{Bv}} = \overbrace{\begin{bmatrix} \lambda x \\ \lambda y \end{bmatrix}}^{\lambda\mathbf{v}}$$

l'unico autovalore possibile di $\mathbf{B}$ è $\lambda = 1$ e deve essere

$$\begin{cases} x + y = x \\ y = y \end{cases} \quad \begin{cases} y = 0 \end{cases} \quad \langle \mathbf{e}_1 \rangle$$

pertanto dato che deve esser $v \neq \mathbf{0}$, se $y = 0$ si deve avere $x \neq 0$. Ossia $\mathbf{v}$ è del tipo $\begin{bmatrix} x \\ 0 \end{bmatrix}$ con $x \neq 0$, gli autovettori sono $\langle \mathbf{e}_1 \rangle \setminus \{\mathbf{0}\}$ ed $E_{\mathbf{B}}(\lambda) = \langle \mathbf{e}_1 \rangle$.

Pertanto $\mathbb{C}^2$ non è somma degli autospazi di $\mathbf{B}$ (sarebbe necessario un altro autovalore λ_2 tale che $E_{\mathbf{B}}(\lambda_2) = \langle \mathbf{e}_2 \rangle$; allora sì che sarebbe $\mathbb{C}^n = E_{\mathbf{B}}(\lambda_1) \oplus E_{\mathbf{B}}(\lambda_2) = \langle \mathbf{e}_1 \rangle \oplus \langle \mathbf{e}_2 \rangle$).

Osservazione 268. Vedremo che anche solo il fatto che una matrice $n \times n$ sia hermitiana (non necessariamente di permutazione) è sufficiente affinché $\mathbb{C}^n$ sia somma diretta di suoi autospazi. Come mostra il seguente esempio, tuttavia, anche tale condizione non è strettamente necessaria.

Esempio 5.3.8. Sia

$$\mathbf{C} = \begin{bmatrix} \lambda & 0 \\ \lambda & 0 \end{bmatrix}$$

con $\lambda \in \mathbb{C} \setminus \{0\}$. Chiaramente $\mathbf{C} \neq \mathbf{C}^H$, quindi la matrice non è hermitiana. Da:

$$\mathbf{C}\mathbf{e}_2 = \mathbf{0} = 0\mathbf{e}_2$$

segue che 0 è autovalore ed $\mathbf{e}_2$ autovettore. Similmente da:

$$\mathbf{C}(\mathbf{e}_1 + \mathbf{e}_2) = \begin{bmatrix} \lambda \\ \lambda \end{bmatrix} = \lambda(\mathbf{e}_1 + \mathbf{e}_2)$$

segue che λ è autovalore per l'autovettore $\mathbf{e}_1 + \mathbf{e}_2$.
In generale si ha che

$$E_{\mathbf{C}}(0) = \langle \mathbf{e}_2 \rangle, \quad E_{\mathbf{C}}(\lambda) = \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle$$

dunque $E_{\mathbf{C}}(0) \oplus E_{\mathbb{C}}(\lambda) = \mathbb{C}^n$; ossia $\mathbb{C}^2$ è somma diretta di autospazi di $\mathbf{C}$ pur non essendo quest'ultima hermitiana.

5.3.4 Polinomio caratteristico, autovalori e autospazi

Definizione 5.3.8 (Polinomio caratteristico di $\mathbf{A}$)**.**

$$p_{\mathbf{A}}(X) = \det(\mathbf{A} - X\mathbf{I}_n)$$

nell'indeterminata X

Definizione 5.3.9 (Equazione caratteristica di $\mathbf{A}$)**.** L'equazione

$$\det(\mathbf{A} - X\mathbf{I}_n) = 0 \tag{5.3}$$

nell'incognita X.

In maxima il polinomio caratteristico si ottiene calcolando direttamente $p_{\mathbf{A}}(X) = \det(\mathbf{A} - X\mathbf{I}_n)$ oppure con il comando `charpoly`:

```
##
## A:matrix([1,2],[3,4])
## determinant(A-X*ident(2))
##                                  (1 - X) (4 - X) - 6
## expand(charpoly(A,X))
##                                      2
##                                     X  - 5 X - 2
```

Proposizione 5.3.3. *Se λ è un autovalore di $\mathbf{A}$ allora l'autospazio di $\mathbf{A}$ relativo a λ risulta*

$$E_{\mathbf{A}}(\lambda) = \mathrm{N}\left(\mathbf{A} - \lambda\mathbf{I}_n\right) \neq \{\mathbf{0}\}$$

Dimostrazione. Fissato $\lambda \in \mathbb{C}$ si può sempre definire il sottospazio $E_{\mathbf{A}}(\lambda)$ di $\mathbb{C}^n$, anche se λ non è un autovalore di $\mathbf{A}$ (al massimo coincide con $\{\mathbf{0}\}$). Per la definizione di autovalore si ha che

$$\lambda \text{ è un autovalore} \overset{(1)}{\iff} \{\mathbf{0}\} \neq E_{\mathbf{A}}(\lambda) \overset{(2)}{=} \mathrm{N}\left(\mathbf{A} - \lambda \mathbf{I}_n\right) \tag{5.4}$$

dove in (1) diciamo che è autovalore se $E_{\mathbf{A}}(\lambda)$ non ha solo $\mathbf{0}$ e (2) deriva da 5.2. □

Proposizione 5.3.4. *Uno scalare $\lambda \in \mathbb{K}$ è un autovalore di $\mathbf{A}$ se e solo se λ è soluzione dell'equazione caratteristica di $\mathbf{A}$*

Dimostrazione. Si ha:

$$\begin{aligned} \lambda \text{ è un autovalore} &\iff \mathrm{N}\left(\mathbf{A} - \lambda \mathbf{I}_n\right) \neq \{\mathbf{0}\} \overset{(1)}{\iff} \mathbf{A} - \lambda \mathbf{I}_n \text{ non è invertibile} \\ &\overset{(2)}{\iff} \det(\mathbf{A} - \lambda \mathbf{I}_n) = 0 \end{aligned}$$

in (1) se invertibile l'unica soluzione ad $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$ è proprio $\mathbf{0}$ (teorema 1.6.12) e (2) è dato da teorema 4.3.3.
Dunque gli autovalori di una matrice $\mathbf{A}$ sono tutte e sole le soluzioni dell'equazione caratteristica o equivalentemente, tutte le radici del polinomio caratteristico nell'indeterminata X. □

Esempio 5.3.9. Il polinomio caratteristico della matrice

$$\mathbf{B} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

dell'esempio 5.3.7 è

$$\begin{aligned} p_{\mathbf{B}}(X) &= \det\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - X\mathbf{I}_2\right) = \det\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix}\right) \\ &= \det\left(\begin{bmatrix} 1-X & 1 \\ 0 & 1-X \end{bmatrix}\right) = (1-X)^2 \end{aligned}$$

la sua equazione caratteristica è $(1-X)^2 = 0$ e quindi, come avevamo già osservato, la matrice $\mathbf{B}$ ha come unico autovalore 1 (contato due volte, dal momento che l'equazione $(1-X)^2 = 0$ è di secondo grado). Inoltre l'autospazio di $\mathbf{B}$ relativo a 1 è

$$E_{\mathbf{B}}(1) = \mathrm{N}\left(\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} - 1 \cdot \mathbf{I}_2\right) = \mathrm{N}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \langle \mathbf{e}_1 \rangle$$

Esempio 5.3.10. Il polinomio caratteristico della matrice

$$\mathbf{C} = \begin{bmatrix} \lambda & 0 \\ \lambda & 0 \end{bmatrix}$$

dell'esempio 5.3.8 dove $\lambda \in \mathbb{C} \setminus \{0\}$ è:

$$\begin{aligned} p_{\mathbf{C}}(X) &= \det\left(\begin{bmatrix} \lambda & 0 \\ \lambda & 0 \end{bmatrix} - X\mathbf{I}_2\right) = \det\left(\begin{bmatrix} \lambda & 0 \\ \lambda & 0 \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & X \end{bmatrix}\right) \\ &= \det\begin{bmatrix} \lambda - X & 0 \\ \lambda & -X \end{bmatrix} = -X(\lambda - X) \end{aligned}$$

la sua equazione caratteristica è $-X(\lambda - X) = 0$, e quindi, come avevamo già osservato, la matrice $\mathbf{C}$ ha come autovalori i numeri 0 e λ. Gli autospazi di $\mathbf{C}$ relativi a 0 e λ sono

$$E_{\mathbf{C}}(0) = \mathrm{N}(\mathbf{C}) = \mathrm{N}\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\right) = \langle \mathbf{e}_2 \rangle$$
$$E_{\mathbf{C}}(\lambda) = \mathrm{N}(\mathbf{C} - \lambda\mathbf{I}_2) = \mathrm{N}\left(\begin{bmatrix}0 & 0\\ \lambda & -\lambda\end{bmatrix}\right) = \mathrm{N}\left(\begin{bmatrix}1 & -1\\ 0 & 0\end{bmatrix}\right) = \langle \mathbf{e}_1 + \mathbf{e}_2 \rangle$$

Esempio 5.3.11. Il polinomio caratteristico della matrice

$$\mathbf{A} = \begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}$$

considerata nell'esempio 5.3.5 è

$$p_{\mathbf{A}}(X) = \det\left(\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} - X\mathbf{I}_2\right) = \det\left(\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} - \begin{bmatrix}x & 0\\ 0 & x\end{bmatrix}\right)$$
$$= \det\begin{bmatrix}-X & -1\\ 1 & -X\end{bmatrix} = X^2 + 1$$

Poiché $X^2 + 1$ non ha radici reali, $\mathbf{A}$ non ha autovalori reali. Ha però due autovalori complessi distinti: $\lambda_1 = i$ e $\lambda_2 = -i$.
Determiniamo ora $E_{\mathbf{A}}(i)$ e $E_{\mathbf{A}}(-i)$:

$$E_{\mathbf{A}}(i) = \mathrm{N}(\mathbf{A} - i\mathbf{I}) = \mathrm{N}\left(\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix} - \begin{bmatrix}i & 0\\ 0 & i\end{bmatrix}\right) = \mathrm{N}\left(\begin{bmatrix}-i & -1\\ 1 & -i\end{bmatrix}\right)$$
$$E_{\mathbf{A}}(-i) = \mathrm{N}(\mathbf{A} + i\mathbf{I}) = \mathrm{N}\left(\begin{bmatrix}i & -1\\ 1 & i\end{bmatrix}\right)$$

Per la determinazione dello spazio nullo, nel caso $\lambda = i$

$$\begin{bmatrix}-i & -1\\ 1 & -i\end{bmatrix}\begin{bmatrix}-i^2 & -i\\ 1 & -i\end{bmatrix}\begin{bmatrix}1 & -i\\ 1 & -i\end{bmatrix}\begin{bmatrix}1 & -i\\ 0 & 0\end{bmatrix} \to \begin{cases}x_2 = h\\ x_1 - ih = 0\end{cases} \quad \begin{cases}x_1 = ih\\ x_2 = h\end{cases} \quad \to h\begin{bmatrix}i\\ 1\end{bmatrix}$$

dunque $E_{\mathbf{A}}(i) = \left\langle \begin{bmatrix}i\\ i\end{bmatrix} \right\rangle$.
Nel caso $\lambda = -i$ specularmente

$$\begin{bmatrix}i & -1\\ 1 & i\end{bmatrix}\begin{bmatrix}i^2 & -i\\ 1 & i\end{bmatrix}\begin{bmatrix}-1 & -i\\ 1 & i\end{bmatrix}\begin{bmatrix}0 & 0\\ 1 & i\end{bmatrix} \to \begin{cases}x_2 = h\\ x_1 + i \cdot x_2 = 0\end{cases} \quad \begin{cases}x_2 = h\\ x_1 = -hi\end{cases} \quad \to h\begin{bmatrix}-i\\ 1\end{bmatrix}$$

quindi $E_{\mathbf{A}}(-i) = \left\langle \begin{bmatrix}-i\\ 1\end{bmatrix} \right\rangle$.
Verifichiamo con R i risultati:

```
m <- rmatrix(c(0, -1,
               1, 0), ncol = 2)
(res <- eigen(m))
```

```
## eigen() decomposition
## $values
## [1] 0+1i 0-1i
##
## $vectors
##                       [,1]                  [,2]
## [1,] 0.7071068+0.0000000i 0.7071068+0.0000000i
## [2,] 0.0000000-0.7071068i 0.0000000+0.7071068i
```

```
res$vectors[1, ] / res$vectors[2, ]
```

```
## [1] 0+1i 0-1i
```

5.3.5 Alcune considerazioni pratiche

Osservazione 269. Dal punto di vista teorico le proposizioni 5.3.3 e 5.3.4 risolvono entrambi i problemi posti all'inizio del paragrafo; da quello computazionale però, per la determinazione degli autovalori anche quando si ha un'espressione esplicita del polinomio caratteristico il calcolo delle sue radici può presentare problemi (non sappiamo risolvere equazioni polinomiali di grado qualsiasi).

Osservazione 270. La ricerca degli autovalori può risultare più accessibile se si individuano delle proprietà della matrice esprimibili tramite "equazioni" coinvolgenti somme di sue potenze e loro prodotti per scalari: da queste si ottengono altre equazioni (più facilmente risolvibili, in quanto di grado inferiore rispetto all'equazione caratteristica) che ogni autovalore deve soddisfare.

Esempio 5.3.12. Un'illustrazione di questo è nell'esempio 5.3.4 (terzo caso): sfruttando il fatto che $\mathbf{P}$ sia idempotente, quindi $\mathbf{P} = \mathbf{P}^2$ abbiamo e trovato che gli autovalori di $\mathbf{A}$ devono essere soluzioni della corrispondente equazione $\lambda^2 - \lambda = 0$, facilmente risolvibile.

5.4 Sul polinomio caratteristico

5.4.1 Numero di autovalori di una matrice

Osservazione 271. Più avanti si mostrerà che l'equazione caratteristica di una matrice di ordine n ha grado n.

Osservazione 272. Ai fini della ricerca degli autovalori, però, vi è una differenza tra la fattorizzazione di polinomi (caratteristici) a coefficienti reali e complessi.

Teorema 5.4.1 (Teorema fondamentale dell'algebra)**.** *Ogni polinomio a coefficienti complessi si fattorizza in fattori di grado 1 a coefficienti complessi.*

Osservazione 273. Lo stesso non si può dire per la fattorizzazione di polinomi a coefficienti reali.

Dimostrazione. In appendice D del libro. □

Osservazione 274. L'implicazione pratica del teorema fondamentale è che:

- esistono matrici reali senza autovalori reali (vedi esempio 5.3.5)
- le matrici complesse (o quelle reali pensate come complesse) hanno sempre autovalori complessi e questi sono (contati con la loro molteplicità) tanti quant'è l'ordine della matrice. Trovarli tutti dipende esclusivamente dalla nostra abilità nel fattorizzare il suo polinomio caratteristico.

Osservazione 275. Altresì per una matrice complessa $\mathbf{A}$ di ordine n esistono sempre vettori non nulli di $\mathbb{C}^n$ (gli autovettori di $\mathbf{A}$) la cui pre-moltiplicazione per $\mathbf{A}$ equivale alla moltiplicazione per un opportuno scalare.

Osservazione 276. Per questi motivo quando si tratta la teoria degli autovalori si preferisce impiegare matrici complesse, piuttosto che quelle reali.

5.4.2 Altre proprietà del polinomio caratteristico

Osservazione 277. Si ricordi che per convenzione si assume che il polinomio nullo (ossia un polinomio in cui tutti i coefficienti sono nulli, come ad esempio $0 + 0X + 0X^2$) abbia grado $-\infty$, minore di qualsiasi numero (che invece è polinomio di grado 0). Cominciamo col fare la seguente osservazione.

Proposizione 5.4.2. *Se $\mathbf{B}(X)$ è una matrice quadrata di ordine m i cui elementi sono polinomi di grado ≤ 1, allora il grado g del determinante $\det(\mathbf{B}(x))$ è $\leq m$.*

Dimostrazione. Per induzione su m, sia $\mathbf{B}(X)$ con elementi (polinomi in X) di grado ≤ 1 :

- se $\mathbf{B}(X)$ è matrice di ordine $1 = m$, il determinante della matrice coincide con il polinomio, che è di grado $g \leq 1 = m$ e quindi $g \leq m$
- supponendo la disuguaglianza vera per matrici di ordine $m - 1$, proviamo che vale anche se l'ordine è m. Sfruttando la formula del determinante per colonne sulla prima colonna (teniamo $\mathbf{B}(X) = \mathbf{A}$ per comodità di notazione) si ha

$$\begin{aligned} \det \mathbf{A} &= \sum_{i=1}^{m} (-1)^{i+1} a_{i1} \det \mathbf{A}_{i1} \\ &= \left(\sum_{i=1}^{m-1} (-1)^{i+1} a_{i1} \det \mathbf{A}_{i1} \right) + (-1)^{m+1} a_{m,1} \det \mathbf{A}_{m,1} \end{aligned}$$

  Ora in quest'ultima equazione il primo termine è il determinante della matrice avente una colonna in meno, e per ipotesi induttiva ha grado $g \leq m - 1$; al secondo termine a_{m1} ha grado ≤ 1 mentre $\det \mathbf{A}_{m,1}$ ha grado $\leq m - 1$, quindi complessivamente il secondo termine ha al più grado m. E per effetto di questo secondo termine il grado complessivo è al più m.

□

Corollario 5.4.3. *Come conseguenza della 5.4.2 se tutti gli elementi di una delle m colonne di $\mathbf{B}(X)$ sono elementi di $\mathbb{K}$ (grado $= 0$), allora il grado di $\det(\mathbf{B}(X))$ è al più $m - 1$.*

Lemma 5.4.4. *Sia* $\mathbf{A} = [a_{ij}]$ *una matrice* $n \times n$ *con* $n \geq 2$*. Allora esiste un polinomio* $f(X)$ *di grado al più* $n-2$ *tale che*

$$p_{\mathbf{A}}(X) = f(X) + \prod_{i=1}^{n}(a_{ii} - X)$$

con secondo termine la produttoria degli elementi della diagonale di $\mathbf{A} - X\mathbf{I}$

Dimostrazione. Per induzione su $n \geq 2$:

- se $n = 2$ (primo passo dell'induzione) allora

  $$p_{\mathbf{A}}(X) = \det(\mathbf{A} - X\mathbf{I}_2) = \det \begin{bmatrix} a_{11} - X & a_{12} \\ a_{21} & a_{22} - X \end{bmatrix} = (a_{11} - X)(a_{22} - X) + f(X)$$

  dove $f(X) = -a_{12}a_{21}$ ha grado $\leq 0 = n - 2$.

- sia $n > 2$ e si supponga la proposizione vera per matrici $(n-1) \times (n-1)$. Sia $\mathbf{A} = [a_{ij}]$ una matrice $n \times n$. Sviluppando

  $$p_{\mathbf{A}} = \det(\mathbf{A} - X\mathbf{I}_n) = \det \begin{bmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - X & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - X \end{bmatrix}$$

  Si nota che gli elementi sulla diagonale sono polinomi di grado 1, mentre gli altri sono scalari di grado 0. Sviluppando il determinante rispetto alla prima riga, spezzando il calcolo sul primo elemento rispetto a quello sulla parte rimanente della riga, si ottiene:

  $$p_{\mathbf{A}}(X) = (a_{11} - X) \det \begin{bmatrix} a_{22} - X & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} - X \end{bmatrix} + \underbrace{\sum_{j \geq 2} (-1)^{i+j} a_{1j} \det((\mathbf{A} - X\mathbf{I}_n)_{1j})}_{g(X)} \tag{5.5}$$

  Considerando il secondo termine $g(X)$, poiché gli elementi di $(\mathbf{A} - X\mathbf{I}_n)_{1j}$ sono polinomi di grado ≤ 1, e inoltre tutti gli elementi della sua prima colonna (ossia $a_{21}, \dots, a_{n1}$ elementi di $\mathbb{K}$, di grado 0) dalla 5.4.2 si ottiene che $\det((\mathbf{A} - X\mathbf{I}_n)_{1j})$ ha grado al più $n-2$. Quindi anche $g(X)$ ha grado al più $n-2$ (a_{12} e seguenti hanno grado 0, e quindi non modificano il grado totale). Venendo al determinante del primo termine, per ipotesi induttiva si ha che

  $$\det \begin{bmatrix} a_{22} - X & \dots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \dots & a_{nn} - X \end{bmatrix} = h(X) + \prod_{i=2}^{n}(a_{ii} - X) \tag{5.6}$$

con $h(X)$ di grado al più $n-3$ ($n-2$ per ipotesi base, -1 dato che togliamo una colonna). Allora:

$$\begin{aligned} p_{\mathbf{A}}(X) &= (a_{11}-X)\left(h(X)+\prod_{i=2}^{n}(a_{ii}-X)\right)+g(X) \\ &= (a_{11}-X)h(X)+\left((a_{11}-X)\prod_{i=2}^{n}(a_{ii}-X)\right)+g(X) \\ &= \underbrace{(a_{11}-X)h(X)+g(X)}_{f(X)}+\prod_{i=1}^{n}(a_{ii}-X) \end{aligned}$$

con $f(X)$ che ha grado al più $n-2$

□

Proposizione 5.4.5 (Proprietà del polinomio caratteristico di matrice quadrata)**.** *Sia* $\mathbf{A} = [a_{ij}]$ *una matrice* $n \times n$*. Allora* $p_{\mathbf{A}}(X) = \det(\mathbf{A} - X\mathbf{I})$*:*

1. *ha grado* n*;*
2. *ha coefficiente di* X^n *pari a* $(-1)^n$*;*
3. *ha coefficiente di* X^{n-1} *pari a* $(-1)^n \operatorname{Tr}\mathbf{A} = (-1)^{n-1}\sum_{i=1}^{n} a_{ii}$
4. *ha termine noto pari a* $\det\mathbf{A}$

Dimostrazione. Se:

- $\mathbf{A} = [a_{11}]$ allora
$$p_{\mathbf{A}}(X) = \det([a_{11}] - X\mathbf{I}) = \det([a_{11} - X]) = (a_{11} - X)$$
ha grado 1, coefficiente di X uguale a -1 e termine noto $a_{11} = \det\mathbf{A}$

- se $n \geq 2$. Poiché $\prod_{i=1}^{n}(a_{ii}-X)$ ha grado n, ed essendo (dal lemma 5.4.4) il grado del polinomio il massimo tra quello di $f(X)$ ($n-2$) e quello di $\prod_{i=1}^{n}(a_{ii}-X)$ (di grado n), si ha che il grado del determinante è n (abbiamo dimostrato il punto 1).
Inoltre il coefficiente di X^n e il coefficiente di X^{n-1} coincidono con quelli di $\prod_{i=1}^{n}(a_{ii}-X)$ rispettivamente: infatti si sa che
$$p_{\mathbf{A}}(X) = f(X) + \prod_{i=1}^{n}(a_{ii}-X)$$
ha grado n e che $f(x)$ ha al più grado $n-2$ pertanto $\prod_{i=1}^{n}(a_{ii}-X)$ contiene i termini di X con grado $n-1$ ed n e dunque i coefficienti di grado $n-1$ ed n di $p_{\mathbf{A}}(X)$ coincideranno con quelli di $\prod_{i=1}^{n}(a_{ii}-X)$. Sfruttando questa cosa e sviluppando $\prod_{i=1}^{n}(a_{ii}-X)$ si ottiene che il coefficiente di X^n è $(-1)^n$ (infatti il termine con X^n si ottiene moltiplicando tutti gli X dei vari termini quindi il coefficiente sarà dato dal -1 elevato al numero di fattori n). Ciò prova il punto 2.

Il coefficiente di X^{n-1} coincide con il coefficiente di X^{n-1} di $\sum_{i=1}^{n} a_{ii} \prod_{j\neq i}(a_{jj} - X)$, infatti sviluppandolo si nota come

$$\sum_{i=1}^{n} a_{ii} \prod_{j\neq i}(a_{jj} - X) = a_{11}\left[(a_{22} - X) \cdot \ldots \cdot (a_{22} - X)\right] + \ldots + a_{nn}\left[(a_{11} - X) \cdot \ldots \cdot (a_{n-1,n-1} - X)\right] \tag{5.7}$$

questi siano tutti termini di grado $n - 1$. In altre parole

$$\prod_{i=1}^{n}(a_{ii} - X) = (a_{11} - X)(a_{22} - X) \cdot \ldots \cdot (a_{nn} - X)$$

i termini di grado $n-1$ si ottengono moltiplicando $n-1$ X elementi a turno ed un a_{ii} a giro; il coefficiente dunque coincide con quello della somma di questi, che è appunto quanto visto in equazione 5.7.
Poiché per ogni $i = 1, \ldots, n$ si ha che il coefficiente di X^{n-1} di $\prod_{j\neq i}(a_{jj} - X)$ è $(-1)^{n-1}$, allora il coefficiente di X^{n-1} di $\sum_{i=1}^{n} a_{ii} \prod_{j\neq i}(a_{jj} - X)$ è $(-1)^{n-1} \sum_{i=1}^{n} a_{ii} = -(1)^{n-1} \operatorname{Tr} \mathbf{A}$, e anche la 3 è provata.
Per provare 4 si osservi che attribuendo il valore 0 all'indeterminata X si ottiene che il termine noto di $p_{\mathbf{A}}(X)$ è:

$$p_{\mathbf{A}}(0) = \det(\mathbf{A} - 0\mathbf{I}) = \det(\mathbf{A})$$

□

Osservazione 278. Se $\mathbf{A}$ è una matrice quadrata complessa di ordine n e $p_{\mathbf{A}}(X)$ il suo polinomio caratteristico, per il teorema fondamentale dell'algebra $p_{\mathbf{A}}(X)$ si fattorizza in fattori di grado 1, per cui esistono numeri complessi distinti $\lambda_1, \ldots, \lambda_r$, gli autovalori di $\mathbf{A}$, e numeri naturali $m_1, \ldots, m_r$ tali che

$$p_{\mathbf{A}}(X) = c(X - \lambda_1)^{m_1}(X - \lambda_2)^{m_2} \ldots (X - \lambda_r)^{m_r} \tag{5.8}$$

per un opportuno numero complesso c, che come si vedrà è $(-1)^n$.

Definizione 5.4.1 (Molteplicità algebrica dell'autovalore λ_i)**.** Si intende il numero naturale m_i in equazione 5.8, per ogni $i = 1, \ldots, r$.

Osservazione 279. La molteplicità algebrica può non coincidere con la molteplicità geometrica. Per esempio, l'autovalore 1 della matrice $\mathbf{B}$ nell'esempio 5.3.9 ha molteplicità algebrica uguale a 2 e geometrica uguale a 1.

Proposizione 5.4.6. *La somma delle molteplicità algebriche degli autovalori di una matrice quadrata è uguale all'ordine della stessa*

Dimostrazione. Per il punto 1 della proposizione 5.4.5 $p_{\mathbf{A}}(X)$ ha grado n, per cui si ha

$$n = m_1 + m_2 + \ldots + m_r$$

□

Osservazione 280. Inoltre per il punto 2 di proposizione 5.4.5 il coefficiente di X^n di $p_{\mathbf{A}}(X)$ è $(-1)^n$, quindi nella 5.8 si ha che $c = (-1)^n$; analogamente a quanto avviene per la fattorizzazione $ax^2+bx+c = 0$ posta ad $a(x-x_1)(x-x_2) = 0$ ed a è il coefficiente di termine di grado massimo, allo stesso modo avviene qui per cui il c della 5.8 che è $(-1)^n$.

Osservazione 281. Alla luce della precedente considerazione e mediante raccoglimento dei segni dai fattori una fattorizzazione alternativa del polinomio caratteristico è la seguente:

$$\begin{aligned}
p_{\mathbf{A}}(X) &= c \cdot (X-\lambda_1)^{m_1} \cdot (X-\lambda_2)^{m_2} \cdot \ldots \cdot (X-\lambda_r)^{m_r} \\
&= (-1)^n \cdot (-1)^{m_1} \cdot (\lambda_1 - X)^{m_1} \cdot (-1)^{m_2} \cdot (\lambda_2 - X)^{m_2} \cdot \ldots \cdot (-1)^{m_r} \cdot (\lambda_r - X)^{m_r} \\
&= (-1)^{n+m_1+\ldots+m_r} \cdot (\lambda_1 - X)^{m_1} \cdot (\lambda_2 - X)^{m_2} \cdot \ldots \cdot (\lambda_r - X)^{m_r} \\
&= (-1)^{n+n} \cdot (\lambda_1 - X)^{m_1} \cdot (\lambda_2 - X)^{m_2} \cdot \ldots \cdot (\lambda_r - X)^{m_r} \\
&= (-1)^{2n} \cdot (\lambda_1 - X)^{m_1} \cdot (\lambda_2 - X)^{m_2} \cdot \ldots \cdot (\lambda_r - X)^{m_r} \\
&= (\lambda_1 - X)^{m_1} \cdot (\lambda_2 - X)^{m_2} \cdot \ldots \cdot (\lambda_r - X)^{m_r} \qquad (5.9)
\end{aligned}$$

Proposizione 5.4.7. *Se* $\mathbf{D}$ *se la matrice diagonale di ordine* n *in cui elementi diagonali sono gli autovalori di* $\mathbf{A}$ *(ciascuno contato con la sua molteplicità algebrica) ossia*

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & 0 & 0 & \ldots & 0 \\ 0 & \lambda_1 & 0 & \ldots & 0 \\ 0 & 0 & \lambda_2 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_r \end{bmatrix}$$

con r_1 *termini* λ_1 *sulla diagonale (e così via per gli altri) si ha che*

$$p_{\mathbf{A}}(X) = p_{\mathbf{D}}(X)$$

Dimostrazione. Infatti

$$p_{\mathbf{D}}(X) = \det(\mathbf{D} - X\mathbf{I}) = \det \begin{bmatrix} \lambda_1 - X & 0 & 0 & \ldots & 0 \\ 0 & \lambda_1 - X & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \lambda_r - X \end{bmatrix}$$

da cui (per questo esempio) il determinante si calcola ricorsivamente come prodotto di un termine sulla diagonale per il determinante della matrice eliminando prima riga e colonna della matrice considerata, sino a giungere al determinante dell'ultima

$$\begin{aligned}
\det(\mathbf{D} - X\mathbf{I}) &= (-1)^{1+1}(\lambda_1 - X) \cdot \det(\mathbf{D} - X\mathbf{I})_{11} + 0 \cdot \ldots \\
\det(\mathbf{D} - X\mathbf{I})_{11} &= (-1)^{1+1}(\lambda_1 - X) \cdot \det((\mathbf{D} - X\mathbf{I})_{11})_{11} + 0 \cdot \ldots \\
\ldots &= \ldots \\
\det([\lambda_r - X]) &= \lambda_r - X
\end{aligned}$$

quindi

$$p_{\mathbf{D}}(X) = \det(\mathbf{D} - X\mathbf{I}) = (\lambda_1 - X)^{m_1} \cdot \ldots \cdot (\lambda_r - X)^{m_r}$$

e $p_{\mathbf{A}}(X) = p_{\mathbf{D}}(X)$ □

Proposizione 5.4.8. *Sia* $\mathbf{A}$ *una matrice complessa: la traccia di* $\mathbf{A}$ *coincide con la somma degli autovalori di* $\mathbf{A}$ *ciascuno contato con la sua molteplicità algebrica*

$$\operatorname{Tr} \mathbf{A} = \sum_{i=1}^{r} \lambda_i m_i$$

Dimostrazione. Stando a quanto predicato dalla proposizione 5.4.5 si ha che:

$$p_{\mathbf{A}}(X) = (-1)^n(X^n - \operatorname{Tr}\mathbf{A} \cdot X^{n-1} + \ldots + (-1)^n \det \mathbf{A}) \tag{5.10}$$

Ma al contempo si ha anche che (osservazione 278):

$$p_A(\mathbf{X}) = (-1)^n(X-\lambda_1)^{m_1} \cdot \ldots \cdot (X-\lambda_r)^{m_r}$$

Quindi deve essere che (elidendo i segni in entrambi i membri):

$$\begin{aligned}(X^n - \operatorname{Tr}\mathbf{A} \cdot X^{n-1} + \ldots + (-1)^n \det \mathbf{A}) &= (X-\lambda_1)^{m_1} \cdot \ldots \cdot (X-\lambda_r)^{m_r} \\ &= \underbrace{(X-\lambda_1)\cdot\ldots\cdot(X-\lambda_1)}_{m_1 \text{ volte}} \cdot \ldots \cdot \underbrace{(X-\lambda_r)\cdot\ldots\cdot(X-\lambda_r)}_{m_r \text{ volte}}\end{aligned}$$

Considerando ciascun blocco $(X-\lambda_i)\cdot\ldots\cdot(X-\lambda_i)$, in questo il coefficiente di X^{n-1} si ottiene sommando tutti i prodotti che coinvolgono $n-1$ termini X e un termine λ_i a turno; a livello complessivo il termine di grado $n-1$ si ha mediante

$$\begin{aligned}&= \overbrace{-\lambda_1 \cdot \underbrace{X\cdot\ldots\cdot X}_{n-1 \text{ volte}} - \ldots - \lambda_1 \cdot \underbrace{X\cdot\ldots\cdot X}_{n-1 \text{ volte}}}^{m_1 \text{ termini}} + \ldots + \overbrace{-\lambda_r \cdot \underbrace{X\cdot\ldots\cdot X}_{n-1 \text{ volte}} - \ldots - \lambda_r \cdot \underbrace{X\cdot\ldots\cdot X}_{n-1 \text{ volte}}}^{m_r \text{ termini}} \\ &= -\lambda_1 \cdot m_1 \cdot X^{n-1} + \ldots + (-\lambda_r)\cdot m_r X^{n-1} \\ &= X^{n-1}(-\lambda_1 m_1 - \ldots - \lambda_r m_r) \\ &= -(\lambda_1 m_1 + \ldots + \lambda_r m_r)X^{n-1}\end{aligned}$$

e dal confronto del coefficiente di quest'ultima con la 5.10 si ha che

$$\operatorname{Tr}\mathbf{A} = \lambda_1 m_1 + \ldots + \lambda_r m_r$$

□

Proposizione 5.4.9. *Sia* $\mathbf{A}$ *una matrice complessa: il determinante di* $\mathbf{A}$ *coincide con il prodotto degli autovalori di* $\mathbf{A}$ *ciascuno contato con la sua molteplicità algebrica:*

$$\det \mathbf{A} = \lambda_1^{m_1} \cdot \ldots \cdot \lambda_r^{m_r}$$

Dimostrazione. Si ha:

$$\det \mathbf{A} = \det(\mathbf{A} - 0\mathbf{I}) = P_{\mathbf{A}}(0) \overset{(1)}{=} (\lambda_1 - 0)^{m_1} \cdot \ldots \cdot (\lambda_r - 0)^{m_r} = \lambda_1^{m_1} \cdot \ldots \cdot \lambda_r^{m_r}$$

dove in (1) abbiamo applicato l'equazione 5.9: □

Proposizione 5.4.10. *L'insieme degli autovalori di* $\mathbf{A}$ *e* $\mathbf{A}^T$ *coincidono.*

Dimostrazione. Infatti si avrà che

$$\det(\mathbf{A} - \lambda\mathbf{I}) \overset{(1)}{=} \det\left[(\mathbf{A} - \lambda\mathbf{I})^T\right] = \det(\mathbf{A}^T - \lambda\mathbf{I}^T) = \det(\mathbf{A}^T - \lambda\mathbf{I})$$

(con (1) per teorema 4.3.4) dunque $\mathbf{A}$ e $\mathbf{A}^T$ hanno lo stesso polinomio caratteristico e gli stessi autovettori. □

Proposizione 5.4.11. *Se λ è un autovalore della matrice quadrata* $\mathbf{A}$ *allora* $\overline{\lambda}$ *è un autovalore della matrice coniugata* $\overline{\mathbf{A}}$

Dimostrazione. Sia $\mathbf{Av} = \lambda\mathbf{v}$; coniugando entrambi i membri si ha

$$\overline{\mathbf{Av}} = \overline{\lambda\mathbf{v}} \overset{(1)}{=} \overline{\lambda}\overline{\mathbf{v}}$$

con (1) grazie alla proprietà $\overline{ab} = \overline{a} \cdot \overline{b}$ del coniugio.
Ora riusciamo a dimostrare la proposizione se mostriamo che $\overline{\mathbf{Av}} = \overline{\mathbf{A}}\overline{\mathbf{v}}$ in quanto in tal caso si avrebbe

$$\overline{\mathbf{A}}\overline{\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$$

Il posto $(j, 1)$ del vettore $\overline{\mathbf{Av}}$ è

$$\overline{\sum_{i=1}^{n} a_{ji}v_{i1}} \overset{(1)}{=} \sum_{i=1}^{n} \overline{a_{ji}} \cdot \overline{v_{i1}}$$

(con (1) per le proprietà del coniugio $\overline{ab} = \overline{a} \cdot \overline{b}$ e $\overline{a+b} = \overline{a} + \overline{b}$), ma quest'ultimo termine è proprio il coefficiente di posto $(j, 1)$ di $\overline{\mathbf{A}}\overline{\mathbf{v}}$; pertanto possiamo dire che $\overline{\mathbf{Av}} = \overline{\mathbf{A}}\overline{\mathbf{v}}$ e si conclude. □

5.5 Proprietà degli autospazi

Osservazione 282. Sia $\mathbf{A}$ una matrice complessa $n \times n$. L'idea di fondo per lo studio dell'applicazione lineare $f_{\mathbf{A}} : \mathbb{C}^n \to \mathbb{C}^n$ $f(\mathbf{v}) = \mathbf{Av}$ è quella di cercare un'opportuna base $\mathscr{B}$ di $\mathbb{C}^n$ rispetto alla quale la matrice $\mathbf{B}$ associata a f abbia molti coefficienti nulli. In tal caso il calcolo di $f_{\mathbf{B}}$ è meno problematico.

Osservazione 283. Se $\mathbf{A}$ è matrice associata ad f rispetto alla base $\mathscr{E}$ su dominio e codominio, mentre $\mathbf{B}$ è matrice associata ad f rispetto alla base $\mathscr{B}$ su dominio e codominio, si ha:

$$\mathrm{C}_{\mathscr{E}_n}(f(\mathbf{v})) = \mathbf{A}\,\mathrm{C}_{\mathscr{E}_n}(\mathbf{v})$$
$$\mathrm{C}_{\mathscr{B}}(f(\mathbf{v})) = \mathbf{B}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v})$$

e applicando il teorema 2.6.5 si ottiene la relazione

$$\mathbf{A} = \mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}\mathbf{B}\mathbf{M}^{-1}_{\mathscr{E}\leftarrow\mathscr{B}}$$

con $\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}$ matrice del cambiamento di base definita come

$$\mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}} = [\mathrm{C}_{\mathscr{E}}(\mathbf{b}_1) \dots \mathrm{C}_{\mathscr{E}}(\mathbf{b}_n)] = [\mathbf{b}_1 \dots \mathbf{b}_n]$$

con $\{\mathbf{b}_1, \dots, \mathbf{b}_n\} \in \mathscr{B}$.
Stante questo, ponendo $\mathbf{S} = \mathbf{M}_{\mathscr{E}\leftarrow\mathscr{B}}$ con $\mathbf{S}$ matrice che ha come colonne i vettori della base $\mathscr{B}$, la relazione che intercorre tra $\mathbf{A}$ e $\mathbf{B}$ è $\mathbf{A} = \mathbf{SBS}^{-1}$.
In altre parole $f_{\mathbf{A}}$ ammette una fattorizzazione del tipo $f_{\mathbf{A}} = g f_{\mathbf{B}} g^{-1}$ con g isomorfismo di $\mathbb{C}^n$.

Definizione 5.5.1 (Matrici quadrate simili)**.** Due matrici quadrate $\mathbf{A}$ e $\mathbf{B}$ si dicono simili se esiste una matrice invertibile $\mathbf{S}$ tale che $\mathbf{A} = \mathbf{SBS}^{-1}$

Proposizione 5.5.1. *La relazione di similitudine è una relazione di equivalenza (è riflessiva, simmetrica e transitiva)*

Dimostrazione. Per mostrare la riflessività basta porre $\mathbf{S} = \mathbf{I}$ e concludere osservando che

$$\mathbf{A} = \mathbf{IAI}^{-1} = \mathbf{IAI}$$

La simmetria si deriva dal fatto che:

$$\mathbf{A} = \mathbf{SBS}^{-1} \iff \mathbf{AS} = \mathbf{SB} \iff \mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$$

e la matrice ricercata è $\mathbf{S}^{-1}$. Infine per la transitività se $\mathbf{A} = \mathbf{SBS}^{-1}$, e $\mathbf{B} = \mathbf{TCT}^{-1}$ allora

$$\mathbf{A} = \mathbf{S}\underbrace{\mathbf{TCT}^{-1}}_{\mathbf{B}}\mathbf{S}^{-1} = \mathbf{ST} \cdot \mathbf{C} \cdot (\mathbf{ST})^{-1}$$

e la matrice ricercata è $\mathbf{ST}$. □

Proposizione 5.5.2. *Due matrici simili hanno lo stesso polinomio caratteristico e gli stessi autovalori con le stesse molteplicità algebriche e geometriche.*

Dimostrazione. Siano $\mathbf{A}$ e $\mathbf{B}$ due matrici simili di ordine n ed $\mathbf{S}$ una matrice invertibile tale che $\mathbf{A} = \mathbf{SBS}^{-1}$. Allora:

$$\begin{aligned} p_{\mathbf{A}}(X) = \det(\mathbf{A} - X\mathbf{I}_n) = \det(\mathbf{SBS}^{-1} - X\mathbf{I}_n) = \det(\mathbf{SBS}^{-1} - X\mathbf{SI}_n\mathbf{S}^{-1}) &= \det(\mathbf{S}(\mathbf{B} - X\mathbf{I}_n)\mathbf{S}^{-1}) \\ &\overset{(1)}{=} \det\mathbf{S} \cdot \det(\mathbf{B} - X\mathbf{I}_n) \cdot (\det\mathbf{S})^{-1} \\ &\overset{(2)}{=} \det(\mathbf{B} - X\mathbf{I}_n) \\ &= p_{\mathbf{B}}(X) \end{aligned}$$

(dove in (1) si è applicato 4.3.2, mentre in (2) si è applicato 4.3.12) per cui $\mathbf{A}$ e $\mathbf{B}$ hanno lo stesso polinomio caratteristico, e quindi anche gli stessi autovalori con le stesse molteplicità algebriche.
Passiamo all'analisi delle molteplicità geometriche; sia ora λ un autovalore di $\mathbf{A}$ e $\mathbf{B}$; considerando un elemento del suo autospazio

$$\begin{aligned} \mathbf{v} \in E_{\mathbf{B}}(\lambda) &\iff \mathbf{Bv} = \lambda\mathbf{v} \overset{(1)}{\iff} \mathbf{S}^{-1}\mathbf{ASv} = \lambda\mathbf{v} \overset{(2)}{\iff} \mathbf{ASv} = \lambda\mathbf{Sv} \\ &\iff \mathbf{Sv} \in E_{\mathbf{A}}(\lambda) \end{aligned}$$

(dove in (1) abbiamo sostituito $\mathbf{B} = \mathbf{S}^{-1}\mathbf{AS}$ e in (2) pre-moltiplicato per $\mathbf{S}$) segue che la restrizione a $E_{\mathbf{B}}(\lambda)$ della pre-moltiplicazione per $\mathbf{S}$ è un'applicazione lineare da $E_{\mathbf{B}}(\lambda)$ a $E_{\mathbf{A}}(\lambda)$. Tale applicazione lineare è un isomorfismo: la sua inversa è la restrizione a $E_{\mathbf{A}}(\lambda)$ della pre-moltiplicazione per $\mathbf{S}^{-1}$. Essendo biiettiva i due insiemi che congiunge ($E_{\mathbf{A}}(\lambda)$ ed $E_{\mathbf{B}}(\lambda)$) hanno medesima cardinalità e per medesima dimensione (per proposizione 2.6.6), ossia la molteplicità geometrica di λ in quanto autovalore di $\mathbf{B}$. □

Osservazione 284. Nel teorema che segue, impiegando la relazione di similitudine tra matrici, proviamo che la molteplicità geometrica di un autovalore è minore o uguale a quella algebrica. Questo fatto ci permetterà nel prossimo paragrafo di caratterizzare le matrici diagonalizzabili.

Teorema 5.5.3. *Siano* $\mathbf{A}$ *una matrice* $n \times n$ *e* λ*un autovalore di* $\mathbf{A}$ *con molteplicità geometrica* d *e algebrica* m. *Allora* $d \leq m$.

Dimostrazione. Si estenda una base $\{\mathbf{v}_1, \ldots, \mathbf{v}_d\}$ dell'autospazio $E_{\mathbf{A}}(\lambda)$ (si noti la molteplicità geometrica a pedice dell'ultimo vettore) a una base $\mathscr{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_d, \mathbf{v}_{d+1}, \ldots, \mathbf{v}_n\}$ di $\mathbb{C}^n$ e si consideri la matrice $\mathbf{B}$ associata all'applicazione lineare $f_{\mathbf{A}} : \mathbb{C}^n \to \mathbb{C}^n$ rispetto alla base $\mathscr{B}$ (su dominio e codominio), ossia $\mathrm{C}_{\mathscr{B}}(f_{\mathbf{A}}(\mathbf{v})) = \mathbf{B}\,\mathrm{C}_{\mathscr{B}}(\mathbf{v})$. Per le prime $i = 1, \ldots, d$ colonne si ha, per quanto visto nella dimostrazione del teorema 2.6.3, che la colonna i-esima di $\mathbf{B}$

$$\mathbf{B}\mathbf{e}_i = \mathrm{C}_{\mathscr{B}}(f(\mathbf{v}_i)) = \mathrm{C}_{\mathscr{B}}(\mathbf{A}\mathbf{v}_i) \overset{(1)}{=} \mathrm{C}_{\mathscr{B}}(\lambda\mathbf{v}_i) \overset{(2)}{=} \lambda\,\mathrm{C}_{\mathscr{B}}(\mathbf{v}_i) = \lambda\mathbf{e}_i \tag{5.11}$$

(in (1) si è applicata la definizione di autovalore; (2) dato che è funzione lineare) dove $\mathbf{e}_1, \ldots, \mathbf{e}_d$ sono le prime d colonne della matrice identità $\mathbf{I}_n$.
Pertanto $\mathbf{B}$ è una matrice triangolare superiore a blocchi del tipo

$$\mathbf{B} = \begin{bmatrix} \lambda\mathbf{I}_d & \mathbf{C} \\ \mathbb{O} & \mathbf{D} \end{bmatrix}$$

con $\lambda\mathbf{I}_d$ di tipo $d \times d$ e $\mathbf{D}$ matrice quadrata $(n-d) \times (n-d)$. Quindi si ha che $p_{\mathbf{B}}(X) = (\lambda - X)^d p_{\mathbf{D}}(X)$, infatti ad esempio sviluppando il determinante sulla prima colonna:

$$\begin{aligned} p_{\mathbf{B}}(X) = \det(\mathbf{B} - X\mathbf{I}) &= \det \begin{bmatrix} \lambda - X & 0 & \ldots & \mathbf{c}_1^T \\ 0 & \lambda - X & \ldots & \mathbf{c}_2^T \\ \ldots & \ldots & \ldots & \ldots \\ & \mathbb{O} & & \mathbf{D} - X\mathbf{I} \end{bmatrix} \\ &= (\lambda - X)^d \det(\mathbf{D} - X\mathbf{I}) \\ &= (\lambda - X)^d p_{\mathbf{D}}(X) \end{aligned}$$

Poiché $\mathbf{A}$ è la matrice associata all'applicazione lineare $f_{\mathbf{A}}$ rispetto alla base canonica di $\mathbb{C}^n$ (su dominio e codominio) $\mathbf{B}$ è simile ad $\mathbf{A}$ (osservazione 283); dalla proposizione 5.5.2 segue dunque che:

$$p_{\mathbf{B}}(X) = p_{\mathbf{A}}(X) = (\lambda - X)^d p_{\mathbf{D}}(X)$$

Pertanto il polinomio $(\lambda - X)^d$ divide il polinomio caratteristico di $\mathbf{A}$; da questo discende che per cui la molteplicità geometrica d di λ è $\leq$ a quella algebrica m (nel senso che magari $p_{\mathbf{D}}(X)$ può essere fattorizzato ponendo in evidenza ulteriori termini $(\lambda - X)$, per un totale di $m - d$). □

Osservazione 285. Ricordando definizione 2.1.7, se U e Z sono due sottospazi dello spazio vettoriale V, la somma di U e Z

$$U + Z = \{\mathbf{u} + \mathbf{z} : \mathbf{u} \in U, \mathbf{z} \in Z\}$$

è il più piccolo sottospazio di Vcontenente U e Z.
Inoltre se ogni elemento di $U + Z$ si scrive in modo unico come somma di un vettore di U e un vettore di Z, allora la somma di U e Z si chiama somma diretta (definizione 2.1.8) e si indica con il simbolo $U \oplus Z$.
Quanto detto si può estendere a un numero finito di sottospazi di V, come già accennato nella proposizione 2.3.13.

Definizione 5.5.2 (Somma di un numero finito di sottospazi)**.** Siano $U_1, \dots, U_n$ sottospazi vettoriali dello spazio vettoriale V. La somma:

$$\sum_{i=1}^{n} U_i = \{\mathbf{u}_1 + \ldots + \mathbf{u}_n : \mathbf{u}_i \in U_i, i = 1, \ldots, n\} \tag{5.12}$$

è il più piccolo sottospazio di V contenente $U_1, \dots, U_n$.

Definizione 5.5.3 (Somma diretta di un numero finito di sottospazi)**.** Se ogni elemento di $\sum_{i=1}^{n} U_i$ si scrive in modo unico come somma di vettori di $U_1, \dots, U_n$ la somma $\sum_{i=1}^{n} U_i$ si chiama diretta e si indica con il simbolo $\bigoplus_{i=1}^{n} U_i$

Osservazione 286. Per verificare che $\sum_{i=1}^{n} U_i = \bigoplus_{i=1}^{n} U_i$ è sufficiente verificare che i sottospazi $U_1, \dots, U_n$ sono indipendenti, secondo definizione 2.3.5

$$\sum_{i=1}^{n} \mathbf{u}_i = \mathbf{0}, \quad \mathbf{u}_i \in U_i \quad \text{per ogni } i = 1, \dots, n \implies \mathbf{u}_i = \mathbf{0} \quad \text{per ogni } i = 1, \dots, n$$

Osservazione 287. Si osservi che se U_i sono sottospazi indipendenti, ossia $\sum_{i=1}^{n} U_i = \bigoplus_{i=1}^{n} U_i$ e $\mathscr{B}_i$ è una base di U_i per ogni $i = 1, \dots, n$ allora $\mathscr{B} = \mathscr{B}_1 \cup \ldots \cup \mathscr{B}_n$ è una base di $\bigoplus_{i=1}^{n} U_i$ (Esercizio 5.17).

TODO: fixme

Teorema 5.5.4 (Somma e somma diretta di autospazi coincidono)**.** *Siano* $\mathbf{A}$ *una matrice* $n \times n$ *e* $\lambda_1, \dots, \lambda_k$ *autovalori distinti di* $\mathbf{A}$*. Allora si ha*

$$\sum_{i=1}^{n} E_{\mathbf{A}}(\lambda_i) = \bigoplus_{i=1}^{n} E_{\mathbf{A}}(\lambda_i)$$

Dimostrazione. Abbiamo osservato che è sufficiente provare che se $\mathbf{v}_i \in E_{\mathbf{A}}(\lambda_i)$ allora

$$\sum_{i=1}^{k} \mathbf{v}_i = \mathbf{0} \implies \mathbf{v}_i = \mathbf{0} \quad \text{per ogni } i = 1, \dots, k$$

Procediamo per induzione su k. Per $k = 1$ non c'è nulla da dimostrare. Sia allora $k > 1$ e si supponga che la somma degli autospazi relativi a $k - 1$ autovalori distinti di $\mathbf{A}$ sia diretta, quindi che i sottospazi siano indipendenti e si abbia

$$\sum_{i=1}^{k-1} \mathbf{v}_i = \mathbf{0} \implies \mathbf{v}_i = \mathbf{0} \quad \text{per ogni } i = 1, \dots, k-1$$

Ipotizzando che sia $\sum_{i=1}^{k} \mathbf{v}_i = \mathbf{0}$ andiamo a mostrare l'implicazione; pre-moltiplicando ambo i membri dell'uguaglianza per λ_1 si ottiene

$$\sum_{i=1}^{k} \lambda_1 \mathbf{v}_i = \lambda_1 \mathbf{0} = \mathbf{0} \tag{5.13}$$

mentre pre-moltiplicandoli per $\mathbf{A}$ si ottiene

$$\mathbf{0} = \mathbf{A}\mathbf{0} = \mathbf{A} \sum_{i=1}^{k} \mathbf{v}_i = \sum_{i=1}^{k} \mathbf{A}\mathbf{v}_i \overset{(1)}{=} \sum_{i=1}^{k} \lambda_i \mathbf{v}_i \tag{5.14}$$

dove in (1) si è sostituito essendo λ_i autovalore. Ora sottraendo membro a membro la 5.14 dalla 5.13 si ricava che

$$\sum_{i=2}^{k}(\lambda_i - \lambda_1)\mathbf{v}_i = \mathbf{0}$$

si noti che la sommatoria parte dal secondo, poiché al primo item abbiamo $(\lambda_1 - \lambda_1)\mathbf{v} = 0\mathbf{v} = \mathbf{0}$, che tralasciamo. Sfruttando l'ipotesi induttiva, prendendo i $k-1$ che vanno da 2 a k, dato che la loro somma è $\mathbf{0}$, allora anche i singoli elementi $(\lambda_i - \lambda_1)\mathbf{v} = \mathbf{0}$ per ogni $i \geq 2$ (possiamo applicarla perché se $\mathbf{v}_i \in E_{\mathbf{A}}(\lambda_i)$ anche $c\mathbf{v}_i \in E_{\mathbf{A}}(\lambda_i)$, essendo $E_{\mathbf{A}}(\lambda_i)$ uno spazio vettoriale).
Ora, dato che gli autovalori sono distinti si ha che $\lambda_i - \lambda_1 \neq 0$ per ogni $i \geq 2$, pertanto deve essere $\mathbf{v}_i = \mathbf{0}$ per ogni $i \geq 2$. Da ciò segue che anche $\mathbf{v}_1 = \mathbf{0}$ essendo $\sum_{i=1}^{k} \mathbf{v}_i = \mathbf{0}$. La dimostrazione è così completata. □

Osservazione 288. Come conseguenza del teorema 5.5.4 si ha il seguente risultato

Corollario 5.5.5. *Autovettori relativi ad autovalori distinti sono linearmente indipendenti*

Dimostrazione. Siano $\mathbf{A}$ una matrice quadrata $\lambda_1, \ldots, \lambda_k$ suoi autovalori distinti, con $\mathbf{v}_1, \ldots, \mathbf{v}_k$ autovettori ad essi corrispondenti e $\alpha_1\mathbf{v}_1 + \ldots + \alpha_k\mathbf{v}_k = \mathbf{0}$ una loro combinazione lineare nulla.
Poiché $E_{\mathbf{A}}(\lambda_i)$ è uno spazio vettoriale per ogni $i = 1, \ldots, k$ allora $\alpha_i\mathbf{v}_i \in E_{\mathbf{A}}(\lambda_i)$ (sono altri autovettori), per cui dal teorema 5.5.4 (da cui deriva $\sum \mathbf{u}_i = 0 \implies \mathbf{u}_i = \mathbf{0}$) si ottiene che $\alpha_i\mathbf{v}_i = \mathbf{0}$ per ogni $i = 1, \ldots, k$.
Dal fatto che ciascun $\mathbf{v}_i$, in quanto autovettore di $\mathbf{A}$, è non nullo, segue che $\alpha_i = 0$ per ogni $i = 1, \ldots, k$, ossia che i vettori $\mathbf{v}_1, \ldots, \mathbf{v}_k$ sono linearmente indipendenti. □

5.6 Matrici diagonalizzabili e triangolarizzabili

Definizione 5.6.1. $\mathbf{A}$ è diagonalizzabile se è simile a una matrice diagonale, ossia se esistono una matrice diagonale $\mathbf{D}$ e una matrice invertibile $\mathbf{S}$ tali che $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$.

Osservazione 289. In tal caso si dice che $\mathbf{S}$ diagonalizza $\mathbf{A}$.

Osservazione 290 (Applicazioni matrici diagonalizzabili). Le maggiori applicazioni del concetto si trovano:

- per lo studio dell'applicazione lineare indotta: nella sezione 5.3 abbiamo espresso il concetto di matrice diagonalizzabile in relazione all'applicazione lineare da essa indotta

- quando si vogliono calcolare le sue potenze (vedi sezione 5.1), dal momento per ogni numero naturale k si ha che $\mathbf{A}^k = \mathbf{S}\mathbf{D}^k\mathbf{S}^{-1}$, e se $\mathbf{D} = \mathbf{Diag}(d_1, \ldots, d_n)$, allora

$$\mathbf{D}^k = \mathbf{Diag}\big(d_1^k, \ldots, d_n^k\big)$$

Osservazione 291. Il calcolo della potenza di una matrice è trattabile più facilmente anche nel caso di matrici triangolari: sappiamo infatti che se $\mathbf{T}$ è una matrice triangolare, per ogni numero naturale k anche $\mathbf{T}^k$ è triangolare e i coefficienti diagonali di $\mathbf{T}^k$ sono le potenze k-esime dei coefficienti diagonali di $\mathbf{T}$. Possiamo allora dare una definizione analoga.

Definizione 5.6.2 (Matrice triangolarizzabile)**.** $\mathbf{A}$ è triangolarizzabile se simile a una matrice triangolare, ossia se esistono una matrice triangolare $\mathbf{T}$ e una matrice invertibile $\mathbf{S}$ tali che $\mathbf{A} = \mathbf{STS}^{-1}$.

Osservazione 292. In tal caso si dice che $\mathbf{S}$ triangolarizza $\mathbf{A}$.

Teorema 5.6.1 (Caratterizzazione delle matrici diagonalizzabili)**.** *Siano* $\mathbf{A}$ *una matrice complessa* $n \times n$ *e* $\lambda_1, \ldots, \lambda_k$ *i suoi autovalori distinti. I seguenti fatti sono equivalenti:*

1. $\mathbf{A}$ *è diagonalizzabile;*
2. $\mathbb{C}^n$ *ha una base costituita da autovettori di* $\mathbf{A}$
3. $\mathbb{C}^n = \bigoplus_{i=1}^k E_{\mathbf{A}}(\lambda_i)$
4. *la molteplicità geometrica di ciascun autovalore* λ_i *coincide con la sua molteplicità algebrica.*

Dimostrazione. Rispettivamente:

- la $1 \implies 2$ poiché:

$$\mathbf{A} = \mathbf{SDS}^{-1} \iff \mathbf{AS} = \mathbf{SD} \overset{(1)}{\iff} \mathbf{ASe}_i = \mathbf{S}\,\underbrace{\mathbf{De}_i}_{=d_i\mathbf{e}_i} = d_i\mathbf{Se}_i \quad \text{per ogni } i = 1, \ldots, n$$

  (dove in (1) siamo passati alla i-esima colonna). L'insieme $\{\mathbf{Se}_1, \ldots, \mathbf{Se}_n\}$ delle colonne di $\mathbf{S}$, che essendo $\mathbf{S}$ invertibile è una base di $\mathbb{C}^n$ (da 1.6.12, punto 2, con le colonne posso generare qualsiasi $\mathbf{b} \in \mathbb{C}^n$), è costituito da autovettori di $\mathbf{A}$ (dato che $\mathbf{A}(\mathbf{Se}_i) = d_i(\mathbf{Se}_i)$).
  Si osservi che in tal caso gli elementi diagonali di $\mathbf{D}$ sono gli autovalori di $\mathbf{A}$, ripetuti tante volte quanto è la loro molteplicità algebrica.

- la $2 \implies 4$: sia $\mathscr{B}$ una base di $\mathbb{C}^n$ costituita da autovettori di $\mathbf{A}$. Per ogni $i = 1, \ldots, k$ (con k numero di autovalori) sia s_i il numero degli elementi di $\mathscr{B}$ appartenenti a $E_{\mathbf{A}}(\lambda_i)$.
  Allora $s_i \leq d_i$, poiché ricordando che la molteplicità geometrica d_i è il numero di elementi di una base di $E_{\mathbf{A}}(\lambda_i)$, quei s_i possono esser sufficienti a generare $E_A(\lambda_i)$ o viceversa ne possono dover servire anche altri. Per il teorema 5.5.3, poi, si ha che $d_i \leq m_i$ per ogni $i = 1, \ldots, k$.
  Poiché $\mathscr{B}$ ha n elementi, allora $\sum_{i=1}^k s_i = n$. D'altra parte per la proposizione 5.4.6 si ha anche che $\sum_{i=1}^k m_i = n$ per cui $s_i = d_i = m_i$ per ogni $i = 1, \ldots, k$.
  In altre parole si sa che $s_i \leq d_i \leq m_i$; passando alle sommatorie e aggiungendo gli n ad inizio e fine:

$$n = \sum_{i=1}^k s_i \leq \sum_{i=1}^k d_i \leq \sum_{i=1}^k m_i = n$$

e dunque

$$\sum_{i=1}^{k} s_i = \sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i$$

da cui se $s_i \leq d_i \leq m_i$, $i = 1, \dots, k$ non può essere che $s_i = d_i = m_i$.

- per la 4 $\implies$ 3: se ogni autovalore λ_i ha molteplicità geometrica e algebrica uguali, $d_i = m_i$, allora $\sum_{i=1}^{k} m_i = \sum_{i=1}^{k} d_i$ per cui si ottiene:

$$\dim \mathbb{C}^n = n \overset{(1)}{=} \sum_{i=1}^{k} m_i = \sum_{i=1}^{k} d_i \overset{(2)}{=} \dim \left(\bigoplus_{i=1}^{k} E_{\mathbf{A}}(\lambda_i) \right)$$

  con (1) dalla proposizione 5.4.6 e (2) dato dal fatto che d_i è il numero di elementi della base di $E_{\mathbf{A}}(\lambda_i)$ e gli $E_{\mathbf{A}}(\lambda_i)$ sono sottospazi indipendenti (quindi le cardinalità delle basi si sommano per avere la cardinalità dello spazio generato dalla somma diretta).
  Pertanto da $\dim \mathbb{C}^n = \dim \left(\bigoplus_{i=1}^{k} E_{\mathbf{A}}(\lambda_i) \right)$ deriva $\mathbb{C}^n = \bigoplus_{i=1}^{k} E_{\mathbf{A}}(\lambda_i)$ per la 2.3.10

- la 3 $\implies$ 2: sia $\mathscr{B}_i$ una base di $E_{\mathbf{A}}(\lambda_i)$ per ogni $i = 1, \dots, k$. Allora $\mathscr{B} = \mathscr{B}_1 \sqcup \ldots \sqcup \mathscr{B}_k$ (con $\mathscr{B}$ di cardinalità n) è una base di $\bigoplus_{i=1}^{k} E_{\mathbf{A}}(\lambda_i) = \mathbb{C}^n$ (che ha base di cardinalità n) costituita da autovettori di $\mathbf{A}$ per la proposizione 2.3.13 (esercizio 5.17)

TODO: fixme

$\square$

Osservazione 293. Dal teorema precedente si deduce che le matrici di proiezione sono matrici diagonalizzabili (esempio 5.3.4) così come la matrice dell'esempio 5.3.8. Non è invece diagonalizzabile la matrice dell'esempio 5.3.7.

Osservazione 294. Osservando che se $\mathbf{A}$ è reale un autovalore λ di $\mathbf{A}$ è reale se e solo se gli autovettori di $\mathbf{A}$ relativi a λ possono essere scelti a coordinate reali, si ottiene il caso reale del teorema precedente.

Teorema 5.6.2. *Siano $\mathbf{A}$ una matrice reale $n \times n$ e $\lambda_1, \dots, \lambda_k$ i suoi autovalori complessi distinti. Sono equivalenti i seguenti fatti:*

1. *$\mathbf{A}$ è diagonalizzabile con una matrice reale*
2. *$\lambda_1, \dots, \lambda_k \in \mathbb{R}$ e $\mathbb{R}^n$ ha una base costituita da autovettori di $\mathbf{A}$*
3. *$\lambda_1, \dots, \lambda_k \in \mathbb{R}$ e $\mathbb{R}^n = \bigoplus_{i=1}^{k} (E_{\mathbf{A}}(\lambda_i) \cap \mathbb{R}^n)$*
4. *$\lambda_1, \dots, \lambda_k \in \mathbb{R}$ e per ogni $i = 1, \dots, k$ la dimensione del sottospazio $E_{\mathbf{A}}(\lambda_i) \cap \mathbb{R}^n$ di $\mathbb{R}^n$ coincide con la molteplicità algebrica m_i di λ_i*

Corollario 5.6.3. *Una matrice $n \times n$ con n autovalori distinti è diagonalizzabile*

Dimostrazione. Se gli n autovalori sono distinti hanno tutti molteplicità algebrica 1; per il teorema 5.5.3 dunque hanno molteplicità geometrica $d_i \leq m_i = 1$, ossia $d_i = m_i = 1$. Ma allora dato che le molteplicità di ciascun autovalore coincidono, allora per la 5.6.1 la matrice è diagonalizzabile $\square$

Osservazione 295. Naturalmente le matrici diagonali (e quindi in particolare le scalari), sono diagonalizzabili (basta porre $\mathbf{S} = \mathbf{I}$); per cui esistono matrici diagonalizzabili che non hanno autovalori distinti

Esempio 5.6.1. La matrice dell'esempio 5.3.5 è una matrice reale con autovalori complessi distinti e per il corollario 5.6.3 è diagonalizzabile come matrice complessa. Dal momento però che i suoi autovalori, i e $-i$ non sono reali, essa non è diagonalizzabile con una matrice reale.

Osservazione 296. Per passare da una matrice a una a essa simile, occorre calcolare l'inversa della matrice $\mathbf{S}$ che realizza la similitudine. In alcuni casi questo calcolo può essere particolarmente semplice, come quando, per esempio, si riduce a una trasposizione o a una H-trasposizione.

Definizione 5.6.3 (Matrice ortogonale)**.** $\mathbf{A}$ si dice *ortogonale* se $\mathbf{A}^{-1} = \mathbf{A}^T$

Definizione 5.6.4 (Matrice unitaria)**.** $\mathbf{A}$ si dice *unitaria* se $\mathbf{A}^{-1} = \mathbf{A}^H$

Osservazione 297. Ovviamente una matrice reale è ortogonale se e solo se è unitaria, ma in generale esistono matrici ortogonali che non sono unitarie e viceversa.

Osservazione 298. Le matrici reali ortogonali vengono chiamate in questo modo perché (come le complesse unitarie) l'insieme delle loro colonne è un insieme ortonormale di vettori di $\mathbb{R}^n$ (rispettivamente di $\mathbb{C}^n$). Si approfondirà questo nel capitolo 6.

Definizione 5.6.5 (Matrici ortogonalmente simili)**.** Due matrici $\mathbf{A}$ e $\mathbf{B}$ simili per le quali la matrice $\mathbf{S}$ che realizza la similitudine è ortogonale.

Definizione 5.6.6 (Matrici unitariamente simili)**.** Due matrici $\mathbf{A}$ e $\mathbf{B}$ simili per le quali la matrice $\mathbf{S}$ che realizza la similitudine è unitaria.

Definizione 5.6.7 (Similitudine ortogonale)**.** La relazione intercorrente tra due matrici ortogonalmente simili

Definizione 5.6.8 (Similitudine unitaria)**.** La relazione intercorrente tra due matrici unitariamente simili

Osservazione 299. Come visto, se $\mathbf{A}$ e $\mathbf{B}$ sono simili e $\mathbf{S}$ è la matrice che realizza la similitudine (ossia $\mathbf{A} = \mathbf{S}\mathbf{B}\mathbf{S}^{-1}$), allora $\mathbf{B}$ è la matrice associata a $f_\mathbf{A}$ rispetto alla base ordinata $\mathscr{B}$ di $\mathbb{C}^n$, avente come i-esimo vettore la i-esima colonna di $\mathbf{S}$. Se $\mathbf{A}$ e $\mathbf{B}$:

- sono reali e ortogonalmente simili la base $\mathscr{B}$ di $\mathbb{R}^n$ è una base ortonormale
- sono unitariamente simili la base $\mathscr{B}$ di $\mathbb{C}^n$ è una base ortonormale

Osservazione 300. Diamo ora una caratterizzazione delle matrici unitariamente simili a matrici triangolari.

Teorema 5.6.4 (Schur)**.** *Sia* $\mathbf{A}$ *una matrice in* $M_n(\mathbb{K})$*;* $\mathbf{A}$ *è unitariamente simile in* $M_n(\mathbb{K})$ *a una matrice triangolare superiore* $\mathbf{T}$ *(ossia* $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1} = \mathbf{U}\mathbf{T}\mathbf{U}^H$*) se e solo se tutti gli* n *autovalori di* $\mathbf{A}$ *sono elementi di* $\mathbb{K}$*.*

Osservazione 301. E dunque se $\mathbf{A}$ è *complessa* allora è unitariamente simile a una matrice triangolare[3]; viceversa, se *reale*, non è detto che sia ortogonalmente simile a una matrice triangolare.
Il teorema fondamentale dell'algebra assicura che ogni polinomio (caratteristico) complesso si fattorizzi in fattori lineari, e pertanto ogni matrice complessa di ordine n abbia n autovalori complessi (contati con la loro molteplicità). Le matrici reali di ordine n, invece, non necessariamente hanno n autovalori reali (anzi possono addirittura non averne alcuno, vedi esempio 5.3.5).

Dimostrazione. Partiamo dimostrando che se $\mathbf{A}$ è unitariamente simile allora gli n autovalori sono elementi di $\mathbb{K}$: sia $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$ con $\mathbf{U}$ unitaria e $\mathbf{T} = [t_{ij}] \in M_n(\mathbb{K})$ triangolare. Allora si ha

$$p_{\mathbf{A}}(X) \overset{(1)}{=} p_{\mathbf{T}}(X) = (t_{11} - X)(t_{22} - X)\dots(t_{nn} - X)$$

(con (1) per la proposizione 5.5.2) per cui gli autovalori di $\mathbf{A}$ sono i coefficienti diagonali di $\mathbf{T}$, e pertanto elementi di $\mathbb{K}$ (dato che $T \in M_n(\mathbb{K})$.

Viceversa si supponga che tutti gli n autovalori di $\mathbf{A} \in M_n(\mathbb{K})$ siano elementi di $\mathbb{K}$ e andiamo, per induzione su n, a provare che $\mathbf{A}$ è unitariamente simile a una matrice triangolare superiore.
Se $n = 1$ non c'è nulla da dimostrare: sarà $\mathbf{A} = [a]$, con $a \in \mathbb{K}$ (autovalore ed elemento), e se $\mathbf{U} = [u], \mathbf{U}^{-1} = [u^{-1}]$ (dato che $\mathbf{U}\mathbf{U}^{-1} = 1$) ci sarà un $\mathbf{T} = [t]$ tale che $[a] = [u][t][u^{-1}]$ e risolvendo tale t è proprio $t = a$; ossia $\mathbf{T} = \mathbf{A}$, ed essendo $\mathbf{A}$ matrice 1×1 è triangolare superiore.
Supponiamo allora $n > 1$ e che ogni matrice $(n-1) \times (n-1)$ a coefficienti in $\mathbb{K}$ soddisfacente l'ipotesi ($n-1$ autovalori $\in \mathbb{K}$) sia unitariamente simile in $M_{n-1}(\mathbb{K})$ a una matrice triangolare superiore. Siano λ_1 un autovalore di $\mathbf{A}$ e $\mathbf{v}_1$ un autovettore di $\mathbf{A}$ relativo a λ_1 (ossia si abbia $\mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$) con $\|\mathbf{v}_1\|_2 = 1$. Completando $\mathbf{v}_1$ a una base ortonormale $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ di $\mathbb{K}^n$, per la matrice composta di tali colonne ortonormali $\mathbf{U}_1 = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]$ (dal lemma 3.6.2) si ha che $\mathbf{U}_1^H\mathbf{U}_1 = \mathbf{I}$; ma allora per definizione di matrice inversa si ha che $\mathbf{U}_1^{-1} = \mathbf{U}_1^H$ e dunque $\mathbf{U}_1$ è unitaria.
Focalizzandosi sulla prima colonna di $\mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1$:

$$\mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1\mathbf{e}_1 \overset{(1)}{=} \mathbf{U}_1^{-1}\mathbf{A}\mathbf{v}_1 \overset{(2)}{=} \mathbf{U}_1^{-1}\lambda_1\mathbf{v}_1 = \lambda_1\mathbf{U}_1^{-1}\mathbf{v}_1 \overset{(3)}{=} \lambda_1\mathbf{e}_1$$

(dove in (1) si è sostituito con la prima colonna di $\mathbf{U}_1$, (2) dato che $\mathbf{A}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$, (3) dato che $\mathbf{U}_1^{-1}\mathbf{v}_1$ è la prima colonna di $\mathbf{U}_1^{-1}\mathbf{U}_1 = \mathbf{I}$). Per cui battezzando $\mathbf{A}' = \mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1$ e scrivendola isolandone la prima colonna si ha che:

$$\mathbf{A}' = \mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1 = \begin{bmatrix} \lambda_1 & \mathbf{w}^T \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$$

per opportune matrici $\mathbf{w}$ e $\mathbf{C}$. Ne segue che, dato che $\mathbf{A}'$ e $\mathbf{A}$ sono simili (con $\mathbf{U}_1^{-1}$ a realizzare la somiglianza, dato che $(\mathbf{U}_1^{-1})^{-1} = \mathbf{U}$, per 1.6.13) e quindi accomunati dal polinomio caratteristico:

$$p_{\mathbf{A}}(X) = p_{\mathbf{A}'}(X) = (\lambda_1 - X) \cdot p_{\mathbf{C}}(X)$$

[3]Questo verrà usato nel capitolo 6 per caratterizzare le matrici unitariamente diagonalizzabili, ossia unitariamente simili a una matrice diagonale

per cui gli autovalori di $\mathbf{C}$ sono autovalori di $\mathbf{A}$ e quindi, per ipotesi induttiva, elementi di $\mathbb{K}$.
Focalizzandoci sulla $\mathbf{C}$, per l'ipotesi induttiva esisterà una matrice unitaria $\mathbf{V} \in M_{n-1}(\mathbb{K})$ e una matrice triangolare superiore $\mathbf{T}_1 \in M_{n-1}(\mathbb{K})$ tali che

$$\mathbf{C} = \mathbf{V}\mathbf{T}_1\mathbf{V}^H$$

Posto $\mathbf{U}_2 = \mathbf{Diag}(1, \mathbf{V}) = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix}$ si ha che $\mathbf{U}_2$ è unitaria essendolo $\mathbf{V}$: infatti

$$\mathbf{U}_2\mathbf{U}_2^H = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \begin{bmatrix} \overline{1} & \overline{\mathbf{0}}^T \\ \overline{\mathbf{0}} & \mathbf{V}^H \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ 0 & \mathbf{V}\mathbf{V}^H \end{bmatrix} \overset{(1)}{=} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \mathbf{I}$$

(con (1) dato che $\mathbf{V}$ è unitaria quindi $\mathbf{V}^H = \mathbf{V}^{-1}$); pertanto $\mathbf{U}_2\mathbf{U}_2^H = \mathbf{I}$, ma allora $\mathbf{U}_2^H = \mathbf{U}_2^{-1}$ e $\mathbf{U}_2$ è unitaria.
Inoltre posto $\mathbf{U} = \mathbf{U}_1\mathbf{U}_2$, si ha che $\mathbf{U}$ è unitaria in quanto prodotto di due matrici unitarie (si vedrà ciò proposizione 6.1.2). Infine:

$$\begin{aligned}
\mathbf{A}' = \mathbf{U}^{-1}\mathbf{A}\mathbf{U} &\overset{(1)}{=} \mathbf{U}_2^{-1}\mathbf{U}_1^{-1}\mathbf{A}\mathbf{U}_1\mathbf{U}_2 = \mathbf{U}_2^{-1}\mathbf{A}'\mathbf{U}_2 = \mathbf{U}_2^{-1} \begin{bmatrix} \lambda_1 & \mathbf{w}^T \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \mathbf{U}_2 \\
&\overset{(2)}{=} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{V}^H \end{bmatrix} \begin{bmatrix} \lambda_1 & \mathbf{w}^T \\ \mathbf{0} & \mathbf{C} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{V} \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{w}^T \\ \mathbf{0} & \mathbf{V}^H\mathbf{C} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{V} \end{bmatrix} \\
&= \begin{bmatrix} \lambda_1 & \mathbf{w}^T\mathbf{V} \\ \mathbf{0} & \mathbf{V}^H\mathbf{C}\mathbf{V} \end{bmatrix} = \begin{bmatrix} \lambda_1 & \mathbf{w}^T\mathbf{V} \\ \mathbf{0} & \mathbf{T}_1 \end{bmatrix}
\end{aligned}$$

(con (1) dato che $\mathbf{U} = \mathbf{U}_1\mathbf{U}_2$ e $\mathbf{U}^{-1} = (\mathbf{U}_1\mathbf{U}_2)^{-1} = \mathbf{U}_2^{-1}\mathbf{U}_1^{-1}$, (2) dato che $\mathbf{U}_2^{-1} = \mathbf{U}_2^H$) è triangolare superiore essendolo $\mathbf{T}_1$. Ciò completa la dimostrazione. □

Proposizione 5.6.5. *Se* $\mathbf{A}$ *è unitariamente simile a una matrice triangolare superiore, è anche unitariamente simile a una matrice triangolare inferiore.*

Dimostrazione. Se $\mathbf{A}$ ha n autovalori $\in \mathbb{K}$ allora per le proposizioni 5.4.10 e 5.4.11 $\mathbf{A}^H$ ha ancora n autovalori $\in \mathbb{K}$; infatti l'unica cosa che li modifica nella H-trasposizione è il coniugio, ma se quelli di $\mathbf{A}$ sono reali anche i loro coniugati lo saranno e idem, rispettivamente, si avrà nel caso di autovalori di $\mathbf{A}$ complessi. Pertanto avendo n autovalori $\in \mathbb{K}$, $\mathbf{A}^H$ sarà, per il teorema di Schur, unitariamente simile ad una triangolare superiore $\mathbf{T}_1$ opportuna:

$$\mathbf{A}^H = \mathbf{U}\mathbf{T}_1\mathbf{U}^H$$

H-trasponendo entrambi i membri si ottiene

$$\mathbf{A} = \mathbf{U}\mathbf{T}_1^H\mathbf{U}^H$$

con $\mathbf{T}_1^H$ triangolare inferiore, quindi $\mathbf{A}$ è anche unitariamente simile ad una matrice triangolare inferiore.
Si ha quindi complessivamente il seguente corollario. □

Corollario 5.6.6. *Ogni matrice complessa è unitariamente simile a una matrice triangolare.*

Osservazione 302. Data una matrice $\mathbf{A}$ con polinomio caratteristico che si fattorizza in fattori lineari, la dimostrazione del teorema di Schur fornisce un procedimento effettivo per costruire una matrice triangolare $\mathbf{T}$ e una matrice unitaria $\mathbf{U}$ tali che $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$. Dal punto di vista computazionale, però, tale costruzione dipende dalla possibilità di calcolare gli autovalori di $\mathbf{A}$, e tale calcolo, come abbiamo già sottolineato, può essere estremamente problematico, se non impossibile.
Ma dal punto di vista teorico il teorema di Schur ha applicazioni notevoli, permettendo di ricondursi al caso di matrici triangolari ogniqualvolta si vogliano studiare delle proprietà delle matrici che siano invarianti rispetto alla relazione di similitudine.
Il corollario che segue è un esempio di applicazione teorica del teorema di Schur.

Corollario 5.6.7. *Sia $\mathbf{A}$ una matrice $n \times n$ con n autovalori (ripetuti con le loro molteplicità) $\lambda_1, \dots, \lambda_n$. Per ogni intero positivo k gli autovalori di $\mathbf{A}^k$ sono i numeri $\lambda_1^k, \dots, \lambda_n^k$.*

Dimostrazione. Per il teorema di Schur (corollario 5.6.6) esistono una matrice triangolare $\mathbf{T}$ e una matrice unitaria $\mathbf{U}$ tali che $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1}$ per cui si ha

$$\mathbf{A}^k = \mathbf{U}\mathbf{T}^k\mathbf{U}^{-1}$$

Infatti, procedendo per induzione:

- il passo base è verificato se $k = 1$ in quanto $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1}$
- nell'ipotesi induttiva che la formula valga per $n-1$ moltiplicando entrambi i membri per $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1}$ si ottiene

$$\begin{aligned} \mathbf{A}^{k-1}\mathbf{A} &= \mathbf{U}\mathbf{T}^{k-1}\mathbf{U}^{-1}\mathbf{U}\mathbf{T}\mathbf{U}^{-1} \\ \mathbf{A}^k &= \mathbf{U}\mathbf{T}^k\mathbf{U}^{-1} \end{aligned}$$

  e così si ritrova la formula generale.

Dalla proposizione 5.5.2, essendo matrici simili, segue allora che gli autovalori di $\mathbf{A}$ coincidono con quelli di $\mathbf{T}$, e gli autovalori di $\mathbf{A}^k$ coincidono con quelli di $\mathbf{T}^k$. Poiché anche $\mathbf{T}^k$ è triangolare, essendolo $\mathbf{T}$, e gli autovalori di una matrice triangolare sono i suoi coefficienti diagonali, allora $\lambda_1, \dots, \lambda_n$ sono i coefficienti diagonali di $\mathbf{T}$, e gli autovalori di $\mathbf{A}^k$ sono i coefficienti diagonali di $\mathbf{T}^k$. La conclusione segue quindi dal fatto che i coefficienti diagonali di $\mathbf{T}^k$ sono le potenze k-esime dei coefficienti diagonali di $\mathbf{T}$, essendo $\mathbf{T}$ triangolare. □

5.7 I teoremi di Hamilton-Cayley e di Gerschgorin

Osservazione 303. Dei molti risultati sul polinomio caratteristico di una matrice quadrata e sui suoi autovalori, due sono di basilare importanza: il teorema di Hamilton-Cayley e il teorema dei cerchi di Gerschgorin.

5.7.1 Hamilton-Cayley

Osservazione 304. Il primo mostra che, partendo dal polinomio caratteristico di una generica matrice $\mathbf{A}$ (ossia da una particolare equazione soddisfatta dagli autovalori di $\mathbf{A}$) si ottiene una equazione che deve essere soddisfatta dalla matrice $\mathbf{A}$ stessa.

Teorema 5.7.1 (Hamilton-Cayley)**.** *Sia* $\mathbf{A}$ *una matrice* $n \times n$ *con polinomio caratteristico* $p_{\mathbf{A}}(X) = (-1)^n X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0 X^0$*. Allora* $p_{\mathbf{A}}(\mathbf{A}) = \mathbb{O}$ *cioè*

$$(-1)^n \mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \ldots + a_1\mathbf{A} + a_0\mathbf{I} = \mathbb{O}$$

Dimostrazione. Dato che per il teorema fondamentale dell'algebra il polinomio è fattorizzabile generando n autovalori complessi, la matrice sia anch'essa complessa (al massimo se reale la si considera complessa, direi); allora per il teorema di Schur $\mathbf{A}$ è unitariamente simile a una matrice triangolare superiore $\mathbf{T} = [t_{ij}]$:

$$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1}$$

con $\mathbf{U}$ matrice unitaria ($\mathbf{U}^{-1} = \mathbf{U}^H$). Per la proposizione 5.5.2, essendo simili si ha che $p_{\mathbf{A}}(X) = p_{\mathbf{T}}(X)$ perciò basta provare che $p_{\mathbf{T}}(\mathbf{A}) = \mathbb{O}$. Osserviamo che

$$\begin{aligned}
p_{\mathbf{A}}(\mathbf{A}) &= (-1)^n\mathbf{A}^n + a_{n-1}\mathbf{A}^{n-1} + \ldots + a_1\mathbf{A} + a_0 \overbrace{\mathbf{A}^0}^{\mathbf{I}} \\
&= (-1)^n(\mathbf{U}\mathbf{T}\mathbf{U}^{-1})^n + a_{n-1}(\mathbf{U}\mathbf{T}\mathbf{U}^{-1})^{n-1} + \ldots + a_1(\mathbf{U}\mathbf{T}\mathbf{U}^{-1}) + a_0\mathbf{I} \\
&= (-1)^n\mathbf{U}\mathbf{T}^n\mathbf{U}^{-1} + a_{n-1}\mathbf{U}\mathbf{T}^{n-1}\mathbf{U}^{-1} + \ldots + a_1\mathbf{U}\mathbf{T}\mathbf{U}^{-1} + a_0\mathbf{U}\mathbf{I}\mathbf{U}^{-1} \\
&= \mathbf{U}\left[(-1)^n\mathbf{T}^n\mathbf{U}^{-1} + a_{n-1}\mathbf{T}^{n-1}\mathbf{U}^{-1} + \ldots + a_1\mathbf{T}\mathbf{U}^{-1} + a_0\mathbf{I}\mathbf{U}^{-1}\right] \\
&= \mathbf{U}\underbrace{\left[(-1)^n\mathbf{T}^n + a_{n-1}\mathbf{T}^{n-1} + \ldots + a_1\mathbf{T} + a_0\mathbf{I}\right]}_{p_{\mathbf{T}}(\mathbf{T})}\mathbf{U}^{-1}
\end{aligned}$$

quindi $p_{\mathbf{A}}(\mathbf{A}) = p_{\mathbf{T}}(\mathbf{A}) = \mathbf{U}p_{\mathbf{T}}(\mathbf{T})\mathbf{U}^{-1} = \mathbb{O}$ se $p_{\mathbf{T}}(\mathbf{T}) = \mathbb{O}$, ovvero, basta provare il teorema per la matrice triangolare $\mathbf{T}$. Essendo gli autovalori di $\mathbf{T}$ i suoi elementi diagonali, si ha che il polinomio caratteristico è:

$$p_{\mathbf{T}}(X) = (t_{11} - X)(t_{22} - X)\ldots(t_{nn} - X)$$

con t_{ii} elementi della diagonale di $\mathbf{T}$. Quindi facendo qualche passaggio algebrico

$$\begin{aligned}
p_{\mathbf{T}}(\mathbf{T}) &= \mathbb{O} \\
(t_{11} - \mathbf{T})(t_{22} - \mathbf{T})\ldots(t_{nn} - \mathbf{T}) &= \mathbb{O} \\
\mathbf{I}^n(t_{11} - \mathbf{T})(t_{22} - \mathbf{T})\ldots(t_{nn} - \mathbf{T}) &= \mathbf{I}^n\mathbb{O} \\
(t_{11}\mathbf{I} - \mathbf{T}\mathbf{I})(t_{22}\mathbf{I} - \mathbf{T}\mathbf{I})\ldots(t_{nn}\mathbf{I} - \mathbf{T}\mathbf{I}) &= \mathbb{O} \\
(t_{11}\mathbf{I} - \mathbf{T})(t_{22}\mathbf{I} - \mathbf{T})\ldots(t_{nn}\mathbf{I} - \mathbf{T}) &= \mathbb{O}
\end{aligned}$$

A questo punto raccolgo tutti i segni al primo membro; se n è pari siamo a posto così, se è dispari rimarrà fuori parentesi un -1 che elimino moltiplicando ambo i lati per -1. Siamo pertanto ricondotti a provare l'uguaglianza:

$$(\mathbf{T} - t_{11}\mathbf{I})(\mathbf{T} - t_{22}\mathbf{I})\ldots(\mathbf{T} - t_{nn}\mathbf{I}) = \mathbb{O} \tag{5.15}$$

Ora mostriamo che che gli n fattori di tale matrice commutano tra di loro; basta mostrare che ne commutano due (dopodiché tutti potranno commutare, facendo uno scambio di coppia alla volta) ossia che per ogni i e j

$$(\mathbf{T} - t_{ii}\mathbf{I})(\mathbf{T} - t_{jj}\mathbf{I}) = (\mathbf{T} - t_{jj}\mathbf{I})(\mathbf{T} - t_{ii}\mathbf{I})$$
$$\mathbf{T}^2 - t_{jj}\mathbf{T} - t_{ii}\mathbf{T} + t_{ii}t_{jj} = \mathbf{T}^2 - t_{ii}\mathbf{T} - t_{jj}\mathbf{T} + t_{ii}t_{jj}$$

ed abbiamo dimostrato che tali fattori commutano fra loro.
Per dimostrare l'uguaglianza in equazione 5.15 dobbiamo mostrare che la generica colonna j-esima della matrice $(\mathbf{T} - t_{11}\mathbf{I})(\mathbf{T} - t_{22}\mathbf{I}) \dots (\mathbf{T} - t_{nn}\mathbf{I})$ coincide con il vettore nullo. Tenuto conto del fatto che gli n fattori di tale matrice commutano tra di loro, basta provare che per ogni $j \leq n$ si ha:

TODO: Non mi è chiarissimo qui

$$(\mathbf{T} - t_{11}\mathbf{I})(\mathbf{T} - t_{22}\mathbf{I}) \dots (\mathbf{T} - t_{jj}\mathbf{I})\mathbf{e}_j = \mathbf{0} \tag{5.16}$$

Per $j = 1$ l'uguaglianza 5.16 è banale, perché la prima colonna della matrice $(\mathbf{T} - t_{11}\mathbf{I})$ è ovviamente nulla. Sia allora (1) vera per $j = 1, \dots, k-1$ e proviamo che è vera per k. Si ha infatti:

$$(\mathbf{T} - t_{kk}\mathbf{I})\mathbf{e}_k = \mathbf{T}\mathbf{e}_k - t_{kk}\mathbf{e}_k = t_{1k}\mathbf{e}_1 + t_{2k}\mathbf{e}_2 + \dots t_{k-1,k}\mathbf{e}_{k-1}$$

TODO: Non mi è chiarissimo qui

Da ciò si ricava, sfruttando sempre la commutatività dei fattori $(\mathbf{T} - t_{ii}\mathbf{I})(\mathbf{T} - t_{jj}\mathbf{I}) = (\mathbf{T} - t_{jj}\mathbf{I})(\mathbf{T} - t_{ii}\mathbf{I})$ per ogni i e j:

$$\begin{aligned}&(\mathbf{T} - t_{11}\mathbf{I}) \dots (\mathbf{T} - t_{k-1,k-1}\mathbf{I})(\mathbf{T} - t_{k,k}\mathbf{I})\mathbf{e}_k = \\ &\qquad (\mathbf{T} - t_{11}\mathbf{I}) \dots (\mathbf{T} - t_{k-1,k-1}\mathbf{I}) \cdot (t_{1k}\mathbf{e}_1 + t_{2k}\mathbf{e}_2 + \dots + t_{k-1,k}\mathbf{e}_{k-1}) = \mathbf{0}\end{aligned}$$

dato che $(\mathbf{T} - t_{11}\mathbf{I}) \dots (\mathbf{T} - t_{jj}\mathbf{I})\mathbf{e}_j = \mathbf{0}$ per $j = 1, 2, \dots, k-1$. □

Osservazione 305. Una immediata applicazione del teorema di Hamilton-Cayley si da al calcolo delle potenze di una matrice quadrata e al calcolo della sua inversa, nel caso esista (esercizi 5.26 e 5.27).

TODO: fixme

5.7.2 Gerschgorin

Osservazione 306. Il teorema dei cerchi di Gerschgorin localizza gli autovalori di una matrice $n \times n$ in una regione del piano complesso racchiusa in n cerchi, con centri e raggi deducibili dai coefficienti della matrice. Inoltre dice qualcosa di più nel caso in cui i cerchi siano disposti in modo opportuno.

Definizione 5.7.1 (i-esimo cerchio di Gerschgorin)**.** Data $\mathbf{A} = [a_{ij}] \in M_n(\mathbb{C})$ è chiamato così la regione del piano complesso

$$C_i(\mathbf{A}) = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i\}$$

con

$$R_i = \sum_{j \neq i} |a_{ij}|$$

Osservazione 307. A parole $C_i(\mathbf{A})$ è costituito dai complessi nel cerchio avente per centro l'i-esimo elemento diagonale di $\mathbf{A}$ e come raggio R_i la somma dei moduli degli elementi nella i-esima riga di $\mathbf{A}$ ad eccezione di a_{ii}.

Teorema 5.7.2 (Gerschgorin)**.** *Sia* $\mathbf{A} \in M_n(\mathbb{C})$*. Allora:*

1. *ogni autovalore di* $\mathbf{A}$ *appartiene ad almeno un cerchio di Gerschgorin;*
2. *se l'unione di* $k \leq n$ *cerchi di Gerschgorin forma una regione connessa e disgiunta dall'unione nei restanti* $n-k$ *cerchi, essa contiene esattamente* k *autovalori di* $\mathbf{A}$ *(contati con la loro molteplicità algebrica).*

Dimostrazione. Rispettivamente:

1. sia λ un autovalore di $\mathbf{A}$, per cui $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ per un autovettore $\mathbf{v} \neq \mathbf{0}$. Se $\mathbf{v} = [v_1 \ \ldots \ v_n]^T$, per ogni $i \leq n$ risulta (ci si concentra su una riga i del prodotto $\mathbf{A}\mathbf{v}$):
$$\sum_j a_{ij} v_j = \lambda v_i$$
da cui, dato che $\sum_j a_{ij} v_j = \sum_{j \neq i} a_{ij} v_j + a_{ii} v_i$ si ricava
$$(\lambda - a_{ii}) v_i = \sum_{j \neq i} a_{ij} v_j$$
e, passando ai moduli
$$|\lambda - a_{ii}|\,|v_i| = \left|\sum_{j \neq i} a_{ij} v_j\right| \leq \sum_{j \neq i} |a_{ij} v_j| = \sum_{j \neq i} |a_{ij}|\,|v_j|$$
e in conclusione
$$|\lambda - a_{ii}|\,|v_i| \leq \sum_{j \neq i} |a_{ij}|\,|v_j| \tag{5.17}$$
Si scelga una coordinata v_k di $\mathbf{v}$ di modulo massimo, tale cioè che $|v_k| \geq |v_j|$ per ogni j. Risulta $|v_k| > 0$, perché $\mathbf{v} \neq \mathbf{0}$. Dividendo ambo i membri della disuguaglianza 5.17 riferiti alla riga k-esima ($i = k$) per $|v_k|$ si ha
$$|\lambda - a_{kk}| \leq \sum_{j \neq k} |a_{kj}| \underbrace{\left|\frac{v_j}{v_k}\right|}_{<1} \leq \sum_{j \neq k} |a_{kj}| = R_k$$
il che dice appunto che $\lambda \in C_k(\mathbf{A})$
2. per una traccia di dimostrazione vedere gli esercizi 5.28-5.30 **TODO**: fixme

□

Osservazione 308. Un'utile applicazione del teorema 5.7.2 segue

Definizione 5.7.2 (Matrici strettamente diagonalmente dominanti)**.** Matrici $\mathbf{A} = [a_{ij}] \in M_n(\mathbb{C})$ tali che, per ogni $i \leq n$ risulta
$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

Corollario 5.7.3. *Una matrice* $\mathbf{A} \in M_n(\mathbb{C})$ *strettamente diagonalmente dominante è invertibile.*

Dimostrazione. Nessuno dei cerchi di Gerschgorin di $\mathbf{A}$ contiene l'origine del piano complesso, perché i centri a_{ii} hanno distanza dall'origine maggiore del raggio R_i. Allora 0 non è un autovalore di $\mathbf{A}$ cioè l'equazione $\mathbf{A}\mathbf{x} = \mathbf{0}$ ha solo la soluzione nulla. Per il teorema 1.6.12 $\mathbf{A}$ risulta invertibile. □

Capitolo 6

Matrici normali

Contents

6.1 Matrici unitarie e similitudini unitarie

Osservazione 309. Tra le matrici complesse invertibili sono di particolare interesse le matrici $\mathbf{A}$ complesse normali ($\mathbf{A}^{-1} = \mathbf{A}^H$) e reali ortogonali ($\mathbf{A}^{-1} = \mathbf{A}^T$) dato che per calcolarne l'inversa è sufficiente trasporre e coniugare.
Allo stesso modo le similitudini più interessanti sono quelle unitarie (o ortogonali).

Osservazione 310. Una matrice complessa $\mathbf{A}$ può essere ortogonale (cioè $\mathbf{A}^{-1} = \mathbf{A}^T$), ma in tal caso è di minore interesse rispetto alle matrici unitarie.

6.1.1 Matrici unitarie

Osservazione 311. Raccogliamo nel seguente teorema le principali proprietà delle matrici unitarie (per le reali, basta sostituire la parole "unitaria" con la parola "ortogonale", $\mathbb{C}^n$ con $\mathbb{R}^n$ e l'H-trasposizione con la trasposizione).

Teorema 6.1.1. *Per una matrice complessa* $\mathbf{U}$ *di tipo* $n\times n$ *le proprietà seguenti sono equivalenti:*

1. $\mathbf{U}$ *è unitaria:* $\mathbf{U}^{-1} = \mathbf{U}^H$*;*
2. $\mathbf{U}\mathbf{U}^H = \mathbf{I}_n = \mathbf{U}^H\mathbf{U}$*;*
3. $\mathbf{U}^H$ *è unitaria;*
4. *le colonne di* $\mathbf{U}$ *formano una base ortonormale di* $\mathbb{C}^n$*;*
5. *le righe di* $\mathbf{U}$ *formano una base ortonormale di* $\mathbb{C}_n$*;*
6. *per ogni coppia di vettori* $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$*, il loro prodotto interno* $\mathbf{u}^H\mathbf{v}$ *coincide con quello dei vettori trasformati mediante pre-moltiplicazione:* $(\mathbf{U}\mathbf{u})^H(\mathbf{U}\mathbf{v})$*;*
7. *per ogni vettore* $\mathbf{v} \in \mathbb{C}^n$ *la norma euclidea di* $\mathbf{v}$ *coincide con quella di* $\mathbf{U}\mathbf{v}$

Dimostrazione. Per:

- 1 $\iff$ 2: si ha $\mathbf{U}\mathbf{U}^H = \mathbf{U}\mathbf{U}^{-1} = \mathbf{I} = \mathbf{U}^{-1}\mathbf{U} = \mathbf{U}^H\mathbf{U}$
- 1 $\iff$ 3: si ha che $\mathbf{U}^{-1} = \mathbf{U}^H \iff (\mathbf{U}^{-1})^H = (\mathbf{U}^H)^H$
- 2 $\iff$ 4: dire che le colonne di $\mathbf{U} = [\mathbf{u}_1 \ \dots \ \mathbf{u}_n]$ formano una base ortonormale di $\mathbb{C}^n$ equivale a dire che $\mathbf{u}_i^H\mathbf{u}_j = \delta_{ij}$ (simbolo di Kronecker, pari 1 se $i = j$, 0 altrimenti: ossia il prodotto di due vettori ortonormali è 1 se questi coincidono, altrimenti sono ortogonali ed è quindi 0). Ciò equivale a dire, prendendo tutti i prodotti congiuntamente, che $\mathbf{U}^H\mathbf{U} = \mathbf{I}_n$ (e quindi ci si è riallacciati al punto 2)
- 3 $\iff$ 5: applicando la precedente dimostrazione a $\mathbf{U}^H$, le righe di $\mathbf{U} = \begin{bmatrix}\mathbf{u}_1\\ \dots\\ \mathbf{u}_n\end{bmatrix}$ sono le colonne di $\mathbf{U}^H = \begin{bmatrix}\mathbf{u}_1^H & \dots & \mathbf{u}_n^H\end{bmatrix}$. Le colonne di $\mathbf{U}^H$ formano una base ortonormale di $\mathbb{C}^n$ se e solo se $\mathbf{u}_i\mathbf{u}_j = \delta_{ij}$ ossia se $\mathbf{U}\mathbf{U}^H = \mathbf{I}$
- 2 $\implies$ 6: si ha $(\mathbf{U}\mathbf{u})^H(\mathbf{U}\mathbf{v}) = \mathbf{u}^H\mathbf{U}^H\mathbf{U}\mathbf{v} = \mathbf{u}^H\mathbf{I}_n\mathbf{v} = \mathbf{u}^H\mathbf{v}$
- 6 $\implies$ 7: si ha $\|\mathbf{v}\|_2^2 = \mathbf{v}^H\mathbf{v} = \mathbf{v}^H\mathbf{U}^H\mathbf{U}\mathbf{v} = (\mathbf{U}\mathbf{v})^H\mathbf{U}\mathbf{v} = \|\mathbf{U}\mathbf{v}\|_2^2$
- 7 $\implies$ 2: data una generica matrice $\mathbf{A} = [a_{ij}]$ si ha che $a_{ij} = \mathbf{e}_i^T\mathbf{A}\mathbf{e}_j$ (scelgo prima l'i-esima riga e poi la j-esima colonna). Poiché $\mathbf{e}_i^T\mathbf{U}^H$ è la i-esima riga di $\mathbf{U}^H$ e $\mathbf{U}\mathbf{e}_j$ è la j-esima colonna di $\mathbf{U}$, dall'ipotesi si ha che gli elementi della diagonale di $\mathbf{A} = \mathbf{U}^H\mathbf{U}$ sono

$$\mathbf{e}_i^H\mathbf{U}^H\mathbf{U}\mathbf{e}_i = \|\mathbf{U}\mathbf{e}_i\|_2^2 \overset{(1)}{=} \|\mathbf{e}_i\|_2^2 = 1$$

(con (1) dovuto alla proprietà 7) ossia la matrice $\mathbf{A} = \mathbf{U}^H\mathbf{U}$ ha gli elementi diagonali uguali a 1[1]. Se poi $i \neq j$ la seguente scrittura risulta:

$$\begin{aligned}[\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j)]^H [\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j)] = (\mathbf{e}_i^H + \mathbf{e}_j^H)\mathbf{U}^H\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j) = \|\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j)\|_2^2 \\ \stackrel{(1)}{=} \|\mathbf{e}_i + \mathbf{e}_j\|_2^2 = 2\end{aligned}$$

(con (1) dovuto alla proprietà 7), ma al contempo si ha che:

$$\begin{aligned}[\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j)]^H [\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j)] &= (\mathbf{e}_i^H + \mathbf{e}_j^H)\mathbf{U}^H\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j) \stackrel{(1)}{=} (\mathbf{e}_i^H + \mathbf{e}_j^H)\mathbf{A}(\mathbf{e}_i + \mathbf{e}_j) \\ &= \mathbf{e}_i^H\mathbf{A}\mathbf{e}_i + \mathbf{e}_j^H\mathbf{A}\mathbf{e}_j + \mathbf{e}_i^H\mathbf{A}\mathbf{e}_j + \mathbf{e}_j^H\mathbf{A}\mathbf{e}_i \\ &= 1 + 1 + \underbrace{\mathbf{e}_i^H\mathbf{A}\mathbf{e}_j}_{a_{ij}} + \underbrace{\mathbf{e}_j^H\mathbf{A}\mathbf{e}_i}_{a_{ji}}\end{aligned}$$

(con (1) ricordando che abbiamo definito $\mathbf{A} = \mathbf{U}^H\mathbf{U}$). Uguagliando le ultime due risulta

$$a_{ij} + a_{ji} = 0 \tag{6.1}$$

Ma tenuto conto che $\mathbf{A}$ è hermitiana, in quanto

$$\mathbf{A} = \mathbf{U}^H\mathbf{U} = (\mathbf{U}^H\mathbf{U})^H = \mathbf{A}^H$$

si ha che $a_{ij} = \overline{a_{ji}}$ e in particolare $a_{ji} = \overline{a_{ij}}$ per cui sostituendo questa seconda nella 6.1 giungiamo ad avere $a_{ij} + \overline{a_{ij}} = 0$.
Similmente, considerando al posto del vettore $\mathbf{e}_i + \mathbf{e}_j$, il vettore $\mathbf{e}_i + i\mathbf{e}_j$ (facendo attenzione a non fare confusione tra l'indice e il numero immaginario) si ha rispettivamente

$$[\mathbf{U}(\mathbf{e}_i + i\mathbf{e}_j)]^H [\mathbf{U}(\mathbf{e}_i + \mathbf{e}_j)] = \|\mathbf{U}(\mathbf{e}_i + i\mathbf{e}_j)\|_2^2 = \|\mathbf{e}_i + i\mathbf{e}_j\|_2^2 = 2$$

e

$$\begin{aligned}[\mathbf{U}(\mathbf{e}_i + i\mathbf{e}_j)]^H [\mathbf{U}(\mathbf{e}_i + i\mathbf{e}_j)] &= (\mathbf{e}_i^T - i\mathbf{e}_j^T)\mathbf{U}^H\mathbf{U}(\mathbf{e}_i + i\mathbf{e}_j) = (\mathbf{e}_i^T - i\mathbf{e}_j^T)\mathbf{A}(\mathbf{e}_i + i\mathbf{e}_j) \\ &= \mathbf{e}_i^T\mathbf{A}\mathbf{e}_i + \mathbf{e}_i^T\mathbf{A}i\mathbf{e}_j - i\mathbf{e}_j^T\mathbf{A}\mathbf{e}_i + -i\mathbf{e}_j^T\mathbf{A}i\mathbf{e}_i \\ &= \underbrace{a_{ii}}_{=1} + ia_{ij} - ia_{ji} + \underbrace{a_{jj}}_{=1} \\ &= 2 + i(a_{ij} - a_{ji})\end{aligned}$$

Dunque

$$0 = i(a_{ij} - a_{ji}) \iff 0 = -(a_{ij} - a_{ji}) \iff a_{ij} - a_{ji} = 0 \stackrel{(1)}{\iff} a_{ij} - \overline{a_{ij}} = 0$$

con (1) dato che $\mathbf{A}$ è hermitiana. Pertanto si ha anche $a_{ij} - \overline{a_{ij}} = 0$. Dunque, riassumendo, da tutto questo deriva che:

$$\begin{cases} a_{ij} + \overline{a_{ij}} = 0 \\ a_{ij} - \overline{a_{ij}} = 0 \end{cases} \begin{cases} a_{ij} = -\overline{a_{ij}} \\ a_{ij} + a_{ij} = 0 \end{cases} \implies 2a_{ij} = 0 \implies a_{ij} = 0$$

ossia $a_{ij} = 0$ (se $i \neq j$), quindi $\mathbf{U}^H\mathbf{U} = \mathbf{A} = \mathbf{I}$ (coincide con la matrice identità).

[1] Noi sapremmo che $\mathbf{U}^H\mathbf{U} = \mathbf{I}$ ma non sfruttiamo tale conoscenza poiché ci basiamo solo su 7

□

Definizione 6.1.1 (Isometria)**.** Una trasformazione lineare $f : \mathbb{C}^m \to \mathbb{C}^n$ è detta una isometria se conserva le norme euclidee dei vettori, cioè se

$$\|\mathbf{v}\|_2 = \|f(\mathbf{v})\|, \quad \forall \mathbf{v} \in \mathbb{C}^m \tag{6.2}$$

Osservazione 312. Il teorema 6.1.1 ($7 \iff 1$) dice che la trasformazione lineare di $\mathbb{C}^n$ in sé indotta da una matrice è una isometria se e solo se la matrice unitaria.

Proposizione 6.1.2. *L'insieme delle matrici unitarie $n \times n$ è chiuso rispetto al passaggio alle inverse e al prodotto, e contiene la matrice identità* $\mathbf{I}$*.*

Dimostrazione. Per la chiusura rispetto ad inverse, se $\mathbf{U}$ è unitaria, tale è $\mathbf{U}^H = \mathbf{U}^{-1}$ per il teorema 6.1.1 (punto 3); se anche $\mathbf{V}$ è unitaria si ha

$$(\mathbf{UV})^H = \mathbf{V}^H\mathbf{U}^H = \mathbf{V}^{-1}\mathbf{U}^{-1} = (\mathbf{UV})^{-1}$$

da cui la chiusura rispetto al prodotto. Il fatto che $\mathbf{I}_n$ è unitaria:

$$\mathbf{I}^H = \mathbf{I} = \mathbf{I}^{-1}$$

□

Osservazione 313. Con terminologia algebrica, la proposizione 6.1.2 dice che le matrici unitarie di ordine n formano un gruppo, che è chiamato gruppo unitario $n \times n$; analogamente, le matrici reali ortogonali di ordine n formano un gruppo, che è chiamato gruppo ortogonale $n \times n$.

Proposizione 6.1.3. *Se λ è un autovalore di una matrice unitaria* $\mathbf{U}$*, allora* $|\lambda| = 1$

Dimostrazione. Sia $\mathbf{Uv} = \lambda\mathbf{v}$ per un opportuno vettore non nullo $\mathbf{v}$. Si ha allora:

$$\mathbf{v}^H\mathbf{v} = \mathbf{v}^H\mathbf{I}\mathbf{v} = \mathbf{v}^H\underbrace{\mathbf{U}^H\mathbf{U}}_{\mathbf{I}}\mathbf{v} = (\mathbf{Uv})^H(\mathbf{Uv}) = (\lambda\mathbf{v})^H(\lambda\mathbf{v}) \overset{(1)}{=} \overline{\lambda}\lambda\mathbf{v}^H\mathbf{v}$$

dove (1) deriva dalla proprietà del coniugio (coniugio del prodotto è il prodotto dei coniugati).
Pertanto dato che

$$\mathbf{v}^H\mathbf{v} = \overline{\lambda}\lambda\mathbf{v}^H\mathbf{v} \iff \mathbf{v}^H\mathbf{v}(1 - \overline{\lambda}\lambda) = 0$$

per cui $\mathbf{v}^H\mathbf{v} = 0 \vee 1 - \overline{\lambda}\lambda = 0$ ma essendo $\|\mathbf{v}\| \neq 0$ deve essere $\mathbf{v}^H\mathbf{v} \neq 0$ dunque

$$1 - \overline{\lambda}\lambda = 0 \iff 1 = \overline{\lambda}\lambda = |\lambda|^2$$

da cui $|\lambda| = 1$ □

Osservazione 314. Nel seguito vedremo anche l'implicazione contraria, a testimoniare l'equivalenza logica, ossia che una matrice è unitaria se e solo se i suoi autovalori hanno modulo 1.

6.1.2 Diagonalizzazioni unitarie

Osservazione 315. Il significato geometrico delle diagonalizzazioni unitarie è chiarito dal seguente risultato.

Teorema 6.1.4. *Per una matrice quadrata* $\mathbf{A}$ *di ordine* n*, i fatti seguenti sono equivalenti:*

1. $\mathbf{A}$ *è unitariamente diagonalizzabile:*

$$\mathbf{A} = \mathbf{UDU}^{-1} = \mathbf{UDU}^H$$

con $\mathbf{D}$ *diagonale e* $\mathbf{U}$ *unitaria;*

2. $\mathbb{C}^n$ *ha una base ortonormale costituita da autovettori di* $\mathbf{A}$*;*

3. $\mathbf{A}$ *è diagonalizzabile (anche senza l'utilizzo di matrici unitarie) e autospazi relativi ad autovettori distinti sono tra di loro ortogonali.*

Dimostrazione. Rispettivamente:

- $1 \implies 2$: se $\mathbf{A}$ è unitariamente diagonalizzabile, $\mathbf{A} = \mathbf{UDU}^{-1}$ con $\mathbf{D}$ diagonale e $\mathbf{U}$ unitaria ($\mathbf{U}^{-1} = \mathbf{U}^H$) si ha

$$\begin{aligned} \mathbf{A} = \mathbf{UDU}^{-1} &\iff \mathbf{AU} = \mathbf{UD} \\ &\iff \mathbf{AUe}_i = \mathbf{U}\underbrace{\mathbf{De}_i}_{=d_i\mathbf{e}_i}, \quad \forall 1 \le i \le n \\ &\iff \mathbf{AUe}_i = d_i\mathbf{Ue}_i, \quad \forall 1 \le i \le n \end{aligned}$$

  L'insieme $\{\mathbf{Ue}_1, \dots, \mathbf{Ue}_n\}$ delle colonne di $\mathbf{U}$ è una base ortonormale di $\mathbb{C}^n$ (per teorema 6.1.1, punto 4) ed è costituito da autovettori di $\mathbf{A}$, dato che $\mathbf{A}(\mathbf{Ue}_i) = d_i(\mathbf{Ue}_i)$

- $2 \implies 3$: il fatto che $\mathbf{A}$ sia diagonalizzabile consegue ancora dal teorema 5.6.1 ($2 \iff 1$). Gli autovettori della base (colonne della matrice) relativi a un medesimo autovalore λ formano una base dell'autospazio $E_{\mathbf{A}}(\lambda)$; ne consegue che, se μ è un autovalore diverso da λ, $E_{\mathbf{A}}(\lambda)$ ed $E_{\mathbf{A}}(\mu)$ sono generati da vettori a due a due ortogonali, perciò ogni vettore in $E_{\mathbf{A}}(\lambda)$ è ortogonale a ogni vettore in $E_{\mathbf{A}}(\mu)$

- $3 \implies 1$: scegliendo in ogni autospazio una base ortonormale, l'ipotesi assicura che l'insieme che si ottiene dall'unione di tali basi costituisce una base ortonormale di $\mathbb{C}^n$, quindi la matrice che ha come colonne tale base e che diagonalizza la matrice $\mathbf{A}$ è unitaria.

□

Osservazione 316. La base di $\mathbb{C}^n$ composta da autovettori della matrice $\mathbf{A}$ (unitariamente diagonalizzabile) deve essere costruita in maniera ortonormale, altrimenti non è detto che la matrice che crea la similitudine risulti unitaria.
Il teorema 6.1.4 assicura che risultano tra di loro ortogonali solo autovettori appartenenti ad autospazi distinti.
Pertanto affinché la base di autovettori risulti ortonormale occorre scegliere ogni singola base di ciascun autospazio ortonormale.

Esempio 6.1.1. La matrice diagonale $\mathbf{D} = \mathbf{Diag}(-1, -1, 2)$ è banalmente unitariamente diagonalizzabile. I suoi autovalori sono $\lambda_1 = -1$ con autospazio $E_{\mathbf{A}}(\lambda_1)$ generato dai vettori $\mathbf{e}_1$ ed $\mathbf{e}_2$ e $\lambda_2 = 2$ con autospazio $E_{\mathbf{A}}(\lambda_2)$ generato dal vettore $\mathbf{e}_3$.
Se scegliamo come base di $E_{\mathbf{A}}(\lambda_1)$ i vettori $[1\ 2\ 0]^T$ e $[2\ 1\ 0]^T$ (non ortonormali) e come base di $E_{\mathbf{A}}(\lambda_2)$ il vettore $[0\ 0\ 3]^T$ (non ortonormale), abbiamo la seguente diagonalizzazione di $\mathbf{D} = \mathbf{U}\mathbf{D}_1\mathbf{U}^{-1}$ con

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \mathbf{D}_1 = \mathbf{D} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Per $\mathbf{U}^{-1}$ si ha che

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 0 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} 2 & 4 & 0 & 2 & 0 & 0 \\ 0 & -3 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/3 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2/3 & -1/3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & -1/3 & 2/3 & 0 \\ 0 & 1 & 0 & 2/3 & -1/3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/3 \end{bmatrix}$$

pertanto

$$\mathbf{U}^{-1} = \begin{bmatrix} -1/3 & 2/3 & 0 \\ 2/3 & -1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$

Quindi la diagonalizzazione è

$$\mathbf{D} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} -1/3 & 2/3 & 0 \\ 2/3 & -1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$

ma qui $\mathbf{U}$ non è unitaria, dato che

$$\mathbf{U}^{-1} = \begin{bmatrix} -1/3 & 2/3 & 0 \\ 2/3 & -1/3 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} \neq \mathbf{U}^H = \mathbf{U}^T = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \mathbf{U}$$

6.1.3 Esempi di matrici ortogonali e unitarie

6.1.3.1 Matrici di permutazione

Definizione 6.1.2 (Matrice di permutazione)**.** Una matrice di permutazione $\mathbf{P}$ è definita come un prodotto di matrici di trasposizione $\mathbf{E}_{ij}$

Osservazione 317. Si tratta di una matrice che ha in ogni riga e in ogni colonna un unico 1 e tutti gli altri coefficienti uguali a 0.

Osservazione 318. Le matrici di permutazione intervengono nell'ambito dell'eliminazione di Gauss quando bisogna (meramente) scambiare le righe di una matrice per avere dei pivot non nulli.

Proposizione 6.1.5. *Le matrici di permutazioni* $\mathbf{P}$ *sono matrici reali ortogonali:* $\mathbf{P}^{-1} = \mathbf{P}^T$

Dimostrazione. $\mathbf{P}$ è invertibile poiché prodotto di matrici $\mathbf{E}_{ij}$ invertibili, con $\mathbf{E}_{ij}^{-1} = \mathbf{E}_{ij}^{T}$, pertanto

$$\mathbf{P}^{-1} = (\mathbf{E}_1 \dots \mathbf{E}_n)^{-1} = \mathbf{E}_n^{-1} \dots \mathbf{E}_1^{-1} = \mathbf{E}_n^{T} \dots \mathbf{E}_1^{T} = (\mathbf{E}_1 \dots \mathbf{E}_n)^{T} = \mathbf{P}^{T}$$

□

Definizione 6.1.3 (Cogredienza)**.** Data una matrice di permutazione $\mathbf{P}$ e una qualunque matrice del medesimo ordine $\mathbf{A}$ è chiamata così l'operazione che trasforma la matrice $\mathbf{A}$ nella matrice $\mathbf{PAP}^T$

Definizione 6.1.4 (Matrice cogrediente)**.** La matrice $\mathbf{B} = \mathbf{PAP}^T$ si dice cogrediente alla matrice $\mathbf{A}$.

Definizione 6.1.5 (Congruenza)**.** L'operazione che tramite una matrice invertibile $\mathbf{S}$ trasforma la matrice $\mathbf{A}$ nella matrice $\mathbf{SAS}^T$

Osservazione 319*.* Una cogredienza è sia una similitudine che una congruenza.

Definizione 6.1.6 (Matrice doppiamente stocastica)**.** Matrice i cui coefficienti sono numeri reali non negativi tali che la somma di ogni riga e di ogni colonna è uguale a 1.

Osservazione 320*.* La matrici di permutazione sono ovviamente doppiamente stocastiche.

Osservazione 321*.* Il teorema di Birkhoff (**??**) asserisce che le matrici doppiamente stocastiche sono esattamente le combinazioni convesse delle matrici di permutazione

6.1.3.2 Matrici di rotazione

Definizione 6.1.7 (Matrici di rotazione)**.** Matrice 2×2 del tipo

$$\mathbf{R}_\alpha = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

dove $0 \leq \alpha < 2\pi$.

Proposizione 6.1.6. *Una matrice di permutazione $\mathbf{R}_\alpha$ è una matrice ortogonale.*

Dimostrazione. Dalla relazione trigonometrica $\cos^2\alpha + \sin^2\alpha = 1$ segue che

$$\begin{aligned} \mathbf{R}_\alpha \mathbf{R}_\alpha^T &= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \\ &= \begin{bmatrix} \cos^2\alpha + \sin^2\alpha & \cos\alpha\sin\alpha - \sin\alpha\cos\alpha \\ \sin\alpha\cos\alpha - \cos\alpha\sin\alpha & \sin^2\alpha + \cos^2\alpha \end{bmatrix} = \mathbf{I}_2 \end{aligned}$$

Analogamente si ha che $\mathbf{R}_\alpha^T \mathbf{R}_\alpha = \mathbf{I}_2$, pertanto $\mathbf{R}_\alpha^T = \mathbf{R}_\alpha^{-1}$ ed $\mathbf{R}_\alpha$ è una matrice ortogonale. □

Osservazione 322*.* È un facile esercizio di geometria elementare provare che la pre-moltiplicazione per $\mathbf{R}_\alpha$ in $\mathbb{R}^2$ ha l'effetto di ruotare i vettori di α radianti in senso antiorario.

Definizione 6.1.8 (Matrice di rotazione generalizzata)**.** Matrice $n \times n$ $(n > 2)$ contenente come sottomatrice principale una matrice di rotazione e che al di fuori di tale sottomatrice coincide con la matrice identica.

Esempio 6.1.2. Una matrice 5×5 con sottomatrice principale in righe/colonne 2 e 4, ed una 4×4 con sottomatrice principale seconda

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \cos\alpha & 0 & -\sin\alpha & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \sin\alpha & 0 & \cos\alpha & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 & 0 \\ \sin\alpha & \cos\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Osservazione 323. Una matrice di rotazione generalizzata ha nei posti i e j $(i \neq j)$ della diagonale coefficienti uguali a $\cos\alpha$, nel posto (i, j) coefficiente uguale a $-\sin\alpha$, nel posto (j, i) coefficiente uguale a $\sin\alpha$ e nei restanti posti coincide con $\mathbf{I}_n$.

Proposizione 6.1.7. *Le matrici di rotazione generalizzate sono matrici ortogonali.*

Dimostrazione. Sia $\mathbf{A}$ matrice di rotazione generalizzata. Si ha che $\mathbf{A}\mathbf{A}^T = [\mathbf{r}_i\mathbf{r}_j^T]$ con $\mathbf{r}_i$ denotiamo la riga i-esima di $\mathbf{A}$. Si ha che se $i = j$ (ossia sulla diagonale della matrice risultante) $\mathbf{r}_i\mathbf{r}_i^T = \cos^2\alpha + \sin^2\alpha = 1$ oppure $\mathbf{r}_i\mathbf{r}_i^T = 1^2 = 1$, a seconda della riga considerata. Invece se $i \neq j$ si ha $\mathbf{r}_i\mathbf{r}_j^T = \cos\alpha\sin\alpha - \sin\alpha\cos\alpha = 0$ oppure $\mathbf{r}_i\mathbf{r}_j^T = 0$. Quindi $\mathbf{A}\mathbf{A}^T = \mathbf{I}$. Allo stesso modo si ha che $\mathbf{A}^T\mathbf{A} = \mathbf{I}$. Dunque $\mathbf{A}^{-1} = \mathbf{A}^T$. □

Osservazione 324. La pre-moltiplicazione per una tale matrice in $\mathbb{R}^n$ ha l'effetto di ruotare la proiezione di un vettore sul sottospazio $\langle \mathbf{e}_i; \mathbf{e}_j \rangle$ di α radianti in senso anti-orario e di lasciare fissa la proiezione sul sottospazio complemento ortogonale di $\langle \mathbf{e}_i; \mathbf{e}_j \rangle$.

6.1.3.3 Matrici di Householder

Definizione 6.1.9. Sia $\mathbf{w} \in \mathbb{C}^n$ un vettore di norma euclidea 1: la matrice

$$\mathbf{H}_\mathbf{w} = \mathbf{I}_n - 2\mathbf{w}\mathbf{w}^H$$

si chiama matrice di Householder associata a $\mathbf{w}$.

Osservazione 325. $\mathbf{H}_\mathbf{w}$ è una matrice $n \times n$ complessa e, se $\mathbf{w} \in \mathbb{R}^n$, è una matrice reale.

Osservazione 326. In quel che segue torna utile $\mathbf{w}^H\mathbf{w} = 1$ dato che essendo il vettore $\mathbf{w}$ di norma unitaria si ha $\sqrt{\mathbf{w}^H\mathbf{w}} = 1$

Proposizione 6.1.8. $\mathbf{H}_\mathbf{w}$ *è hermitiana e unitaria.*

Dimostrazione. È hermitiana in quanto:

$$\mathbf{H}_\mathbf{w}^H = (\mathbf{I} - 2\mathbf{w}\mathbf{w}^H)^H = \mathbf{I}^H - (2\mathbf{w}\mathbf{w}^H)^H = \mathbf{I}^H - 2\mathbf{w}\mathbf{w}^H = \mathbf{H}_\mathbf{w}$$

È unitaria in quanto:

$$\begin{aligned}
\mathbf{H}_{\mathbf{w}}^H \mathbf{H}_{\mathbf{w}} = \mathbf{H}_{\mathbf{w}} \mathbf{H}_{\mathbf{w}}^H = \mathbf{H}_{\mathbf{w}}^2 &= (\mathbf{I}_n - 2\mathbf{w}\mathbf{w}^H)(\mathbf{I}_n - 2\mathbf{w}\mathbf{w}^H) \\
&= \mathbf{I}_n - 2\mathbf{w}\mathbf{w}^H - 2\mathbf{w}\mathbf{w}^H + 4\mathbf{w}\underbrace{\mathbf{w}^H\mathbf{w}}_{=1}\mathbf{w}^H \\
&= \mathbf{I}_n - 4\mathbf{w}\mathbf{w}^H + 4\mathbf{w}\mathbf{w}^H \\
&= \mathbf{I}_n
\end{aligned}$$

□

Proposizione 6.1.9. *Si ha che* $\det(\mathbf{H}_{\mathbf{w}}) = -1$.

Dimostrazione. Sia $\mathbf{A} = [\mathbf{w}\ \mathbf{v}_2\ \dots\ \mathbf{v}_n]$ la matrice con prima colonna il vettore $\mathbf{w}$ e successive colonne i vettori di una base ortonormale dello spazio $\langle \mathbf{w} \rangle^{\perp}$. Calcoliamo le colonne della matrice $\mathbf{H}_{\mathbf{w}}\mathbf{A}$:

$$\begin{aligned}
\mathbf{H}_{\mathbf{w}}\mathbf{A} &= (\mathbf{I}_n - 2\mathbf{w}\mathbf{w}^H)\mathbf{A} = \mathbf{A} - 2\mathbf{w}\mathbf{w}^H\mathbf{A} \\
&= \begin{bmatrix}\mathbf{w} & \mathbf{v}_2 & \dots & \mathbf{v}_n\end{bmatrix} - \begin{bmatrix}2\mathbf{w}\mathbf{w}^H\mathbf{w} & 2\mathbf{w}\mathbf{w}^H\mathbf{v}_2 & \dots & 2\mathbf{w}\mathbf{w}^H\mathbf{v}_n\end{bmatrix} \\
&= \begin{bmatrix}\mathbf{w} - 2\mathbf{w}\underbrace{\mathbf{w}^H\mathbf{w}}_{=1} & \mathbf{v}_2 - 2\mathbf{w}\underbrace{\mathbf{w}^H\mathbf{v}_2}_{=0} & \dots & \mathbf{v}_n - 2\mathbf{w}\underbrace{\mathbf{w}^H\mathbf{v}_n}_{=0}\end{bmatrix} \\
&= [-\mathbf{w}\ \mathbf{v}_2\ \dots\ \mathbf{v}_n]
\end{aligned}$$

dove si è tenuto conto che $\mathbf{w}^H\mathbf{w} = 1$ e $\mathbf{w}^H\mathbf{v}_i = 0$ (i $\mathbf{v}_i$ sono perpendicolari a $\mathbf{w}$). Ne consegue che

$$\det(\mathbf{H}_{\mathbf{w}})\det(\mathbf{A}) = \det(\mathbf{H}_{\mathbf{w}}\mathbf{A}) \overset{(1)}{=} -\det(\mathbf{A})$$

l'uguaglianza 1 è dovuta al fatto che $\mathbf{H}_w\mathbf{A}$ e $\mathbf{A}$ differiscono solo per il segno della prima colonna. Sviluppando $\det \mathbf{H}_w\mathbf{A}$ su questa colonna si può raccogliere un segno $-$ portandolo fuori sommatoria, dopodiché lo sviluppo è lo stesso di $\det \mathbf{A}$, che quindi differiscono solamente per un segno.
Ora

$$\begin{aligned}
\det(\mathbf{H}_{\mathbf{w}})\det(\mathbf{A}) = -\det(\mathbf{A}) &\iff \det(\mathbf{A})(\det(\mathbf{H}_{\mathbf{w}}) + 1) = 0 \\
&\iff \det(\mathbf{A}) = 0 \vee \det(\mathbf{H}_{\mathbf{w}}) = -1
\end{aligned}$$

Ora $\mathbf{A}$ è composta da una base ortonormale, quindi per il teorema 6.1.1 è unitaria ($\mathbf{A}^{-1} = \mathbf{A}^H$) ossia invertibile e dunque $\det \mathbf{A} \neq 0$. Quindi, in conclusione, deve essere $\det(\mathbf{H}_{\mathbf{w}}) = -1$. □

Osservazione 327. Dal fatto che una matrice di Householder $\mathbf{H}_{\mathbf{w}}$ è unitaria e dal teorema 6.1.1, punto 7, si ricava che, dato un vettore $\mathbf{u} \in \mathbb{C}^n$ la norma euclidea di $\mathbf{u}$ coincide con quella di $\mathbf{H}_{\mathbf{w}}\mathbf{u}$.

Osservazione 328. Inoltre si ha che $\mathbf{u}^H\mathbf{H}_{\mathbf{w}}\mathbf{u}$ è un numero reale, perché la matrice $\mathbf{H}_{\mathbf{w}}$ è hermitiana: infatti $\mathbf{u}^H\mathbf{H}_{\mathbf{w}}\mathbf{u}$ è chiaramente un numero, però si ha che

$$\overline{\mathbf{u}^H\mathbf{H}_{\mathbf{w}}\mathbf{u}} = (\mathbf{u}^H\mathbf{H}_{\mathbf{w}}\mathbf{u})^H = \mathbf{u}^H\mathbf{H}_{\mathbf{w}}^H\mathbf{u} = \mathbf{u}^H\mathbf{H}_{\mathbf{w}}\mathbf{u}$$

per cui se un numero coincide con il proprio coniugato è reale.

Osservazione 329. La proprietà seguente da in un certo senso l'inverso di quest'ultimo fatto.

TODO: Da sviluppar qui sino al termine matrici $\mathbf{H_w}$

Proposizione 6.1.10. *Dati due vettori $\mathbf{u} \neq \mathbf{v} \in \mathbb{C}^n$ di uguale norma euclidea e tali che $\mathbf{u}^H\mathbf{v} \in \mathbb{R}$, esiste una matrice di Householder $\mathbf{H_w}$ tale che $\mathbf{H_w u} = \mathbf{v}$.*

Dimostrazione. Si ponga $w = (\mathbf{u} - \mathbf{v}) / \|\mathbf{u} - \mathbf{v}\|_2$. Si lascia al lettore la verifica che $\mathbf{H_w u} = \mathbf{v}$, tenuto conto che $\mathbf{u}^H\mathbf{v} = \mathbf{v}^H\mathbf{u}$ (esercizio 6.4) □

Osservazione 330. Diamo ora una interpretazione geometrica dell'azione svolta dalla pre-moltiplicazione in $\mathbb{C}^n$ per una matrice di Householder $\mathbf{H_w}$. Denoteremo con V lo spazio vettoriale complemento ortogonale del vettore $\mathbf{w}$, cioè $V = \left\{\mathbf{v} \in \mathbb{C}^n : \mathbf{v}^H\mathbf{w} = 0\right\} = \langle \mathbf{w} \rangle^\perp$, e con $\mathbf{P}_V$ la matrice di proiezione di $\mathbb{C}^n$ su V. Osserviamo che se $\mathbf{u} \in \mathbb{C}^n$, la riflessione $\mathbf{u}'$ di $\mathbf{u}$ rispetto al sottospazio V è caratterizzata dalle due proprietà:

$$\mathbf{P}_V\mathbf{u}' = \mathbf{P}_V\mathbf{u}, \quad (\mathbf{I}_n - \mathbf{P}_V)\mathbf{u}' = -(\mathbf{I}_n - \mathbf{P}_V)\mathbf{u}$$

perciò $\mathbf{u}' = -\mathbf{u} + 2\mathbf{P}_V\mathbf{u}$. Ne deduciamo la seguente proprietà

Proposizione 6.1.11. *La pre moltiplicazione in $\mathbb{C}^n$ per la matrice di Householder $\mathbf{H_w}$ opera come la riflessione rispetto al sottospazio $V = \langle \mathbf{w} \rangle^\perp$.*

Dimostrazione. Se $\mathbf{u} \in \mathbb{C}^n$

$$\mathbf{H_w u} = \mathbf{u} - 2\mathbf{w}\mathbf{w}^H\mathbf{u} = \mathbf{u} - 2(\mathbf{I}_n - \mathbf{P}_V)\mathbf{u} = -\mathbf{u} + 2\mathbf{P}_V\mathbf{u} = \mathbf{u}'$$

□

Esempio 6.1.3 (Matrici di Householder generalizzate)**.** Le matrici di Householder possono essere generalizzate considerando, anziché un vettore $\mathbf{w}$ di norma 1 e il sottospazio suo complemento ortogonale $V = \langle \mathbf{w} \rangle^\perp$, un sottospazio W di dimensione > 1 e il suo complemento ortogonale $V = W^\perp$.

Se denotiamo ancora con $\mathbf{P}_W$ e $\mathbf{P}_V = \mathbf{I}_n - \mathbf{P}_W$ le due matrici di proiezione su tali sottospazi, definiamo come matrice di Householder generalizzata associata al sottospazio W di $\mathbb{C}^n$ la matrice

$$\mathbf{H}_W = \mathbb{I}_n - 2\mathbf{P}_W$$

Poiché le matrici di proiezione sono hermitiane e idempotenti, segue facilmente come nella proprietà 1 che $\mathbf{H}_W$ è hermitiana e unitaria. Con ragionamento analogo a quello fatto nella proprietà 2 si prova che $\det(\mathbf{H}_W) = (-1)^d$ dove $d = \dim W$. Infine similmente al punto 4 si vede che la pre-moltiplicazione in $\mathbb{C}^n$ per la matrice $\mathbb{H}_w$ opera sui vettori $\mathbf{u} \in \mathbb{C}^n$ come la riflessione rispetto al sottospazio $V = W^\perp$, cioè, posto $\mathbf{u}' = \mathbf{H}_W\mathbf{u}$:

$$\mathbf{P}_V\mathbf{u} = \mathbf{P}_V\mathbf{u}', \quad \mathbf{P}_W\mathbf{u} = -\mathbf{P}_W\mathbf{u}'$$

6.2 Matrici normali e teorema spettrale

Definizione 6.2.1 (Matrice normale)**.** Ricordiamo che una matrice quadrata $\mathbf{A}$ si dice normale se commuta con la sua H-trasposta, ossia

$$\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A} \tag{6.3}$$

Osservazione 331. Abbiamo già visto in proposizione 1.3.10 che parte hermitiana e anti-hermitiana di una matrice commutano se e solo se la matrice è normale, cioè commuta con la sua H-trasposta.

6.2.1 Esempi di matrici normali

Osservazione 332. Il termine "normale" per una matrice $n \times n$ con $n > 1$ non va inteso nel senso di matrice che si incontra usualmente; la condizione di normalità è una condizione forte, raramente verificata. Alcuni esempi notevoli seguono.

Esempio 6.2.1. Sono normali:

- le matrici diagonali $D = \mathbf{Diag}(d_1, \ldots, d_n)$, in quanto

$$\begin{aligned}\mathbf{DD}^H &= \mathbf{Diag}(d_1, \ldots, d_n)\mathbf{Diag}\big(\overline{d_1}, \ldots, \overline{d_n}\big) \\ &= \mathbf{Diag}\big(d_1\overline{d_1}, \ldots, d_n\overline{d_n}\big) = \mathbf{Diag}\big(\overline{d_1}d_1, \ldots, \overline{d_n}d_n\big) \\ &= \mathbf{Diag}\big(\overline{d_1}, \ldots, \overline{d_n}\big)\mathbf{Diag}(d_1, \ldots, d_n) \\ &= \mathbf{D}^H\mathbf{D}\end{aligned}$$

- le matrici unitarie $\mathbf{U}$ dato che per definizione

$$\mathbf{U}^H\mathbf{U} = \mathbf{I} = \mathbf{UU}^H$$

- le matrici $\mathbf{A}$ hermitiane:

$$\mathbf{A}^H\mathbf{A} = \mathbf{A}^2 = \mathbf{AA}^H$$

- le matrici anti-hermitiane $\mathbf{B}$

$$\mathbf{B}^H\mathbf{B} = -\mathbf{B}^2 = \mathbf{BB}^H$$

Osservazione 333. Matrici unitarie, hermitiane e anti-hermitiane possono essere caratterizzate, all'interno della classe delle matrici normali, tramite proprietà dei loro autovalori, come si vedrà.

6.2.2 Teorema spettrale

Osservazione 334. Uno dei più importanti teoremi su autovalori e autovettori

Definizione 6.2.2 (Matrice unitariamente diagonalizzabile)**.** La matrice quadrata $\mathbf{A}$ è unitariamente diagonalizzabile se può essere fattorizzata come $\mathbf{A} = \mathbf{UDU}^{-1} = \mathbf{UDU}^H$ con $\mathbf{D}$ diagonale e $\mathbf{U}$ unitaria ($\mathbf{U}^{-1} = \mathbf{U}^H$)

Teorema 6.2.1 (Teorema spettrale)**.** *Una matrice complessa* $\mathbf{A}$ *è unitariamente diagonalizzabile se e solo se è normale.*

Dimostrazione. Se la matrice $\mathbf{A} = \mathbf{UDU}^H$ è unitariamente diagonalizzabile con $\mathbf{U}$ unitaria e $\mathbf{D}$ diagonale, risulta

$$\begin{aligned}\mathbf{AA}^H &= \mathbf{UDU}^H\mathbf{UD}^H\mathbf{U}^H \overset{(1)}{=} \mathbf{UDD}^H\mathbf{U}^H \overset{(2)}{=} \mathbf{UD}^H\mathbf{DU}^H \\ &= \mathbf{UD}^H\mathbf{IDU}^H = \mathbf{UD}^H\mathbf{U}^H\mathbf{UDU}^H \\ &= \mathbf{A}^H\mathbf{A}\end{aligned}$$

dove in (1) si è usato $\mathbf{U}^H\mathbf{U} = \mathbf{I}$, seguendo teorema 6.1.1, e in (2) si è usato il fatto che essendo $\mathbf{D}$ diagonale $\mathbf{DD}^H = \mathbf{D}^H\mathbf{D}$.

Viceversa supponendo che $\mathbf{A}$ sia normale proviamo che è unitariamente diagonalizzabile. Dato che $\mathbf{A}$ complessa e i suoi autovalori $\lambda_i \in \mathbb{C}$ è unitariamente simile ad una matrice triangolare superiore per il teorema di Schur (teorema 5.6.4), ossia $\mathbf{A} = \mathbf{UTU}^H$, con $\mathbf{U}$ unitaria e $\mathbf{T}$ triangolare superiore.
Per mostrare che questa è anche una diagonalizzazione unitaria proviamo dapprima che $\mathbf{T}$ è normale, e in secondo luogo che se una matrice è triangolare e normale allora deve esser diagonale:

- per la verifica che $\mathbf{T}$ è normale partiamo derivando una fattorizzazione di $\mathbf{T}$ opportuna. Essendo $\mathbf{U}$ è unitaria ($\mathbf{U}^{-1} = \mathbf{U}^H$) e $\mathbf{A} = \mathbf{UTU}^H$ procedo pre-moltiplicando per $\mathbf{U}^H$ e post-moltiplicando per $\mathbf{U}$

$$\begin{aligned} \mathbf{A} &= \mathbf{UTU}^H \\ \mathbf{U}^H\mathbf{AU} &= \mathbf{U}^H\mathbf{UTU}^H\mathbf{U} \\ \mathbf{U}^H\mathbf{AU} &= \mathbf{T} \end{aligned}$$

  per cui

$$\begin{aligned} \mathbf{T} &= \mathbf{U}^H\mathbf{AU} = \mathbf{U}^{-1}\mathbf{AU} \\ \mathbf{T}^H &= (\mathbf{U}^H\mathbf{AU})^H = \mathbf{U}^H\mathbf{A}^H\mathbf{U} = \mathbf{U}^{-1}\mathbf{A}^H\mathbf{U} \end{aligned}$$

  La verifica allora diviene:

$$\begin{aligned} \mathbf{TT}^H &= \mathbf{U}^{-1}\mathbf{AUU}^{-1}\mathbf{A}^H\mathbf{U} = \mathbf{U}^{-1}\mathbf{AA}^H\mathbf{U} \\ &\overset{(1)}{=} \mathbf{U}^{-1}\mathbf{A}^H\mathbf{AU} = \mathbf{U}^{-1}\mathbf{A}^H\mathbf{IAU} = \mathbf{U}^{-1}\mathbf{A}^H\mathbf{UU}^{-1}\mathbf{AU} \\ &= \mathbf{T}^H\mathbf{T} \end{aligned}$$

  dove in (1) si è usato che, come visto in precedenza $\mathbf{AA}^H = \mathbf{A}^H\mathbf{A}$

- proviamo ora che $\mathbf{T}$ è diagonale per induzione sull'ordine n di $\mathbf{T}$. Se $n = 1$ l'asserto è banale (una matrice 1×1 triangolare è anche diagonale). Sia allora $n > 1$ e l'asserto vero per ipotesi per le matrici triangolari normali di ordine $n - 1$. Scriviamo $\mathbf{T}$ in forma bordata

$$\mathbf{T} = \begin{bmatrix} t & \mathbf{x}^H \\ \mathbf{0} & \mathbf{S} \end{bmatrix}$$

  dove $\mathbf{S}$ è una matrice triangolare di ordine $n-1$. Proviamo che $\mathbf{S}$ è normale, di modo che potremo applicare a $\mathbf{S}$ l'ipotesi induttiva. Si ha infatti:

$$\begin{aligned} \mathbf{TT}^H &= \begin{bmatrix} |t|^2 + \mathbf{x}^H\mathbf{x} & \mathbf{x}^H\mathbf{S}^H \\ \mathbf{Sx} & \mathbf{SS}^H \end{bmatrix} \\ \mathbf{T}^H\mathbf{T} &= \begin{bmatrix} |t|^2 & \bar{t}\mathbf{x}^H \\ t\mathbf{x} & \mathbf{xx}^H + \mathbf{S}^H\mathbf{S} \end{bmatrix} \end{aligned}$$

  Dal fatto che $\mathbf{T}$ è normale si deduce, confrontando i blocchi di posto $(1, 1)$ che $\mathbf{x}^H\mathbf{x} = 0$ e quindi $\mathbf{x} = \mathbf{0}$; confrontando i blocchi di posto $(2, 2)$ si ricava $\mathbf{SS}^H = \mathbf{S}^H\mathbf{S}$, cioè $\mathbf{S}$ è normale. Ma dato che $\mathbf{S}$ è normale ed è triangolare e di ordine $n - 1$, allora per l'ipotesi induttiva è diagonale, quindi, ricordandoci che $\mathbf{x} = \mathbf{0}$, si deduce che anche $\mathbf{T} = \begin{bmatrix} t & \mathbf{0}^T \\ \mathbf{0} & \mathbf{S} \end{bmatrix}$ è diagonale.

$\square$

Esempio 6.2.2. Si consideri la matrice simmetrica reale

$$\mathbf{A} = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}$$

Si ha che

$$\mathbf{A}^H = \mathbf{A}^T = \begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix} = \mathbf{A}$$

pertanto $\mathbf{A}$ è simmetrica. Da questo deriva che $\mathbf{A}^T = \mathbf{A}\mathbf{A}^T = \mathbf{A}^2$ e la matrice $\mathbf{A}$ è normale, e in funzione di ciò unitariamente diagonalizzabile. Procediamo a diagonalizzazione. Vediamo gli autovalori

$$\begin{aligned}
\det(\mathbf{A} - \mathbf{I}X) &= \det \begin{bmatrix} 1-X & -2 & -2 \\ -2 & 1-X & -2 \\ -2 & -2 & 1-X \end{bmatrix} \\
&= (-1)^{1+1}(1-X) \det \begin{bmatrix} 1-X & -2 \\ -2 & 1-X \end{bmatrix} - 2(-1)^{1+2} \det \begin{bmatrix} -2 & -2 \\ -2 & 1-X \end{bmatrix} + \\
&\quad - 2(-1)^{1+3} \det \begin{bmatrix} -2 & 1-X \\ -2 & -2 \end{bmatrix} \\
&= (1-X)\left[(1-X)^2 - 4\right] + 2\left[-2(1-X) - 4\right] - 2\left[4 + 2(1-X)\right] \\
&= (1-X)^3 - 4)1-X) - 4(1-X) - 8 - 8 - 4(1-X) \\
&= (1-X)^3 - 12(1-X) - 16 = -X^3 + 1 + 3X^2 - 3X - 12 + 12X - 16 \\
&= -X^3 + 3X^2 + 9X - 27 = -X^2(X-3) + 9(X-3) \\
&= (X-3)(9-X^2) = (X-3)(3-X)(3+X)
\end{aligned}$$

La matrice ha dunque autovalori reali ($\lambda_1 = \lambda_2 = 3$, $\lambda_3 = -3$). Troviamo gli autovettori di $\mathbf{A}$ determinando basi di $\mathrm{N}(\mathbf{A} + 3\mathbf{I})$ e $\mathrm{N}(\mathbf{A} - 3\mathbf{I})$. Partendo da $\mathbf{A} + 3\mathbf{I}$:

$$\begin{bmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -3 & 3 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & -2 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\to \begin{cases} x_3 = h_3 \\ -x_2 + h_3 = 0 \\ x_1 + x_2 - 2h_3 = 0 \end{cases} \quad \begin{cases} x_3 = h_3 \\ x_2 = h \\ x_1 + h - 2h = 0 \end{cases} \quad \to \begin{bmatrix} h \\ h \\ h \end{bmatrix} = h \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \to \mathrm{N}(\mathbf{A} + 3\mathbf{I}) = \left\langle \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\rangle$$

Mentre per $\mathrm{N}(\mathbf{A} - 3\mathbf{I})$

$$\begin{bmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{cases} x_3 = h_3 \\ x_2 = h_2 \\ x_1 + h_2 + h_3 = 0 \end{cases}$$

Ora se $h_2 = 0, h_3 = 1$ si ha $[-1\ 0\ 1]^T$ se $h_2 = 1, h_3 = 0$ si ha $[-1\ 1\ 0]^T$ quindi

$$\mathrm{N}(\mathbf{A} - 3\mathbf{I}) = \left\langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\rangle$$

Ortogonalizziamo questa seconda base

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ -1/2 \\ 1 \end{bmatrix} \stackrel{(1)}{=} \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \end{bmatrix}$$

dove in (1) si è moltiplicato per -1 al fine di ridurre i segni negativi. Pertanto la matrice ortogonale è

$$\mathbf{Q}' = \begin{bmatrix} -1 & 1/2 & 1 \\ 1 & 1/2 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

Al fine di rendere la matrice unitaria devo normalizzare le colonne già ortogonali

$$\begin{aligned} \|\mathbf{u}_1\| &= \sqrt{2} \\ \|\mathbf{u}_2\| &= \sqrt{\frac{1}{4} + \frac{1}{4} + 1} = \frac{\sqrt{3}}{\sqrt{2}} \\ \|\mathbf{u}_3\| &= \sqrt{3} \end{aligned}$$

$$\mathbf{Q} = \begin{bmatrix} -1/\sqrt{2} & \frac{1}{2}\cdot\frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 1/\sqrt{2} & \frac{1}{2}\cdot\frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & -\frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} = \ldots = \frac{1}{\sqrt{6}}\begin{bmatrix} -\sqrt{3} & 1 & \sqrt{2} \\ \sqrt{3} & 1 & \sqrt{2} \\ 0 & -2 & \sqrt{2} \end{bmatrix}$$

Pertanto una diagonalizzazione di $\mathbf{A}$ con matrice ortonormale $\mathbf{Q} = [\mathbf{u}_1\, \mathbf{u}_2\, \mathbf{u}_3]$ è data da:

$$\begin{aligned} \mathbf{A} &= \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T \\ &= \frac{1}{\sqrt{6}}\begin{bmatrix} -\sqrt{3} & 1 & \sqrt{2} \\ \sqrt{3} & 1 & \sqrt{2} \\ 0 & -2 & \sqrt{2} \end{bmatrix}\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -3 \end{bmatrix}\frac{1}{\sqrt{6}}\begin{bmatrix} -\sqrt{3} & \sqrt{3} & 0 \\ 1 & 1 & -2 \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \end{bmatrix} \end{aligned}$$

Osservazione 335. Ricordiamo che dato un sottospazio vettoriale V di $\mathbb{C}^n$ si denota con $P_V : \mathbb{C}^n \to \mathbb{C}^n$ la proiezione ortogonale su V, realizzabile mediante pre-moltiplicazione per una matrice di proiezione $\mathbf{P}_V$ ottenuta da una qualunque base ortonormale $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ di V come somma di matrici di proiezione di rango 1 nel modo seguente:

$$\mathbf{P}_V = \mathbf{v}_1\mathbf{v}_1^H + \ldots + \mathbf{v}_r\mathbf{v}_r^H$$

(quest'ultima è un altro modo di vedere la $\mathbf{P} = \mathbf{Q}\mathbf{Q}^H$ di teorema 3.6.3, sfruttando il prodotto colonne per righe).

Sappiamo altresì le matrici di proiezione sono hermitiane e idempotenti ($\mathbf{P}^2 = \mathbf{P} = \mathbf{P}^H$) e sono indipendenti dalla base ortonormale scelta per ottenerle.

Osservazione 336. Vedremo ora la versione additiva del teorema spettrale, che dice in sostanza che la moltiplicazione in $\mathbb{C}^n$ per una matrice normale agisce come la somma delle matrici di proiezione sui diversi autospazi ciascuna moltiplicata per il corrispondente autovalore (questo vale per ogni matrice diagonalizzabile, si veda la successiva osservazione 337) e tali matrici di proiezione hanno a due a due prodotto nullo.

Teorema 6.2.2 (Teorema spettrale (versione additiva))**.** *Una matrice complessa $\mathbf{A}$ è unitariamente diagonalizzabile se e solo se è del tipo*

$$\lambda_1 \mathbf{P}_1 + \ldots + \lambda_r \mathbf{P}_r$$

dove gli scalari λ_i sono i suoi distinti autovalori e le matrici $\mathbf{P}_i$ sono le matrici di proiezione sui relativi autospazi e soddisfano le condizioni $\mathbf{P}_i \mathbf{P}_j = \mathbb{O}$ per $i \neq j$

Dimostrazione. Siano $\lambda_1, \ldots, \lambda_r$ gli autovalori distinti di $\mathbf{A}$ con rispettive molteplicità algebriche $m_1, \ldots, m_r$. Se $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$, con $\mathbf{U} = [\mathbf{u}_1 \ldots, \mathbf{u}_n]$ matrice unitaria e $\mathbf{D} = \mathbf{Diag}(d_1, \ldots, d_n)$ matrice diagonale, non è restrittivo supporre che gli autovalori coincidenti compaiano consecutivamente sulla diagonale di $\mathbf{D}$:

$$\lambda_1 = d_1 = \ldots = d_{m_1}, \ldots, \lambda_r = d_{n-m_r+1} = \ldots = d_n$$

Di conseguenza per teorema 5.6.1, l'insieme $\{\mathbf{u}_1, \ldots, \mathbf{u}_{m_1}\}$ è una base ortonormale dell'autospazio $E_{\mathbf{A}}(\lambda_1)$, l'insieme $\{\mathbf{u}_{m_1+1}, \ldots, \mathbf{u}_{m_1+m_2}\}$ è una base ortonormale dell'autospazio $E_{\mathbf{A}}(\lambda_e)$, eccetera.
Un calcolo diretto del prodotto (colonne per righe) $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ mostra che risulta

$$\mathbf{A} = \sum_{i \leq n} d_i \mathbf{u}_i \mathbf{u}_i^H$$

Raggruppando gli scalari d_i coincidenti (sono autovalori che coincidono, es $d_1 = d_2 = \lambda_1$) si ha (avendo posto $m_0 = 0$ e $k_i = m_0 + m_1 + \ldots + m_i$):

$$\mathbf{A} = \sum_{i \leq r} \lambda_i \sum_{k_{i-1} < j \leq k_i} \mathbf{u}_j \mathbf{u}_j^H \tag{6.4}$$

Nella precedente si è switchato agli autovalori considerati una volta sola (indice r), ma moltiplicati si è moltiplicato ciascun autovalore singolo per la sequenza di indici che lo riguardano. Ma come si è ricordato sopra

$$\mathbf{P}_i = \sum_{k_{i-1} < j \leq k_i} \mathbf{u}_j \mathbf{u}_j^H$$

non è altro che la matrice di proiezione sull'autospazio $E_{\mathbf{A}}(\lambda_i)$. Poiché se $i \neq j$ ogni colonna di $\mathbf{P}_i$ è ortogonale a ogni colonna di $\mathbf{P}_j$ ne consegue immediatamente che $\mathbf{P}_i \mathbf{P}_j = \mathbb{O}$ come desiderato.
Viceversa sia $\mathbf{A} = \lambda_1 \mathbf{P}_1 + \ldots + \lambda_r \mathbf{P}_r$ con $\mathbf{P}_i \mathbf{P}_j = \mathbb{O}$ per $i \neq j$. Un calcolo diretto mostra che $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H \mathbf{A}$, in quanto tenuto conto che le P_i sono hermitiane e

idempotenti ($\mathbf{P}_i^2 = \mathbf{P}_i = \mathbf{P}_i^H$):

$$
\begin{aligned}
\mathbf{A}\mathbf{A}^H &= (\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r)(\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r)^H = (\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r)(\overline{\lambda_1}\mathbf{P}_1^H + \ldots + \overline{\lambda_r}\mathbf{P}_r^H) \\
&= (\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r)(\overline{\lambda_1}\mathbf{P}_1 + \ldots + \overline{\lambda_r}\mathbf{P}_r) = (\lambda_1\overline{\lambda_1}\mathbf{P}_1\mathbf{P}_1 + \ldots + \lambda_r\overline{\lambda_r}\mathbf{P}_r\mathbf{P}_r) \\
&= (\lambda_1\overline{\lambda_1}\mathbf{P}_1 + \ldots + \lambda_r\overline{\lambda_r}\mathbf{P}_r) \\
\mathbf{A}^H\mathbf{A} &= (\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r)^H(\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r) = (\overline{\lambda_1}\mathbf{P}_1^H + \ldots + \overline{\lambda_r}\mathbf{P}_r^H)(\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r) \\
&= (\overline{\lambda_1}\mathbf{P}_1 + \ldots + \overline{\lambda_r}\mathbf{P}_r)(\lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r) = (\overline{\lambda_1}\lambda_1\mathbf{P}_1\mathbf{P}_1 + \ldots + \overline{\lambda_r}\lambda_r\mathbf{P}_r\mathbf{P}_r) \\
&= (\overline{\lambda_1}\lambda_1\mathbf{P}_1 + \ldots + \overline{\lambda_r}\lambda_r\mathbf{P}_r)
\end{aligned}
$$

Pertanto $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$, $\mathbf{A}$ è dunque normale, e il fatto che $\mathbf{A}$ sia unitariamente diagonalizzabile segue dal teorema 6.2.1 □

Esempio 6.2.3. Riprendendo il caso di esercizio 6.2.2 la versione additiva del teorema spettrale, nella formula 6.4, porge la decomposizione:

$$\mathbf{A} = 3\mathbf{P}_1 - 3\mathbf{P}_2$$

Calcoliamo $\mathbf{P}_1 = \mathbf{u}_1\mathbf{u}_1^H + \mathbf{u}_2\mathbf{u}_2^H$ (matrice di proiezione sull'autospazio $E_{\mathbf{A}}(3)$) e $\mathbf{P}_2 = \mathbf{u}_3\mathbf{u}_3^H$ (di proiezione sull'autospazio $E_{\mathbf{A}}(-3)$):

$$
\begin{aligned}
\mathbf{P}_1 &= \mathbf{u}_1\mathbf{u}_1^H + \mathbf{u}_2\mathbf{u}_2^H \\
&= \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} + \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -\sqrt{2}/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & -\sqrt{2}/\sqrt{3} \end{bmatrix} \\
&= \begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1/6 & 1/6 & -1/3 \\ 1/6 & 1/6 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix} = \begin{bmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix} \\
&= \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}
\end{aligned}
$$

$$\mathbf{P}_2 = \mathbf{u}_3\mathbf{u}_3^H = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Dunque la decomposizione della versione additiva del teorema spettrale è:

$$\mathbf{A} = 3\mathbf{P}_1 - 3\mathbf{P}_2 = 3 \cdot \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} - 3 \cdot \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Osservazione 337. Il ragionamento fornito nella precedente dimostrazione si applica a ogni matrice diagonalizzabile $\mathbf{A}$, tranne che per il fatto che non è vero che $\mathbf{P}_i\mathbf{P}_j = \mathbb{O}$ per $i \neq j$. Infatti, per una matrice non normale $\mathbf{A}$ non è vero che gli autospazi di autovettori distinti sono ortogonali.
Si possono però ancora scegliere le basi di ciascun autospazio ortonormali, pervenendo a una diagonalizzazione $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$ con $\mathbf{S} = [\mathbf{S}_1 \ldots \mathbf{S}_r]$ avente r blocchi $\mathbf{S}_i$ formati ciascuno da colonne ortonormali (le basi degli r autospazi), soddisfacenti quindi le uguaglianze $\mathbf{S}_i^H\mathbf{S}_i = \mathbf{I}_{m_i}$, dove m_i è la molteplicità algebrica e geometrica dell'autovalore λ_i.

6.2.3 Alcune matrici normali e i loro autovalori

Osservazione 338. Come anticipato in precedenza, caratterizziamo tra le matrici normali le unitarie, le hermitiane e le anti-hermitiane in funzione dei loro autovalori

Teorema 6.2.3. *Sia* $\mathbf{A}$ *una matrice normale:*

1. $\mathbf{A}$ *è unitaria se e solo se i suoi autovalori hanno modulo 1;*
2. $\mathbf{A}$ *è hermitiana se e solo se i suoi autovalori sono numeri reali*
3. $\mathbf{A}$ *è anti-hermitiana se e solo se i suoi autovalori sono numeri immaginari*

Dimostrazione. Rispettivamente

1. la necessità è già stata provata nella proposizione 6.1.3. Viceversa ipotizziamo che gli autovalori della matrice abbiano modulo 1 e mostriamo che è normale; dato che $\mathbf{A}$ è un matrice normale, per il teorema spettrale è unitariamente diagonalizzabile, ossia si ha $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ con $\mathbf{U}$ unitaria e $\mathbf{D}$ diagonale; gli elementi diagonali d_i di $\mathbf{D}$, essendo gli autovalori di $\mathbf{A}$ (dato che le due matrici simili $\mathbf{A}$ e $\mathbf{D}$ hanno stesso polinomio caratteristico e stessi autovalori per proposizione 5.5.2; e gli autovalori di una matrice diagonale $\mathbf{D}$ corrispondono ai suoi elementi diagonali) hanno modulo 1. Risulta allora:
$$\mathbf{A}^H\mathbf{A} = \mathbf{U}\overline{\mathbf{D}}\underbrace{\mathbf{U}^H\mathbf{U}}_{\mathbf{I}}\mathbf{D}\mathbf{U}^H = \mathbf{U}\overline{\mathbf{D}}\mathbf{D}\mathbf{U}^H$$
con $\overline{\mathbf{D}}$ la matrice ottenuta coniugando $\mathbf{D}$. Si ha però che
$$\overline{\mathbf{D}}\mathbf{D} = \mathbf{Diag}\Big(|d_1|^2, \ldots, |d_n|^2\Big) = \mathbf{I}$$
da cui segue che $\mathbf{A}^H\mathbf{A} = \mathbf{I}$. Analogamente vale per
$$\mathbf{A}\mathbf{A}^H = \mathbf{U}\mathbf{D}\underbrace{\mathbf{U}^H\mathbf{U}}_{\mathbf{I}}\overline{\mathbf{D}}\mathbf{U}^H = \mathbf{U}\mathbf{D}\overline{\mathbf{D}}\mathbf{U}^H = \mathbf{I}$$
Per cui $\mathbf{A}$ è unitaria, in accordo a teorema 6.1.1.
2. sia $\mathbf{A}$ hermitiana e sia $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ per un vettore $v \neq \mathbf{0}$. Pre-moltiplicando ambo i membri per $\mathbf{v}^H$ si ricava
$$\mathbf{v}^H\mathbf{A}\mathbf{v} = \lambda \left\|\mathbf{v}\right\|_2^2$$
Tenuto conto che $\mathbf{A} = \mathbf{A}^H$, il numero complesso $z = \mathbf{v}^H\mathbf{A}\mathbf{v}$ coincide con il proprio coniugato,
$$z = \mathbf{v}^H\mathbf{A}\mathbf{v} = (\mathbf{v}^H\mathbf{A}\mathbf{v})^H = \overline{z}$$
quindi è un numero reale; ne consegue che $\lambda = z/\left\|\mathbf{v}\right\|_2^2$ è pure reale.
Viceversa nell'ipotesi che gli autovalori siano reali, mostriamo che $\mathbf{A}$ è hermitiana. $\mathbf{A}$ è normale quindi per il teorema spettrale si diagonalizza in $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ con $\mathbf{U}$ unitaria e $\mathbf{D}$ diagonale; gli elementi diagonali di $\mathbf{D}$, essendo gli autovalori di $\mathbf{A}$ sono numeri reali (per ipotesi). Ne consegue che $\mathbf{D} = \overline{\mathbf{D}}$ e quindi $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H = \mathbf{U}\overline{\mathbf{D}}\mathbf{U}^H = \mathbf{A}^H$, ossia $\mathbf{A}$ è hermitiana.

3. sia $\mathbf{A}$ anti-hermitiana ($\mathbf{A}^H = -\mathbf{A}$, o equivalentemente $\mathbf{A} = -\mathbf{A}^H$) e sia $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ per un vettore $v \neq \mathbf{0}$. Pre-moltiplicando ambo i membri per $\mathbf{v}^H$ si ricava
$$\mathbf{v}^H\mathbf{A}\mathbf{v} = \lambda \|\mathbf{v}\|_2^2$$
Il numero complesso $z = \mathbf{v}^H\mathbf{A}\mathbf{v}$ coincide con l'opposto del coniugato in quanto sviluppando questo secondo
$$-\overline{(\mathbf{v}^H\mathbf{A}\mathbf{v})} = -(\mathbf{v}^H\mathbf{A}\mathbf{v})^H = -\mathbf{v}^H\mathbf{A}^H\mathbf{v} \overset{(1)}{=} \mathbf{v}^H\mathbf{A}\mathbf{v} = z$$
dove in (1) si è sfruttato $\mathbf{A} = -\mathbf{A}^H$. Pertanto z è un numero immaginario, così come $\lambda = z/\|\mathbf{v}\|_2^2$, in quanto al numeratore ha un numero immaginario e al denominatore un reale.
Viceversa nell'ipotesi che gli autovalori siano immaginari, mostriamo che $\mathbf{A}$ è anti-hermitiana. $\mathbf{A}$ è normale quindi per il teorema spettrale si diagonalizza in $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ con $\mathbf{U}$ unitaria e $\mathbf{D}$ diagonale; gli elementi diagonali di $\mathbf{D}$, essendo gli autovalori di $\mathbf{A}$ sono numeri immaginari (per ipotesi). Ne consegue che $-\mathbf{D} = \overline{\mathbf{D}}$, o $\mathbf{D} = -\overline{\mathbf{D}}$, e quindi
$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H = -\mathbf{U}\overline{\mathbf{D}}\mathbf{U}^H = -(\mathbf{U}\mathbf{D}\mathbf{U}^H)^H = -\mathbf{A}^H$$
ossia $\mathbf{A}$ è anti-hermitiana.

□

Proposizione 6.2.4. *Se $\mathbf{A} = [a_{ij}]$ è una matrice qualunque, risulta*
$$\mathrm{Tr}(\mathbf{A}\mathbf{A}^H) = \mathrm{Tr}(\mathbf{A}^H\mathbf{A}) = \sum_{i,j} |a_{ij}|^2 \tag{6.5}$$

Dimostrazione. La prima uguaglianza è garantita dal fatto che $\mathrm{Tr}\,\mathbf{AB} = \mathrm{Tr}\,\mathbf{BA}$; sviluppiamo una delle due tracce per verificare la seconda uguaglianza. Ad esempio il posto i, i di $\mathbf{A}^H\mathbf{A}$ è
$$\mathbf{a}_i^H\mathbf{a}_i = \|\mathbf{a}_i\|^2 = |a_{i1}|^2 + \ldots + |a_{in}|^2 = \sum_j |a_{ij}|^2$$
Sommando per tutti gli elementi i, i sulla diagonale, si ha
$$\sum_i \sum_j |a_{ij}|^2$$

□

Osservazione 339. Le matrici normali hanno una particolarità/caratterizzazione su $\mathrm{Tr}\,\mathbf{A}\mathbf{A}^H$, pari ad una opportuna quantità derivata dagli autovalori.

Proposizione 6.2.5. *Una matrice complessa $\mathbf{A} = [a_{ij}]$ con autovalori $d_1, \ldots, d_n$ (ripetuti con le loro molteplicità algebriche), è normale se e solo se*
$$\mathrm{Tr}(\mathbf{A}\mathbf{A}^H) = \sum_i |d_i|^2$$

Dimostrazione. Se $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ è normale con $\mathbf{U}$ unitaria e $\mathbf{D}$ diagonale (le notazioni sono quelle della dimostrazione del teorema 6.2.2, allora per ogni coppia di matrici $\mathbf{X}$ e $\mathbf{Y}$ moltiplicabili nei due sensi si ha

$$\begin{aligned}\operatorname{Tr}(\mathbf{A}\mathbf{A}^H) &= \operatorname{Tr}(\mathbf{U}\mathbf{D}\mathbf{U}^H\mathbf{U}\mathbf{D}^H\mathbf{U}^H) = \operatorname{Tr}(\mathbf{U}\mathbf{D}\mathbf{D}^H\mathbf{U}^H)\\ &= \operatorname{Tr}(\mathbf{U}^H\mathbf{U}\mathbf{D}\mathbf{D}^H) = \operatorname{Tr}(\mathbf{D}\mathbf{D}^H) = \operatorname{Tr}(\mathbf{D}\overline{\mathbf{D}})\\ &= \sum_i |d_i|^2\end{aligned}$$

dove in (1) abbiamo sfruttato $\operatorname{Tr}(\mathbf{AB}) = \operatorname{Tr}(\mathbf{BA})$ con $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{D}^H$ e $B = \mathbf{U}^H$. Viceversa sia $\operatorname{Tr}(\mathbf{A}\mathbf{A}^H) = \sum_i |d_i|^2$. Da una fattorizzazione $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^H$ ottenuta dal teorema di Schur, con $\mathbf{T} = (t_ij)$ matrice triangolare e $\mathbf{U}$ matrice unitaria, si ricava che gli elementi t_{ii} (sulla diagonale) coincidono (a meno dell'ordine) con gli autovalori d_i e che

$$\mathbf{A}\mathbf{A}^H = \mathbf{U}\mathbf{T}\mathbf{U}^H(\mathbf{U}\mathbf{T}\mathbf{U}^H)^H = \mathbf{U}\mathbf{T}\mathbf{U}^H\mathbf{U}\mathbf{T}^H\mathbf{U}^H = \mathbf{U}\mathbf{T}\mathbf{T}^H\mathbf{U}^H$$

Di conseguenza

$$\sum_i |d_i|^2 = \operatorname{Tr}(\mathbf{A}\mathbf{A}^H) = \operatorname{Tr}(\mathbf{U}\mathbf{T}\mathbf{T}^H\mathbf{U}^H) \overset{(1)}{=} \operatorname{Tr}(\mathbf{U}^H\mathbf{U}\mathbf{T}\mathbf{T}^H) = \operatorname{Tr}(\mathbf{T}\mathbf{T}^H) \overset{(2)}{=} \sum_{i,j} |t_{ij}|^2$$

dove in (1) abbiamo sfruttato $\operatorname{Tr}(\mathbf{AB}) = \operatorname{Tr}(\mathbf{BA})$ con $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{T}^H$ e in (2) equazione 6.5. Pertanto, in virtù del fatto che $\sum_i |d_i|^2 = \sum_i |t_{ii}|^2$, si ricava che $t_{ij} = 0$ per $i \neq j$, ovvero, $\mathbf{T}$ è diagonale e $\mathbf{A}$ è unitariamente diagonalizzabile (dato che $\mathbf{U}$ è unitaria e $\mathbf{T}$ è diagonale. La conclusione che $\mathbf{A}$ è normale segue dal teorema spettrale. □

6.3 Matrici simmetriche e hermitiane

6.3.1 Forme quadratiche e matrici variamente simmetriche

Osservazione 340. Matrici reali simmetriche e complesse hermitiane sono tra le più studiate, anche perché intimamente legate alle forme quadratiche.

Definizione 6.3.1 (Forma quadratica complessa n-dimensionale)**.** Data una matrice $\mathbf{A} = [a_{ij}] \in M_n(\mathbb{C})$, detta *generatrice*, si può associare la forma quadratica complessa n-dimensionale

$$Q : \mathbb{C}^n \to \mathbb{C} \qquad Q(\mathbf{x}) = \mathbf{x}^H\mathbf{A}\mathbf{x} = \sum_{i,j} a_{ij}\overline{x_i}x_j$$

Osservazione 341. Le matrici generatrici hermitiane intervengono quando si considerano forme quadratiche complesse che assumono solo valori reali, come prova il seguente risultato

Proposizione 6.3.1. *La forma quadratica complessa $Q(\mathbf{x}) = \mathbf{x}^H\mathbf{A}\mathbf{x}$ assume solo valori reali se e solo se la matrice $\mathbf{A}$ è hermitiana.*

Dimostrazione. Partiamo ipotizzando che la forma quadratica dia un risultato reale e mostriamo che $\mathbf{A}$ è hermitiana; dato un arbitrario vettore $\mathbf{x} \in \mathbb{C}^n$, si ha

$$\mathbf{x}^H\mathbf{A}\mathbf{x} \in \mathbb{R} \overset{(1)}{\iff} \mathbf{x}^H\mathbf{A}\mathbf{x} = (\mathbf{x}^H\mathbf{A}\mathbf{x})^H = \mathbf{x}^H\mathbf{A}^H\mathbf{x} \overset{(2)}{\iff} \mathbf{A} = \mathbf{A}^H$$

dove in (1) la forma quadratica è un numero reale se e solo se coincide con il suo coniugato e in (2) si confrontano primo e ultimo membro dell'uguaglianza. Viceversa se $\mathbf{A}$ è hermitiana, applicando $\mathbf{x} = \mathbf{e}_i$ nella forma quadratica si ha

$$a_{ii} = \mathbf{e}_i^T\mathbf{A}\mathbf{e}_i \overset{(1)}{=} \mathbf{e}_i^T\mathbf{A}^H\mathbf{e}_i = \overline{a_{ii}}$$

(con (1) dovuto al fatto che $\mathbf{A}$ è hermitiana); quindi gli elementi diagonali di $\mathbf{A}$ sono reali. Prendendo poi $i \neq j$:

- se applichiamo $\mathbf{x} = \mathbf{e}_i + \mathbf{e}_j$ nella forma quadratica $\mathbf{x}^H A\mathbf{x}$ facciamo la somma della i-esima e della j-esima riga di $\mathbf{A}$, e di questa somma sommiamo l'i-esimo e il j-esimo elemento. In questo caso, il fatto che $\mathbf{x}^H\mathbf{A}\mathbf{x} \in \mathbb{R}$ equivale a dire che $a_{ii} + a_{jj} + a_{ij} + a_{ji} \in \mathbb{R}$; ma dato che $a_{ii}, a_{jj} \in \mathbb{R}$ per quanto appena visto, ci basta mostrare che $a_{ij} + a_{ji} \in \mathbb{R}$. Ricordando che $\mathbf{A}$ è hermitiana (quindi $a_{ji} = \overline{a_{ij}}$), se a_{ij} in forma algebrica è $a_{ij} = a + ib$, si ha
$$a_{ij} + a_{ji} = a_{ij} + \overline{a_{ij}} = a + ib + a - bi = 2a \in \mathbb{R}$$
- similmente applicando $\mathbf{y} = \mathbf{e}_i + i\mathbf{e}_j$ nella forma quadratica (non lasciandosi confondere dal fatto che "i" è usata sia come indice che come numero complesso), $\mathbf{y}^H\mathbf{A}\mathbf{y} \in \mathbb{R}$ equivale ad $a_{ii} + a_{jj} + ia_{ij} - ia_{ji} \in \mathbb{R}$ e più precisamente ci basta dimostrare $i(a_{ij} - a_{ji}) \in \mathbb{R}$. Similmente a quanto fatto in precedenza:
$$\begin{aligned} i(a_{ij} - a_{ji}) &= i(a_{ij} - \overline{a_{ij}}) = i(a + bi - a + bi) = i(2bi) = i^2 2b \\ &= -2b \in \mathbb{R} \end{aligned}$$

□

Osservazione 342. Se $\mathbf{A}$ è reale si definisce una forma quadratica n-dimensionale.

Definizione 6.3.2 (Forma quadratica n-dimensionale)**.** Data una matrice $\mathbf{A} = [a_{ij}] \in M_n(\mathbb{R})$, detta *generatrice*, si può associare la forma quadratica n-dimensionale

$$Q : \mathbb{R}^n \to \mathbb{R} \qquad Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x} = \sum_{i,j} a_{ij}x_ix_j$$

Osservazione 343. La matrice generatrice di una tale forma può essere una qualunque matrice reale di ordine n; però, anche in questo caso, le matrici simmetriche giocano un ruolo privilegiato.

Proposizione 6.3.2. *Data la forma quadratica reale $Q : \mathbb{R}^n \to \mathbb{R}$ generata dalla matrice $\mathbf{A}_1$ esiste una e una sola matrice simmetrica $\mathbf{A} \in M_n(\mathbb{R})$ tale che $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}_1\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{x}$, e tale è la parte simmetrica di $\mathbf{A}_1$, ossia $\mathbf{A} = (\mathbf{A}_1 + \mathbf{A}_1^T)/2$.*

Dimostrazione. Se $\mathbf{A} = (\mathbf{A}_1 + \mathbf{A}_1^T)/2$ è la parte simmetrica di $\mathbf{A}_1$, si ha:

$$\begin{aligned}\mathbf{x}^T\mathbf{A}\mathbf{x} &= \frac{\mathbf{x}^T\mathbf{A}_1 x + \mathbf{x}^T\mathbf{A}_1^T\mathbf{x}}{2} \overset{(1)}{=} \frac{\mathbf{x}^T\mathbf{A}_1 x + (\mathbf{x}^T\mathbf{A}_1^T\mathbf{x})^T}{2} = \frac{\mathbf{x}^T\mathbf{A}_1 x + \mathbf{x}^T\mathbf{A}_1\mathbf{x}}{2} \\ &= \mathbf{x}^T\mathbf{A}_1\mathbf{x} = Q(\mathbf{x})\end{aligned}$$

dove in (1) si è sfruttato il fatto che un singolo numero $\mathbf{x}^T\mathbf{A}_1^T\mathbf{x}$ coincide con la sua $(\mathbf{x}^T\mathbf{A}_1^T\mathbf{x})^T$, e in seguito si sono applicate le proprietà della trasposizione. Pertanto, essendo $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}_1\mathbf{x} = \mathbf{x}^T\mathbf{A}\mathbf{x}$, anche $\mathbf{A}$ è generatrice di Q. (la si può usare nella formula, tanto il risultato non cambia)
Quanto all'unicità, la forma quadratica:

- applicata ad $\mathbf{e}_i$
$$Q(\mathbf{e}_i) = \mathbf{e}_i^T\mathbf{A}\mathbf{e}_i = a_{ii}$$
individua univocamente l'elemento diagonale a_{ii} per ogni indice i
- applicata a $\mathbf{e}_i + \mathbf{e}_j$, individua
$$Q(\mathbf{e}_i + \mathbf{e}_j) = a_{ii} + a_{jj} + a_{ij} + a_{ji}$$
ma più utilmente usandola in congiunzione con la precedente individua la somma $a_{ij} + a_{ji}$
$$Q(\mathbf{e}_i + \mathbf{e}_j) - Q(\mathbf{e}_i) - Q(\mathbf{e}_j) = a_{ij} + a_{ji}$$
- essendo $\mathbf{A}$ simmetrica a_{ij} e a_{ji} restano pure univocamente individuati; per la simmetria si ha $a_{ij} = a_{ji}$, dunque
$$a_{ij} = a_{ji} = \frac{Q(\mathbf{e}_i + \mathbf{e}_j) - Q(\mathbf{e}_i) - Q(\mathbf{e}_j)}{2}$$

□

Esempio 6.3.1. La forma quadratica reale 2-dimensionale con $\mathbf{y} = [x_1\ x_2]^T$

$$Q(\mathbf{y}) = Q\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = ax_1^2 + bx_1x_2 + cx_2x_1 + dx_2^2 \quad (a, b, c, d \in \mathbb{R})$$

è generata dalla matrice

$$\mathbf{A}_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Infatti si ha

$$\begin{aligned}\mathbf{y}^T\mathbf{A}_1\mathbf{y} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1 + cx_2 & bx_1 + dx_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= (ax_1 + cx_2)x_1 + (bx_1 + dx_2)x_2 \\ &= ax_1^2 + cx_1x_2 + bx_1x_2 + dx_2^2\end{aligned}$$

La stessa forma quadratica è altresì generata dalla parte simmetrica

$$\mathbf{A} = \frac{\mathbf{A}_1 + \mathbf{A}_1^T}{2} = \frac{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a & c \\ b & d \end{bmatrix}}{2} = \begin{bmatrix} a & (b+c)/2 \\ (b+c)/2 & d \end{bmatrix}$$

Infatti

$$\begin{aligned}
\mathbf{y}^T\mathbf{A}\mathbf{y} &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & (b+c)/2 \\ (b+c)/2 & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= \begin{bmatrix} ax_1 + x_2\frac{b+c}{2} & x_1\frac{b+c}{2} + dx_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\
&= \left(ax_1 + x_2\frac{b+c}{2}\right)x_1 + \left(x_1\frac{b+c}{2} + dx_2\right)x_2 \\
&= ax_1^2 + x_1x_2\frac{b+c}{2} + x_1x_2\frac{b+c}{2} + dx_2^2 \\
&= ax_1^2 + x_1x_2(b+c) + dx_2^2 \\
&= ax_1^2 + bx_1x_2 + cx_1x_2 + dx_2^2
\end{aligned}$$

Osservazione 344. L'importanza delle matrici reali simmetriche deriva oltre che da quanto asserito nella proposizione 6.3.2, soprattutto dal fatto che molti fenomeni modellabili tramite matrici e simmetrie, danno di fatto luogo a matrici simmetriche. Quindi tali matrici si incontrano spesso nelle applicazioni.

6.3.2 Altre proprietà di matrici variamente simmetriche

Osservazione 345. Di seguito alcune proprietà delle matrici hermitiane e anti-hermitiane, simmetriche e anti-simmetriche (ricordando che per questi ultimi due tipi di matrici ci limitiamo a matrici reali).

Proposizione 6.3.3. *Gli elementi diagonali di una matrice hermitiana* $\mathbf{A} = [a_{ij}]$*, la sua traccia e il suo determinante sono numeri reali.*

Dimostrazione. Infatti a_{ii} deve coincidere con $\overline{a_{ii}}$ per ogni i quindi $a_{ii} \in \mathbb{R}$. La traccia, essendo la somma degli elementi diagonali, è essa pure reale. È stato poi provato nel teorema 6.2.3 che gli autovalori di una matrice hermitiana sono reali. Il determinante, essendo il prodotto degli autovalori (proposizione 5.4.9), è esso pure reale. Faremo uso di questo fatto nei teoremi 6.3.8 e 6.3.9. $\square$

Proposizione 6.3.4. *Gli elementi diagonali di una matrice anti-hermitiana* $\mathbf{A} = [a_{ij}]$ *sono numeri immaginari e gli elementi diagonali di una matrice anti-simmetrica reale* $\mathbf{A} = [a_{ij}]$ *sono nulli*

Dimostrazione. Infatti nel primo caso a_{ii} deve coincidere con $-\overline{a_{ii}}$ per ogni i, quindi a_{ii} ha parte reale nulla. Nel secondo caso a_{ii} deve coincidere con $-a_{ii}$, quindi $a_{ii} = 0$ $\square$

Proposizione 6.3.5. *Il determinante di una matrice anti-simmetrica reale* $n \times n$ $\mathbf{A}$ *è nullo per* n *dispari, ed è un numero reale non negativo per* n *pari*

Dimostrazione. Poiché il polinomio caratteristico di $\mathbf{A}$ ha coefficienti reali (essendo $\mathbf{A}$ una matrice reale), se λ è una sua radice lo è anche $\overline{\lambda}$ (se radice è reale la coniugata coincide ed è sempre reale; se viceversa la radice è complessa vi è un teorema di algebra - *complex conjugate root theorem* - che afferma che anche la coniugata di una radice di un polinomio è radice dello stesso).
Tuttavia λ ha parte reale nulla, per il teorema 6.2.3 (poiché una matrice reale antisimmetrica è anche anti-hermitiana), quindi gli autovalori di $\mathbf{A}$ se non sono

nulli (ma hanno parte immaginaria $\neq 0$), vanno a coppie: $ir, -ir$ con $r \in \mathbb{R}$. Poiché il determinante è il prodotto degli autovalori e $-ir \cdot ir = r^2$ segue l'asserto: se ve ne è uno solo a 0 di autovalori, $\det \mathbf{A} = 0$, altrimenti andando a coppie di coniugati $det\mathbf{A}$ è la somma di reali $r^2 + s^2 + \ldots \in \mathbb{R}$ □

6.3.3 Autovalori di matrici hermitiane

Osservazione 346. Ci concentreremo d'ora in avanti sulle matrici hermitiane, tenendo presente che tutti i risultati che proveremo sono validi anche per le matrici simmetriche reali, con le ovvie modifiche.

Osservazione 347 (Notazione). Poiché una matrice hermitiana $\mathbf{A}$ di ordine n ha tutti i suoi n autovalori reali (teorema 6.2.3) possiamo disporre tali autovalori, ripetuti con la loro molteplicità, in ordine non decrescente

$$\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{n-1} \leq \lambda_n$$

Questa notazione resterà fissata e sottintesa per tutto il resto del capitolo.

Osservazione 348. Nostro prossimo scopo è caratterizzare gli autovalori λ_i come massimi o minimi di una certa funzione a valori reali definita su $\mathbb{C}^n$, assunti su opportuni sottospazi di $\mathbb{C}^n$.

Definizione 6.3.3 (Funzione di Rayleigh-Ritz)**.** Data la matrice hermitiana $\mathbf{A}$ di ordine n, si chiama funzione di Rayleigh-Ritz associata ad $\mathbf{A}$ la funzione

$$\rho_{\mathbf{A}} : \mathbb{C}^n \setminus \{0\} \to \mathbb{R}, \quad \rho_{\mathbf{A}}(\mathbf{x}) = \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\mathbf{x}^H \mathbf{x}} = \frac{\mathbf{x}^H \mathbf{A} \mathbf{x}}{\|\mathbf{x}\|_2^2}, \quad \forall \mathbf{0} \neq \mathbf{x} \in \mathbb{C}^n$$

Osservazione 349. La funzione ha come codominio $\mathbb{R}$ proprio grazie al fatto che $\mathbf{A}$ hermitiana, e dunque il numeratore $\mathbf{x}^H \mathbf{A} \mathbf{x} \in \mathbb{R}$ (proposizione 6.3.1).

Definizione 6.3.4 (Funzione di Rayleigh-Ritz (definizione con versore))**.** Denotando con $S_n = \{\mathbf{x} \in \mathbb{C}^n : \|\mathbf{x}\|_2 = 1\}$ la superficie sferica n-dimensionale costituita dai vettori di norma euclidea 1, la funzione si semplifica a:

$$\rho_{\mathbf{A}} : S_n \to \mathbb{R}, \quad \rho_{\mathbf{A}}(\mathbf{y}) = \mathbf{y}^H \mathbf{A} \mathbf{y}, \forall \mathbf{y} \in S_n$$

Osservazione 350. Noi useremo indifferentemente le due differenti definizioni.

Osservazione 351 (Notazione). Dato che $\mathbf{A}$ è hermitiana allora è normale (esercizio 6.2.1) e per il teorema spettrale (teorema 6.2.1) è unitariamente diagonalizzabile. Se $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ è una sua diagonalizzazione unitaria possiamo supporre che $\mathbf{\Lambda} = \mathbf{Diag}(\lambda_1, \ldots, \lambda_n)$ (ossia di avere gli autovalori in ordine crescente) e $\mathbf{U} = [\mathbf{u}_1 \ldots \mathbf{u}_n]$ con $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ una base ortonormale di $\mathbb{C}^n$. Per ogni k compreso tra 1 e n poniamo :

$$V_k = \langle \mathbf{u}_1; \ldots; \mathbf{u}_k \rangle, \quad V^k = \langle \mathbf{u}_k; \ldots; \mathbf{u}_n \rangle,$$

NB: Qui la notazione è differente dal libro, in pratica si invertono V_k e V^k

Si noti che $V^1 = \mathbb{C}^n = V_n$. Con questa notazione si ha il seguente risultato.

Teorema 6.3.6 (Principio di Rayleigh-Ritz)**.** *Data la matrice hermitiana* $\mathbf{A}$ *di ordine* n *si ha*

1. *la funzione, applicata a qualsiasi vettore di* $\mathbb{C}^n \setminus \{\mathbf{0}\}$*, sta tra l'autovalore minimo e il massimo:*

$$\lambda_1 \leq \rho_{\mathbf{A}}(\mathbf{x}) \leq \lambda_n, \forall \mathbf{0} \neq \mathbf{x} \in \mathbb{C}^n$$

2. *l'autovalore minimo è il minimo della funzione*

$$\lambda_1 = \min_{\mathbf{0} \neq \mathbf{x}} \rho_{\mathbf{A}}(\mathbf{x})$$

3. *l'autovalore massimo è il massimo della funzione*

$$\lambda_n = \max_{\mathbf{0} \neq \mathbf{x}} \rho_{\mathbf{A}}(\mathbf{x})$$

4. *se* $2 \leq k \leq n-1$ *si ha:*

$$\lambda_k = \max_{\mathbf{0} \neq \mathbf{x} \in V_k} \rho_{\mathbf{A}}(\mathbf{x}) = \min_{\mathbf{0} \neq \mathbf{x} \in V^k} \rho_{\mathbf{A}}(\mathbf{x})$$

Dimostrazione. Rispettivamente:

1. se $\mathbf{x} \in \mathbb{C}^n$ e $\|\mathbf{x}\|_2 = 1$, posto $\mathbf{U}^H\mathbf{x} = \mathbf{z} = [z_1 \, \dots \, z_n]^T$ si ha (adottando la versione della funzione per versori):

$$\rho_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^H\mathbf{A}\mathbf{x} = \underbrace{\mathbf{x}^H\mathbf{U}}_{(\mathbf{U}^H\mathbf{x})^H} \mathbf{\Lambda}\mathbf{U}^H\mathbf{x} = \mathbf{z}^H\mathbf{\Lambda}\mathbf{z} = \sum_{i=1}^{n} \lambda_i |z_i|^2$$

dove per il teorema 6.1.1 punto 7, applicato ad $\mathbf{U}^H$ si ha che

$$1 = \|\mathbf{x}\|_2 = \left\|\mathbf{U}^H\mathbf{x}\right\|_2 = \|\mathbf{z}\|_2$$

Dal fatto che $\lambda_1 \leq \lambda_i \leq \lambda_n$, e ricordando che $\|\mathbf{z}\|_2 = \sum_i |z_i|^2 = 1$, per ogni i si ha che

$$\lambda_1 = \lambda_1 \underbrace{\sum_{i=1}^{n} |z_i|^2}_{=1} \overset{(1)}{\leq} \underbrace{\sum_{i=1}^{n} \lambda_i |z_i|^2}_{\rho_{\mathbf{A}}(\mathbf{x})} \overset{(2)}{\leq} \lambda_n \underbrace{\sum_{i=1}^{n} |z_i|^2}_{=1} = \lambda_n$$

dove la disuguaglianza (1) è garantita dal fatto che portando dentro sommatoria λ_1 si ha un risultato minore rispetto ad utilizzare i vari λ_i (dato che $\lambda_1 \leq \lambda_i$); specularmente avviene per $\lambda_i \leq \lambda_n$ in (2).

2. l'asserto segue banalmente dal primo punto e per calcolarlo si ha

$$\lambda_1 = \lambda_1 \underbrace{\|\mathbf{u}_1\|_2^2}_{=1(1)} = \lambda_1\mathbf{u}_1^H\mathbf{u}_1 = \mathbf{u}_1^H \underbrace{\lambda_1\mathbf{u}_1}_{=\mathbf{A}\mathbf{u}_1} = \mathbf{u}_1^H\mathbf{A}\mathbf{u}_1 \overset{(2)}{=} \rho_{\mathbf{A}}(\mathbf{u}_1)$$

dove in (1) dato che sono colonne ortonormali e in (2) si è usata la definizione per i versori.
Quindi direi che per calcolare l'autovalore minimo applico la funzione ai vettori della base ortonormale $\mathbf{u}_i$ e tengo il risultato minore (non è detto che siano ordinate, presumo);

3. l'asserto segue dal primo punto per il calcolo si ha, specularmente:

$$\lambda_n = \lambda_n \left\|\mathbf{u}_n\right\|_2^2 = \lambda_n \mathbf{u}_n^H \mathbf{u}_n = \mathbf{u}_n^H \mathbf{A}\mathbf{u}_n = \rho_{\mathbf{A}}(\mathbf{u}_n)$$

Quindi tengo il risultato maggiore della funzione (definita sui versori) applicati alle colonne della base ortonormale;

4. dimostriamo la prima uguaglianza e lasciamo al lettore come esercizio l'analoga verifica della seconda. Se $\|x\|_2 = 1$ e $\mathbf{x} \in V^k = \langle \mathbf{u}_k; \dots; \mathbf{u}_n \rangle = \langle \mathbf{u}_1; \dots; \mathbf{u}_{k-1} \rangle^{\perp}$, allora il vettore $\mathbf{U}^H\mathbf{x} = \mathbf{z}$ ha le prime $k-1$ coordinate nulle. Pertanto, considerando che $\|\mathbf{z}\| = \sum_i |z_i|^2 = 1$, si ha: **TODO**: fixme

$$\lambda_k = \lambda_k \underbrace{\sum_{i=1}^{n} |z_i|^2}_{=1} \overset{(1)}{=} \lambda_k \sum_{i=k}^{n} |z_i|^2 = \sum_{i=k}^{n} \lambda_k |z_i|^2 \overset{(2)}{\leq} \sum_{i=k}^{n} \lambda_i |z_i|^2 \overset{(3)}{=} \sum_{i=1}^{n} \lambda_i |z_i|^2 = \rho_{\mathbf{A}}(\mathbf{x})$$

in (1) ci si è concentrati da k a n dato che le prime $k-1$ sono nulle, in (2) la disuguaglianza vale perché $\lambda_k \leq \lambda_i$ qui (dato che i va da k a n), in (3) ritiriamo dentro anche le prime $k-1$ coordinate.
Si ha infine

$$\lambda_k = \lambda_k \underbrace{\left\|\mathbf{u}_k\right\|_2^2}_{=1} = \lambda_k \mathbf{u}_k^H \mathbf{u}_k = \mathbf{u}_k^H \underbrace{\lambda_k \mathbf{u}_k}_{=\mathbf{A}\mathbf{u}_k} = \mathbf{u}_k^H \mathbf{A}\mathbf{u}_k = \rho_{\mathbf{A}}(\mathbf{u}_k)$$

quindi $\lambda_k = \min_{\mathbf{0} \neq \mathbf{x} \in V^k} \rho_{\mathbf{A}}(\mathbf{x})$ come desiderato; per calcolarlo applico $\rho_{\mathbf{A}}$ a tutti gli elementi di V^k e tengo il minore.

□

6.3.4 Matrici definite e semi-definite positive

Definizione 6.3.5 (Matrice definita positiva)**.** Matrice complessa hermitiane (o simmetrica reale) i cui autovalori reali sono positivi

Definizione 6.3.6 (Matrice semi-definita positiva)**.** Matrice complessa hermitiane (o simmetrica reale) i cui autovalori reali sono non-negativi

Osservazione 352. I due teoremi che seguono caratterizzano queste matrici in vari modi, senza ricorrere agli autovalori, che in generale non si possono conoscere in modo esatto, ma solo approssimato. Prima premettiamo un risultato che ci servirà nel secondo caso.

Lemma 6.3.7. *Se $\mathbf{A} = [a_{ij}]$ è una matrice semi-definita positiva e $a_{ii} = 0$, allora la i-esima riga e la i-esima colonna sono nulle.*

Dimostrazione. Da fare: usare il fatto che le sottomatrici 2×2 hanno determinante 0 □

Teorema 6.3.8 (Caratterizzazione matrici definite positive)**.** *Data la matrice hermitiana $\mathbf{A}$ di ordine n i fatti seguenti sono equivalenti:*

1. *gli autovalori di $\mathbf{A}$ sono positivi;*

2. *la funzione di Rayleigh assume valori solo positivi:* $\rho_{\mathbf{A}}(\mathbf{x}) > 0$ *per ogni vettore* $\mathbf{x} \neq \mathbf{0}$*;*

3. *ogni sottomatrice principale* k*-esima di* $\mathbf{A}$ *ha determinante positivo;*

4. *l'eliminazione di Gauss su* $\mathbf{A}$ *produce senza permutazioni una decomposizione del tipo* $\mathbf{A} = \mathbf{L}_0\mathbf{D}\mathbf{L}_0^H$*, con* $\mathbf{L}_0$ *matrice uni-triangolare inferiore e* $\mathbf{D}$ *matrice diagonale con elementi diagonali reali positivi;*

5. *esiste una matrice invertibile* $\mathbf{W}$ *tale che* $\mathbf{A} = \mathbf{W}\mathbf{W}^H$*;*

6. *esiste una matrice definita positiva* $\mathbf{B}$ *tale che* $\mathbf{A} = \mathbf{B}^2$*;*

7. *se* $\mathbf{A}$ *è in forma bordata*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{x} \\ \mathbf{x}^H & a \end{bmatrix}$$

allora $\mathbf{A}_{n-1}$ *è definita positiva e* $a > \mathbf{x}^H\mathbf{A}_{n-1}^{-1}\mathbf{x}$.

Dimostrazione. Si ha:

- $1 \iff 2$: immediata conseguenza del teorema 6.3.6: si ha che $0 < \lambda_1 \leq \rho_{\mathbf{A}}(\mathbf{x})$ quindi $\rho_{\mathbf{A}}(\mathbf{x}) > 0$;

- $1 \implies 3$: osserviamo innanzitutto che $\det \mathbf{A} > 0$ perché il determinante è il prodotto degli autovalori (proposizione 5.4.9), tutti positivi. È sufficiente allora provare che, se $\mathbf{A}_k$ denota la sottomatrice principale k-esima (ottenuta scegliendo le prime k righe e colonne di $\mathbf{A}$), anche $\mathbf{A}_k$ è definita positiva, ossia ha autovalori positivi e dunque il determinante loro prodotto è positivo. Ovviamente $\mathbf{A}_k$ è hermitiana. Attesa l'equivalenza già provata dei punti 1 e 2, è sufficiente dimostrare che la funzione di Rayleigh-Ritz con $\mathbf{A}_k$ assume solo valori positivi, ossia $\mathbf{y}^H\mathbf{A}_k\mathbf{y} > 0$ per ogni vettore non nullo $\mathbf{y} \in \mathbb{C}^k$.
  Posto $\mathbf{z}^H = [y^H\ 0^H] \in \mathbb{C}^n$ (ossia $\mathbf{z}$ è come $\mathbf{y}$ al quale abbiamo aggiunto $n-k$ termini nulli), si ha (con y_i per le coordinate di $\mathbf{y}$ e i z_i quelle di $\mathbf{z}$):

  $$\mathbf{y}^H\mathbf{A}_k\mathbf{y} = \sum_{i,j \leq k} a_{ij}\overline{y_i}y_j \overset{(1)}{=} \sum_{i,j \leq k} a_{ij}\overline{z_i}z_j \overset{(2)}{=} \mathbf{z}^H\mathbf{A}\mathbf{z} = \rho_{\mathbf{A}}(\mathbf{z}) > 0$$

  dove in (1) l'uguaglianza vale perché da $k+1$ ad n aggiungiamo alla somma solo dei termini nulli (provenienti da $\mathbf{0}^H$), in (2) dato che espandiamo da k ad n ritorna fuori $\mathbf{A}$ invece di $\mathbf{A}_k$.

- $3 \implies 4$: proviamo per induzione su k che il k-esimo pivot d_k incontrato nel corso della EG è reale positivo (quindi poi tutti i pivot sono reali), purché si esegua la EG impiegando solo le matrici elementari del tipo $\mathbf{E}_{ij}(\alpha)$ e pervenendo a una forma a scalini di $\mathbf{A}$, denotata con $\mathbf{U}_0$, contenente sulla diagonale i pivot d_k.

  Per $k = 1$, ossia matrice 1×1 dal determinante (l'unico elemento) positivo; pertanto si ha $d_1 = a_{11} > 0$, perché è il determinante di $\mathbf{A}_1$. Sia $k > 1$ e per ipotesi i pivot sino al $k-1$ siano positivi, ossia $d_i > 0$ per $i \leq k-1$. Le matrici elementari $\mathbf{E}_{ij}(\alpha)$ impiegate non modificano tramite premoltiplicazione il determinante della matrice $\mathbf{A}$, per definizione stessa di determinante e osservazione 236. Le corrispondenti operazioni elementari trasformano $\mathbf{A}_k$ nella sottomatrice principale k-esima di $\mathbf{U}_0$. Pertanto $\det \mathbf{A}_k = \det \mathbf{U}_{0k}$, ed essendo $\mathbf{U}_{0k}$ triangolare superiore con elementi diagonali $d_1, \dots, d_k$, ha come determinante il prodotto $\det \mathbf{U}_{0k} = \prod_{i=1}^k d_i$. Pertanto $\det \mathbf{A}_k = d_1 \cdot \ldots \cdot d_k$. Ne consegue che

$$d_k = \underbrace{\det \mathbf{A}_k}_{>0(1)} / \underbrace{\det \mathbf{A}_{k-1}}_{>0(1)} > 0$$

dove (1) perché sappiamo che ogni sottomatrice principale k-esima ha determinante positivo per ipotesi del punto 3.
A questo punto osserviamo che

$$\mathbf{U}_0 = \mathbf{Diag}(d_1, \dots, d_n)\mathbf{U}$$

dove $\mathbf{U}$ è forma ridotta di Gauss di $\mathbf{A}$, quella dove tutti i pivot sono 1: premoltiplicando per matrice diagonale moltiplico ciascuna riga per i pivot d_i). Si ha quindi

$$\mathbf{A} \overset{(1)}{=} \mathbf{L}_0\mathbf{U}_0 = \mathbf{L}_0\mathbf{Diag}(d_1, \dots, d_n)\mathbf{U} = \mathbf{L}_0\mathbf{D}\mathbf{U}$$

(con (1) dato dalla fattorizzazione LU), dove $\mathbf{L}_0$ è uni-triangolare inferiore (poiché inversa del prodotto di matrici $E_{ij}(\alpha)$, per proposizione 1.8.3 e fine dimostrazione teorema 1.8.7, $\mathbf{U}$ è uni-triangolare superiore e $\mathbf{D} = \mathbf{Diag}(d_1, \dots, d_n)$ con elementi diagonali reali positivi. Infine, da $\mathbf{A} = \mathbf{A}^H$ segue facilmente che $\mathbf{U} = \mathbf{L}_0^H$ da cui $\mathbf{A} = \mathbf{L}_0\mathbf{D}\mathbf{L}_0^H$.

- 4 $\implies$ 5: basta porre $\mathbf{W} = \mathbf{L}_0\mathbf{D}^{1/2}$, dove $\mathbf{D}^{1/2} = \mathbf{Diag}\left(d_1^{1/2}, \dots, d_n^{1/2}\right)$, infatti

$$\mathbf{W}\mathbf{W}^H = (\mathbf{L}_0\mathbf{D}^{1/2})(\mathbf{L}_0\mathbf{D}^{1/2})^H = \mathbf{L}_0 \underbrace{\mathbf{D}^{1/2}\mathbf{D}^{1/2^H}}_{=\mathbf{D}(1)} \mathbf{L}_0^H = \mathbf{L}_0\mathbf{D}\mathbf{L}_0^H = \mathbf{A}$$

con (1) perché moltiplico la radice della diagonale di $\mathbf{D}$ per se stessa

Definizione 6.3.7 (Decomposizione di Choleski)**.** Data una matrice hermitiana $\mathbf{A}$, è chiamata così la sua fattorizzazione:

$$\mathbf{A} = (\mathbf{L}_0\mathbf{D}^{1/2})(\mathbf{L}_0\mathbf{D}^{1/2})^H$$

- 5 $\implies$ 2: per ogni vettore $\mathbf{x}$ di norma euclidea 1 risulta

$$\rho_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^H\mathbf{A}\mathbf{x} = \mathbf{x}^H \underbrace{\mathbf{W}\mathbf{W}^H}_{\mathbf{A}} \mathbf{x} = (\mathbf{x}^H\mathbf{W})(\mathbf{x}^H\mathbf{W})^H = \left\|\mathbf{W}^H\mathbf{x}\right\|_2^2$$

Tale norma risulta positiva, perché la matrice $\mathbf{W}$ è invertibile e da questo si ha che (teorema 1.6.13) $\mathbf{W}^H$ è invertibile e quindi (teorema 1.6.12); $\mathbf{W}^H\mathbf{x} \neq \mathbf{0}$

- 1 $\iff$ 6: $\mathbf{A}$ è hermitiana, quindi normale (esercizio 6.2.1) quindi per il teorema spettrale (teorema 6.2.1) è unitariamente diagonalizzabile. Se $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$ è una diagonalizzazione unitaria di $\mathbf{A}$, essendo gli autovalori (residenti sulla diagonale di $\mathbf{\Lambda}$) λ_i positivi (per ipotesi 1) possiamo porre: $\mathbf{\Lambda}^{1/2} = \mathbf{Diag}\left(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}\right)$. Si ha allora che $\mathbf{A} = \mathbf{B}^2$ in quanto:

$$\mathbf{B}^2 = \mathbf{U}\mathbf{\Lambda}^{1/2}\underbrace{\mathbf{U}\mathbf{U}^H}_{\mathbf{I}}\mathbf{\Lambda}^{1/2}\mathbf{U}^H = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{U}^H = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \mathbf{A}$$

dove $\mathbf{B} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{U}^H$ è evidentemente una matrice definita positiva: infatti $\mathbf{B}$ è diagonalizzata quindi simile a $\mathbf{A}$ da $\mathbf{\Lambda}^{1/2}$ quindi ha gli stessi autovalori (5.5.2), ed essendo $\mathbf{\Lambda}^{1/2}$ diagonale sono gli elementi della diagonale, tutti positivi.
Viceversa, se $\mathbf{A} = \mathbf{B}^2$ con $\mathbf{B}$ matrice definita positiva, poiché gli autovalori di $\mathbf{B}$ sono reali non nulli e quelli di $\mathbf{B}^2$ sono i quadrati degli autovalori di $\mathbf{B}$, si deduce che gli autovalori di $\mathbf{A}$ sono positivi

- 3 $\iff$ 7: se ogni sottomatrice principale k-esima di $\mathbf{A}$ ha determinante positivo, in particolare si ha che $\det \mathbf{A}_{n-1} > 0$, quindi $\mathbf{A}_{n-1}$ è invertibile e inoltre per ipotesi (7) si ha che

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{x} \\ \mathbf{x}^H & a \end{bmatrix}$$

La formula del determinante a blocchi (equazione 4.15) applicato al presente caso conduce a

$$\det \mathbf{A} = \det(\mathbf{A}_{n-1}) \cdot (a - \mathbf{x}^H \mathbf{A}_{n-1}^{-1}\mathbf{x})$$

da cui, sapendo che $\det \mathbf{A} > 0$ e che $\det \mathbf{A}_{n-1} > 0$ deve essere $(a - \mathbf{x}^H \mathbf{A}_{n-1}^{-1}\mathbf{x}) > 0$, pertanto si deduce che $a > \mathbf{x}^H\mathbf{A}^{-1}\mathbf{x}$. Le sottomatrici principali k-esime di $\mathbf{A}_{n-1}$ sono le sottomatrici principali k-esime di $\mathbf{A}$ per $k \leq n-1$. pertanto esse hanno determinante positivo (per ipotesi). Sapendo già che $3 \implies 1$ ($3 \implies 4 \implies 5 \implies 2 \implies 1$) si deduce che $\mathbf{A}_{n-1}$ è definita positiva.
Viceversa, sapendo che $\mathbf{A}_{n-1}$ è definita positiva, da $1 \implies 3$ si deduce che le sottomatrici principali k-esime di $\mathbf{A}$ hanno determinante positivo per $k \leq n-1$ e resta da provare che $\det \mathbf{A} > 0$. Ciò segue dalla ricordata formula del determinante a blocchi e dall'ipotesi che $a > \mathbf{x}^H \mathbf{A}_{n-1}^{-1}\mathbf{x}$

□

Osservazione 353. Il risultato analogo al teorema 6.3.8 per le matrici semi-definite positive si differenzia di poco dal suddetto teorema. Vale la pena di osservare che nell'enunciato del teorema 6.3.9 si richiede

- nel punto (3) che abbiano determinante non negativo *tutte* le sottomatrici principali di $\mathbf{A}$ e non solo le sottomatrici principali k-esime;
- non necessariamente 5 e 6 le matrici $\mathbf{W}$ e $\mathbf{B}$ che fattorizzano $\mathbf{A}$ siano invertibili;
- il punto 7, su cui torneremo dopo la dimostrazione è quello che più si allontana dal medesimo punto del 6.3.8

Teorema 6.3.9 (Caratterizzazione matrici semi-definite positive)**.** *Data la matrice hermitiana* $\mathbf{A}$ *di ordine* n*, i fatti seguenti sono equivalenti:*

1. *gli autovalori di* $\mathbf{A}$ *sono non negativi*
2. *la funzione di Rayleigh assume valori non negativi:* $\rho_{\mathbf{A}}(\mathbf{x}) \geq 0$ *per ogni vettore* $\mathbf{x}$
3. *ogni sottomatrice principale di* $\mathbf{A}$ *(*non solo k-esime*) ha determinante non negativo*
4. $\mathbf{A} = \mathbf{L}_0\mathbf{D}\mathbf{L}_0^H$ *con* $\mathbf{L}_0$ *matrice uni-triangolare inferiore e* $\mathbf{D}$ *matrice diagonale con elementi diagonali reali non negativi*
5. *esiste una matrice (*non necessariamente invertibile*)* $\mathbf{W}$ *tale che* $\mathbf{A} = \mathbf{W}\mathbf{W}^H$
6. *esiste una matrice semi-definita positiva* $\mathbf{B}$ *tale che* $\mathbf{A} = \mathbf{B}^2$
7. *se* $\mathbf{A}$ *è in forma bordata:*

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{x} \\ \mathbf{x}^H & a \end{bmatrix}$$

   allora $\mathbf{A}_{n-1}$ *è semi-definita positiva,* $a \geq \mathbf{x}^H\mathbf{A}_{n-1}^{+}\mathbf{x}$ *e il vettore* $\mathbf{x}$ *appartiene allo spazio delle colonne di* $\mathbf{A}_{n-1}$

Osservazione 354. $\mathbf{B}$ del punto 6 non è necessariamente invertibile dato che il determinante è prodotto degli autovalori e tra questi si può avere anche 0, essendo $\mathbf{B}$ semi-definita positiva.

Dimostrazione. Rispettivamente:

- $1 \iff 2$: si ha che $0 \leq \lambda_1 \leq \rho_{\mathbf{A}}(\mathbf{x})$, quindi $\rho_{\mathbf{A}}(\mathbf{x}) \geq 0$
- $1 \iff 3$: osserviamo che $\det(\mathbf{A}) \geq 0$ (determinante prodotto degli autovalori). Dimostriamo che anche $\mathbf{A}'$, sottomatrice principale di $\mathbf{A}$ di dimensione $m \leq n$ è semi-definita positiva, e quindi ha determinante non negativo. Abbiamo che anche $\mathbf{A}'$ è hermitiana: si ha che gli autovalori di $\mathbf{A}'$ sono ≥ 0 se e solo se la funzione di Rayleigh $\rho_{\mathbf{A}'}(\mathbf{x}) \geq 0$. Procediamo dunque nella dimostrazione di questo secondo punto, ossia $\mathbf{y}^H\mathbf{A}'\mathbf{y} \geq 0$, per ogni $\mathbf{y} \in \mathbb{C}^m$, con $\mathbf{y} \neq 0$.
  Anche qui (sebbene esprimerlo con indici in maniera generale sia laborioso) se $\mathbf{z}$ è $\mathbf{y}$ circondato dagli opportuni 0 in maniera tale che la lunghezza di $\mathbf{z}$ sia n (e notando che $z \neq 0$) si ha che:

$$\mathbf{y}^H\mathbf{A}'\mathbf{y} = \underbrace{\mathbf{z}^H\mathbf{A}\mathbf{z}}_{\rho_{\mathbf{A}}(\mathbf{z})} \overset{(1)}{\geq} 0$$

  dove in (1) la funzione di Rayleigh-Ritz assume solo valori positivi o nulli(per ipotesi ci fondiamo sul punto 1 ma diamo per buono il già dimostrato $1 \implies 2$)
- $4 \implies 5$: ponendo $\mathbf{W} = \mathbf{L}_0\mathbf{D}^{1/2}$ con $\mathbf{D}^{1/2} = \mathbf{Diag}\left(d_1^{1/2}, \ldots, d_n^{1/2}\right)$ si ha che $\mathbf{W}$ è invertibile solo se $\mathbf{D}^{1/2}$ lo è, ossia se tutti gli elementi della diagonale sono > 0). Si ha infatti che

$$\mathbf{W}\mathbf{W}^H = (\mathbf{L}_0\mathbf{D}^{1/2})(\mathbf{L}_0\mathbf{D}^{1/2})^H = \mathbf{L}_0\underbrace{\mathbf{D}^{1/2}\mathbf{D}^{1/2}}_{=\mathbf{D}}\mathbf{L}_0^H = \mathbf{L}_0\mathbf{D}\mathbf{L}_0^H = \mathbf{A}$$

- $5 \implies 2$: per ogni vettore $\mathbf{x}$ di norma euclidea si ha
$$\rho_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^H \mathbf{W}\mathbf{W}^H \mathbf{x} = (\mathbf{W}^H\mathbf{x})^H(\mathbf{W}^H\mathbf{x}) = \left\|\mathbf{W}^H\mathbf{x}\right\|_2^2$$
La norma risulta > 0 se $\mathbf{W}$ è invertibile (e dunque anche $\mathbf{W}^H$ è invertibile); se invece $\mathbf{W}$ non è invertibile $\mathbf{W}^H$ non è invertibile e vi sono vettori $\mathbf{x} \neq 0$ per i quali $\mathbf{W}^H\mathbf{x} = \mathbf{0}$ e dunque $\left\|\mathbf{W}^H\mathbf{x}\right\|_2^2 = \rho_{\mathbf{A}}(\mathbf{x}) = 0$. In conclusione $\rho_{\mathbf{A}}(\mathbf{x}) \geq 0$

- $1 \iff 6$: $\mathbf{A}$ è sempre hermitiana, poniamo sempre $\mathbf{\Lambda}^{1/2} = \mathbf{Diag}\left(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2}\right)$.
Si ha sempre che $\mathbf{A} = \mathbf{B}^2$ con $\mathbf{B} = \mathbf{U}\mathbf{\Delta}^{1/2}\mathbf{U}^H$ è semi-definita positiva, condividendo gli autovalori con $\mathbf{\Delta}^{1/}$.
Viceversa se $\mathbf{A} = \mathbf{B}^2$ con $\mathbf{B}$ semi definita positiva gli autovalori di $\mathbf{A}$ sono nulli o positivi, essendo i quadrati di quelli di $\mathbf{B}$ (nulli o positivi);

- $3 \implies 4$: per induzione sull'ordine n della matrice. Il caso $n = 1$ è banale: se la dimensione di $\mathbf{A}$ è 1, l'unica sottomatrice principale possibile è $\mathbf{A}_1 = \mathbf{A} = [a_{11}]$, $\det(A_1) = a_{11}$. Ipotizzando il determinante ≥ 0 si ha che $a_{11} \geq 0$. A questo punto la fattorizzazione di matrice 1×1 sarà del tipo $a_{11} = 1 \cdot d_0 \cdot 1$ da cui $a_{11} = d_0$.
Sia allora $n > 1$ e l'asserto per ipotesi vero per $n - 1$. Decomponiamo la matrice $\mathbf{A}$ a blocchi nel modo seguente:
$$\mathbf{A} = \begin{bmatrix} a & \mathbf{x}^H \\ \mathbf{x} & \mathbf{A}_{11} \end{bmatrix}$$
con A_{11} che è sottomatrice principale. Per l'ipotesi induttiva, dato che sottomatrice principale, $\mathbf{A}_{11} = \mathbf{L}_1\mathbf{D}_1\mathbf{L}_1^H$ con $\mathbf{L}_1$ matrice uni-triangolare inferiore e $\mathbf{D}_1$ matrice diagonale con elementi diagonali reali non negativi. Se $\mathbf{x} = \mathbf{0}$ basta porre $\mathbf{L}_0 = \mathbf{Diag}(1, L_1)$ e $\mathbf{D} = \mathbf{Diag}(a, \mathbf{D}_1)$ e si ha che
$$\begin{aligned}
\mathbf{L}_0\mathbf{D}\mathbf{L}_0^H &= \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0}^T & \mathbf{L}_1 \end{bmatrix} \begin{bmatrix} a & \mathbf{0} \\ \mathbf{0}^T & \mathbf{D}_1 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0}^T & \mathbf{L}_1^H \end{bmatrix} \\
&= \begin{bmatrix} a & \mathbf{0} \\ \mathbf{0}^T & \mathbf{L}_1\mathbf{D}_1 \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0}^T & \mathbf{L}_1^H \end{bmatrix} \\
&= \begin{bmatrix} a & \mathbf{0} \\ \mathbf{0}^T\mathbf{L}_1\mathbf{D}_1\mathbf{L}_1^H \end{bmatrix} = \begin{bmatrix} a & \mathbf{0} \\ \mathbf{0}^T & \mathbf{A}_{11} \end{bmatrix} = \mathbf{A}
\end{aligned}$$
Se invece $\mathbf{x} \neq \mathbf{0}$, allora per proposizione 6.3.7 necessariamente $a > 0$. Si esegua dunque l'eliminazione simmetrica su prima riga e colonna, premoltiplicando e post-moltiplicando successivamente per le matrici elementari $\mathbf{E}_{i1}(-a^{-1}a_{i1})$ (toglie a tutte le righe la prima, annullando il primo coefficiente risultante, ossia prima colonna tutta nulla ad eccezione di a_{11}) e $\mathbf{E}_{1i}(-a^{-1}a_{1i})$ (che toglie a tutte le colonne la prima moltiplicata in modo da annullare tutta la prima riga ad eccezione di a_{11}) rispettivamente. Si osservi che:
$$\mathbf{E}_{i1}(-a^{-1}a_{i1})^H \overset{(1)}{=} \mathbf{E}_{1i}(-a^{-1}\overline{a_{i1}}) \overset{(2)}{=} \mathbf{E}_{1i}(-a^{-1}a_{1i})$$
dove in (1) abbiamo trasposto e coniugato, mentre in (2) l'uguaglianza vale dato che $\mathbf{A}$ è hermitiana dunque $\overline{a_{1i}} = a_{1i}$.

Se $\mathbf{E}$ denota il prodotto delle matrici elementari che pre-moltiplicano $\mathbf{A}$, a seguito di tutte le varie moltiplicazioni avremo una configurazione come la seguente

$$\mathbf{EAE}^H = \begin{bmatrix} a & \mathbf{0}^H \\ \mathbf{0} & \mathbf{C} \end{bmatrix}$$

dove $\mathbf{E}$ è il prodotto delle varie $\mathbf{E}_{i1}$ e $\mathbf{E}^H$ sono le E_{1i} che, come visto sono uguali alle E_{i1} trasposte.
Ci siamo ricondotti dunque al caso con $\mathbf{x} = \mathbf{0}$ e per quanto già visto la matrice cui siamo giunti può essere fattorizzata come $\mathbf{EAE}^H = \mathbf{LDL}^H$, con $\mathbf{L}$ matrice unitriangolare inferiore e $\mathbf{D}$ matrice diagonale con elementi diagonali reali non negativi. Allora per ottenere una fattorizzazione di $\mathbf{A}$ procediamo come segue

$$\begin{aligned} \mathbf{EAE}^H &= \mathbf{LDL}^H \\ \mathbf{A} &= \mathbf{E}^{-1}\mathbf{LDL}^H(\mathbf{E}^H)^{-1} = \mathbf{E}^{-1}\mathbf{LDL}^H(\mathbf{E}^{-1})^H \\ &= \mathbf{E}^{-1}\mathbf{LD}(\mathbf{E}^{-1}\mathbf{LD})^H \end{aligned}$$

Basta allora porre $\mathbf{L}_0 = \mathbf{E}^{-1}\mathbf{L}$, ricordando che prodotti e inverse di matrici uni-triangolari inferiori sono ancora tali.

- 5 $\implies$ 7: posto $\mathbf{W} = \begin{bmatrix} \mathbf{Y} \\ \mathbf{w}^H \end{bmatrix}$ e $\mathbf{W}^H = [\mathbf{Y}^H \;\; \mathbf{w}]$ dal prodotto a blocchi $\mathbf{WW}^H$

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{w}^H \end{bmatrix} \begin{bmatrix} \mathbf{y}^H & \mathbf{w} \end{bmatrix} = \begin{bmatrix} \mathbf{YY}^H & \mathbf{Yw} \\ w^H Y^H & \mathbf{w}^H\mathbf{w} \end{bmatrix}$$

Uguagliandolo il risultato ad $\mathbf{A}$ si ha che:

$$\mathbf{YY}^H = \mathbf{A}_{n-1}, \quad \mathbf{Yw} = \mathbf{x}, \quad \mathbf{w}^H\mathbf{w} = a$$

Dalla prima uguaglianza segue che $\mathbf{A}_{n-1}$ è semi-definita positiva, perché già sappiamo che 5 $\implies$ 1 (ed abbiamo provato il primo punto di 7); dalla seconda segue che $\mathbf{x} \in \mathrm{C}(\mathbf{Y})$ ed è facile il terzo punto, ossia che $\mathrm{C}(\mathbf{Y}) = \mathrm{C}\left(\mathbf{YY}^H\right) = \mathrm{C}(\mathbf{A}_{11})$. Infine, attesa la terza uguaglianza, vogliamo provare che $a > \mathbf{x}^H\mathbf{A}_{n-1}^+\mathbf{x}$, ma $a = \mathbf{w}^H\mathbf{w} = \|\mathbf{w}\|_2^2$, quindi si tratta di provare che $\|\mathbf{w}\|_2^2 \geq \mathbf{x}^H\mathbf{A}_{n-1}^+\mathbf{x}$. Tenuto conto dell' esercizio 32 del capitolo 1 risulta:

TODO: fixme

TODO: rivedere qui

$$\mathbf{x}^H\mathbf{A}_{n-1}^+\mathbf{x} = \mathbf{x}^H(\mathbf{YY}^H)^+\mathbf{x} \overset{(1)}{=} \underbrace{\mathbf{x}^H(\mathbf{Y}^+)^H}_{=(\mathbf{Y}^+\mathbf{x})^H}\mathbf{Y}^+\mathbf{x} = \left\|\mathbf{Y}^+\mathbf{x}\right\|_2^2 \overset{(2)}{=} \left\|\mathbf{Y}^+\mathbf{Yw}\right\|_2^2$$

non chiarissimo i passaggi (1) e (2) Perciò dobbiamo provare, togliendo equivalentemente la potenza, che $\|\mathbf{w}\|_2 \geq \|\mathbf{Y}^+\mathbf{Yw}\|_2$. Ma $\mathbf{Y}^+\mathbf{Yw}$ coincide con la proiezione di $\mathbf{w}$ sul sottospazio $\mathrm{C}\left(\mathbf{Y}^H\right)$ di $\mathbb{C}^n$, perché la matrice $\mathbf{Y}^+\mathbf{Y}$ risulta essere la matrice di proiezione su $\mathrm{C}\left(\mathbf{Y}^H\right)$ (si veda il corollario 3.7.3), da cui l'asserto segue in modo ovvio (è una proiezione, quindi il modulo diminuisce rispetto al vettore originale o rimane al massimo uguale);

TODO: fixme

- 7 $\implies$ 1: dato che per ipotesi $\mathbf{A}_{n-1}$ è semi-definita positiva (ossia per essa vale (1), e quindi (5)) possiamo trovare una fattorizzazione $\mathbf{A}_{n-1} = \mathbf{Y}\mathbf{Y}^H$ per una opportuna matrice $\mathbf{Y}$. Proviamo che risulta semi-definita positiva la matrice:
$$\mathbf{B} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{x} \\ \mathbf{x}^H & \|\mathbf{Y}^+\mathbf{x}\|_2^2 \end{bmatrix}$$
Ciò deriva dal fatto che $\mathbf{B}$ può essere scritta come $\mathbf{B} = \mathbf{Z}\mathbf{Z}^H$, ossia essendo hermitiana vale anche per essa il punto (5) e possiamo scrivere
$$\mathbf{B} = \mathbf{Z}\mathbf{Z}^H = \begin{bmatrix} \mathbf{Y} \\ (\mathbf{Y}^+\mathbf{x})^H \end{bmatrix} \begin{bmatrix} \mathbf{Y}^H & \mathbf{Y}^+\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}\mathbf{Y}^H & \mathbf{Y}\mathbf{Y}^+\mathbf{x} \\ \mathbf{x}^H(\mathbf{Y}^+)^H\mathbf{Y}^H & \|\mathbf{Y}^+\mathbf{x}\|_2^2 \end{bmatrix} = \begin{bmatrix} \mathbf{Y}\mathbf{Y}^H & \mathbf{Y}\mathbf{Y}^+\mathbf{x} \\ (\mathbf{Y}\mathbf{Y}^+\mathbf{x})^H & \|\mathbf{Y}^+\mathbf{x}\|_2^2 \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{x} \\ \mathbf{x}^H & \|\mathbf{Y}^+\mathbf{x}\|_2^2 \end{bmatrix}$$
dove l'uguaglianza $\mathbf{Y}\mathbf{Y}^+\mathbf{x} = \mathbf{x}$ segue da
$$\mathbf{Y}\mathbf{Y}^+\mathbf{x} \overset{(1)}{=} \mathbf{Y}\mathbf{Y}^+\mathbf{Y}\mathbf{w} \overset{(2)}{=} \mathbf{Y}\mathbf{w} = \mathbf{x}$$
dove (1) dato che se $x \in \mathrm{C}(\mathbf{Y})$, allora $\mathbf{x} = \mathbf{Y}\mathbf{w}$ per un certo vettore $\mathbf{w}$, (2) per la definizione di pseudo-inversa ($\mathbf{Y}\mathbf{Y}^+\mathbf{Y} = \mathbf{Y}$).
Quindi $\mathbf{B}$ è semi-definita positiva perché 5 $\iff$ 1. Ricordiamo che al punto precedente abbiamo visto che $\mathbf{x}^H\mathbf{A}_{n-1}^+\mathbf{x} = \|\mathbf{Y}^+\mathbf{x}\|$. Poniamo
$$\mathbf{C} = \begin{bmatrix} \mathbb{O} & \mathbf{0} \\ \mathbf{0}^H & a - \mathbf{x}^H\mathbf{A}_{n-1}^+\mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbb{O} & \mathbf{0} \\ \mathbf{0}^H & a - \|\mathbf{Y}^+\mathbf{x}\| \end{bmatrix}$$
Si ha che $\mathbf{A} = \mathbf{B} + \mathbf{C}$, in quanto
$$\mathbf{B} + \mathbf{C} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{x} \\ \mathbf{x}^H & \|\mathbf{Y}^+\mathbf{x}\|_2^2 \end{bmatrix} + \begin{bmatrix} \mathbb{O} & \mathbf{0} \\ \mathbf{0}^H & a - \|\mathbf{Y}^+\mathbf{x}\| \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{n-1} & \mathbf{x} \\ \mathbf{x}^H & a \end{bmatrix} = \mathbf{A}$$
Passando alla funzione di Rayleigh-Ritz di $\mathbf{A}$, pertanto:
$$\begin{aligned} \rho_{\mathbf{A}}(\mathbf{x}) &= \frac{\mathbf{x}^H\mathbf{A}\mathbf{x}}{\mathbf{x}^H\mathbf{x}} = \frac{1}{\mathbf{x}^H\mathbf{x}}(\mathbf{x}^H\mathbf{A}\mathbf{x}) = \frac{1}{\mathbf{x}^H}\left[\mathbf{x}^H(\mathbf{B}+\mathbf{C})\mathbf{x}\right] \\ &= \underbrace{\frac{\mathbf{x}^H\mathbf{B}\mathbf{x}}{\mathbf{x}^H\mathbf{x}}}_{\geq 0\ (1)} + \underbrace{\frac{\mathbf{x}^H\mathbf{C}\mathbf{x}}{\mathbf{x}^H\mathbf{x}}}_{\geq 0\ (1)} \geq 0 \end{aligned}$$
dove in (1) gli addendi sono non negativi dato che sia $\mathbf{B}$ che $\mathbf{C}$ sono semi-definite positive. Quindi $\rho_{\mathbf{A}}(\mathbf{x}) \geq 0$ e abbiamo mostrato che anche $\mathbf{A}$ è semi-definita positiva.

□

Osservazione 355. Si osservi che i punti 7 dei due precedenti teoremi danno un criterio induttivo per verificare se una matrice hermitiana è definita o semi-definita positiva.

Osservazione 356. Si noti che nel teorema 6.3.9 la matrice pseudo-inversa $\mathbf{A}_{n-1}^+$ sostituisce la matrice inversa $\mathbf{A}_{n-1}^{-1}$ usata nel teorema 6.3.8, e che la condizione aggiuntiva: $\mathbf{x} \in \mathrm{C}(\mathbf{A}_{n-1})$, è automaticamente verificata quando $\mathbf{A}_{n-1}$ è invertibile.

6.3.4.1 Radice quadrata di matrice semi-definita positiva

Definizione 6.3.8 (Matrice radice quadrata)**.** Una matrice semi-definita positiva $\mathbf{B}$ tale che $\mathbf{A} = \mathbf{B}^2$ (di cui il teorema 6.3.9 punto 6 assicura l'esistenza), si chiama radice quadrata della matrice semi-definita positiva $\mathbf{A}$

Proposizione 6.3.10. *La radice quadrata di una matrice hermitiana semi-definita positiva* $\mathbf{A}$ *è unica.*

Dimostrazione. Supponiamo che $\mathbf{A} = \mathbf{B}^2 = \mathbf{C}^2$ con $\mathbf{B}$ e $\mathbf{C}$ matrici hermitiane semi-definite positive con $\mathbf{B} \neq \mathbf{C}$ per assurdo. Le matrici $\mathbf{B}$ e $\mathbf{C}$ hanno gli stessi autovalori con le stesse molteplicità, perché i loro quadrati sono gli autovalori di $\mathbf{A}$; per esteso (focalizzandoci su $\mathbf{B}$) dato che $\mathbf{B}$ è hermitiana semi-definita positiva può essere diagonalizzata come $\mathbf{B} = \mathbf{L}_0\mathbf{D}\mathbf{L}_0^H$, ma $\mathbf{L}_0$ è una matrice risultato del prodotto di $E_{ij}(\alpha)$ quindi è unitaria e dunque $\mathbf{L}_0\mathbf{L}_0^H = \mathbf{L}_0^H\mathbf{L}_0 = \mathbf{I}$, pertanto

$$\mathbf{A} = \mathbf{B}\mathbf{B} = \mathbf{L}_0\mathbf{D}\underbrace{\mathbf{L}_0^H\mathbf{L}_0}_{\mathbf{I}}\mathbf{D}\mathbf{L}_0^H = \mathbf{L}_0\mathbf{D}^2\mathbf{L}_0^H$$

ed $\mathbf{A}$ ha gli stessi autovalori di $\mathbf{D}^2$, che sono quelli di $\mathbf{B}$ al quadrato. Allo stesso modo avviene per $\mathbf{C}$.
Siano allora $\lambda_1, \ldots, \lambda_r$ i distinti autovalori, dato che $\mathbf{B}$ e $\mathbf{C}$ sono hermitiane quindi normali, quindi unitariamente diagonalizzabili quindi scrivibili mediante decomposizione spettrale (versione additiva) come:

$$\mathbf{B} = \lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r, \qquad \mathbf{C} = \lambda_1\mathbf{Q}_1 + \ldots + \lambda_r\mathbf{Q}_r$$

due decomposizioni spettrali (additive) di $\mathbf{B}$ e $\mathbf{C}$, rispettivamente. Usando il fatto che per $i \neq j$ risulta $\mathbf{P}_i\mathbf{P}_j = \mathbb{O} = \mathbf{Q}_i\mathbf{Q}_j$ (da dimostrazione del teorema 6.2.2), da $\mathbf{A} = \mathbf{B}^2 = \mathbf{C}^2$, e dal fatto che sono matrici di proiezione (quindi $\mathbf{P}_i^2 = \mathbf{P}_i$ e $\mathbf{Q}_i^2 = \mathbf{Q}_i$) si ricava:

$$\underbrace{\lambda_1^2\mathbf{P}_1 + \ldots + \lambda_r^2\mathbf{P}_r}_{\mathbf{B}^2} = \mathbf{A} = \underbrace{\lambda_1^2\mathbf{Q}_1 + \ldots + \lambda_r^2\mathbf{Q}_r}_{\mathbf{C}^2}$$

Se $i \neq j$, moltiplicando primo e ultimo membro a sinistra per $\mathbf{P}_i$ e a destra per $\mathbf{Q}_j$ (ad esempio per quanto riguarda $\mathbf{B}$, il primo membro, così facendo elimino tutti gli addendi che non siano $\lambda_i\mathbf{P}_i$ perché $\mathbf{P}_i\mathbf{P}_j = \mathbb{O}$, poi aggiungo a destra $\mathbf{Q}_j$ ed ottengo $\lambda_i^2\mathbf{P}_i\mathbf{Q}_j$) si ricava

$$\lambda_i^2\mathbf{P}_i\mathbf{Q}_j = \lambda_j^2\mathbf{P}_i\mathbf{Q}_j$$

da cui si ottiene che $\mathbf{P}_i\mathbf{Q}_j = \mathbb{O}$ (dato che l'uguaglianza vale ma $\lambda_i \neq \lambda_j$ deve essere $\mathbf{P}_i\mathbf{Q}_j = \mathbb{O}$). Invertendo i ruoli di $\mathbf{P}_i$ e $\mathbf{Q}_j$ si ricava anche $\mathbf{Q}_j\mathbf{P}_i = \mathbb{O}$.
Consideriamo l'uguaglianza

$$\mathbf{P}_1 + \ldots + \mathbf{P}_r = \mathbf{I} = \mathbf{Q}_1 + \ldots + \mathbf{Q}_r \tag{6.6}$$

questa è ottenuta (credo) perché $\mathbf{I}$ è unitaria ($\mathbf{I}^{-1} = \mathbf{I}^T = \mathbf{I}$), dunque normale e si può scrivere $\mathbf{I} = \lambda_1\mathbf{P}_1 + \ldots + \lambda_r\mathbf{P}_r$; tuttavia si ha che $\mathbf{I}\mathbf{v} = \lambda_i\mathbf{v} \iff \lambda_i = 1$ quindi si ha che $\mathbf{I} = \mathbf{P}_1 + \ldots + \mathbf{P}_r$. Analogamente avviene per $\mathbf{I} = \mathbf{Q}_1 + \ldots + \mathbf{Q}_r$
Tornando all'uguaglianza di equazione 6.6 moltiplicando entrambi i membri prima per $\mathbf{P}_i$ e poi per $\mathbf{Q}_i$, si ricava, per ogni indice i:

$$\mathbf{P}_i = \mathbf{P}_i\mathbf{Q}_i = \mathbf{Q}_i\mathbf{P}_i = \mathbf{Q}_i$$

ossia rimangono solo i termini con indice concorde perché $\mathbf{P}_i\mathbf{Q}_j = \mathbf{Q}_j\mathbf{P}_i = \mathbb{O}$.
Si può allora concludere che $\mathbf{B} = \mathbf{C}$ $\square$

6.4 Decomposizione in valori singolari

Osservazione 357. Data una matrice quadrata $\mathbf{A}$ di ordine n, una matrice hermitiana ad essa naturalmente associata è la cosiddetta parte hermitiana (proposizione 1.3.9), ossia $(\mathbf{A} + \mathbf{A}^H)/2$, avente la stessa dimensione di $\mathbf{A}$

Osservazione 358. Esistono altre due matrici hermitiane naturalmente associate ad $\mathbf{A}$:

$$\mathbf{A}\mathbf{A}^H \qquad \mathbf{A}^H\mathbf{A}$$

È facile verificare che sono hermitiane in quanto, ad esempio $(\mathbf{A}\mathbf{A}^H)^H = \mathbf{A}\mathbf{A}^H$. Essendo hermitiane, quindi normali, queste due matrici sono unitariamente diagonalizzabili.

Osservazione 359 (Decomposizione in valori singolari). Una opportuna combinazione delle diagonalizzazioni di $\mathbf{A}\mathbf{A}^H$ e $\mathbf{A}^H\mathbf{A}$ produce una nuova fattorizzazione della matrice $\mathbf{A}$, chiamata *decomposizione in valori singolari di* $\mathbf{A}$.

Osservazione 360. Iniziamo lo studio delle matrici $\mathbf{A}\mathbf{A}^H$ e $\mathbf{A}^H\mathbf{A}$ con un esame dei loro autovalori e delle relative molteplicità.
Partiamo però da un risultato più generale per una coppia di matrici $\mathbf{A}$ e $\mathbf{B}$ di dimensioni rispettivamente $m \times n$ e $n \times m$, per le quali sono eseguibili quindi i due prodotti $\mathbf{AB}$ e $\mathbf{BA}$.

Proposizione 6.4.1. *Date due matrici complesse* $\mathbf{A}$ *e* $\mathbf{B}$ *di dimensioni rispettivamente* $m \times n$ *e* $n \times m$*, le due matrici* $\mathbf{AB}$ *e* $\mathbf{BA}$ *hanno gli stessi autovalori non nulli con le stesse molteplicità algebriche e geometriche.*

Dimostrazione. Siano $0 \neq \lambda \in \mathbb{C}$ e $\sigma \in \mathbb{C}$ tali che $\sigma^2 = \lambda$ ($\sigma = \sqrt{\lambda}$). Consideriamo le due trasformazioni

$$\varphi_\lambda : E_{\mathbf{AB}}(\lambda) \to E_{\mathbf{BA}}(\lambda), \quad \varphi_\lambda(\mathbf{u}) = \frac{1}{\sigma}\mathbf{Bu} \tag{6.7}$$

$$\psi_\lambda : E_{\mathbf{BA}}(\lambda) \to E_{\mathbf{AB}}(\lambda), \quad \psi_\lambda(\mathbf{v}) = \frac{1}{\sigma}\mathbf{Av} \tag{6.8}$$

Sono trasformazioni lineari in quanto (ad esempio per φ):

$$\varphi(\alpha\mathbf{u}) = \frac{1}{\sigma}\mathbf{B}\alpha\mathbf{u} = \alpha\frac{1}{\sigma}\mathbf{Bu} = \alpha\varphi(\mathbf{u})$$
$$\varphi(\mathbf{u}+\mathbf{z}) = \frac{1}{\sigma}\mathbf{B}(\mathbf{u}+\mathbf{z}) = \frac{1}{\sigma}\mathbf{Bu} + \frac{1}{\sigma}\mathbf{Bz} = \varphi(\mathbf{u}) + \varphi(\mathbf{z})$$

Dimostriamo che le due trasformazioni lineari sono ben definite e una l'inversa dell'altra: da ciò si deduce che vi è una biezione tra $E_{\mathbf{AB}(\lambda)}$ ed $E_{\mathbf{BA}(\lambda)}$ per cui λ è un autovalore di $\mathbf{AB}$ con una certa molteplicità geometrica se e solo se è anche autovalore di $\mathbf{BA}$ con la stessa molteplicità geometrica.
A titolo di esempio affrontiamo inizialmente la buona definizione di φ. Sia $\mathbf{u} \in E_{\mathbf{AB}}(\lambda)$; allora per definizione di autovalore ed autovettore si ha

$$\mathbf{ABu} = \lambda\mathbf{u} \tag{6.9}$$

Quindi:

$$\mathbf{BA}\varphi_\lambda(\mathbf{u}) \stackrel{(1)}{=} \frac{1}{\sigma}\mathbf{BABu} \stackrel{(2)}{=} \frac{1}{\sigma}\mathbf{B}\lambda\mathbf{u} \stackrel{(3)}{=} \lambda\varphi_\lambda(\mathbf{u})$$

dove in (1) e (3) si è sostituito usando equazione 6.7, in (2) mediante equazione 6.9. Guardando a primo ed ultimo membro, per la definizione di autovalore e autovettore si ha che λ è autovalore di $\mathbf{BA}$ e $\varphi_\lambda(\mathbf{u})$ è relativo autovalore, in sintesi $\varphi_\lambda(\mathbf{u}) \in E_{\mathbf{BA}}(\lambda)$ e pertanto φ_λ è ben definita. Analoga dimostrazione per ψ_λ.
Proviamo ora che $\psi_\lambda \cdot \varphi_\lambda$ coincide con l'identità. Se $\mathbf{u} \in E_{\mathbf{AB}}(\lambda)$ risulta:

$$(\psi_\lambda \varphi_\lambda)(\mathbf{u}) = \psi_\lambda \left(\frac{1}{\sigma} \mathbf{Bu} \right) = \frac{1}{\sigma} \frac{1}{\sigma} \mathbf{ABu} = \frac{1}{\lambda} \lambda \mathbf{u} = \mathbf{u}$$

Analogamente si prova che $\varphi_\lambda \cdot \psi_\lambda$ coincide con l'identità. Resta da provare che le molteplicità algebriche coincidono: dovrebbe essere che le funzioni, creando una biezione, legano due spazi isomorfi che debbono avere stessa dimensione per proposizione 2.6.6. □

Osservazione 361. Passiamo alle peculiarità di $\mathbf{AA}^H$ e $\mathbf{A}^H\mathbf{A}$

Osservazione 362. Se in proposizione 6.4.1 la matrice $\mathbf{B}$ coincide con $\mathbf{A}^H$, allora i due prodotti $\mathbf{AB}$ e $\mathbf{BA}$ sono entrambi matrici semi-definite positive per il teorema 6.3.9 ($5 \implies 1$), quindi con autovalori reali non negativi.

Osservazione 363. Una seconda particolarità, in aggiunta a proposizione 6.4.1, riguardante le due applicazioni

$$\varphi_\lambda : E_{\mathbf{AA}^H}(\lambda) \to E_{\mathbf{A}^H\mathbf{A}}(\lambda) \tag{6.10}$$

$$\psi_\lambda : E_{\mathbf{A}^H\mathbf{A}}(\lambda) \to E_{\mathbf{AA}^H}(\lambda) \tag{6.11}$$

a seguire.

Proposizione 6.4.2. *Data la matrice complessa $m \times n$ $\mathbf{A}$ e detto λ un suo autovalore non nullo, le due applicazioni φ_λ e ψ_λ, definite nella dimostrazione della proposizione 6.4.1 sostituendo $\mathbf{A}^H$ a $\mathbf{B}$, conservano i prodotti interni dei vettori ($\|\mathbf{v}\|_2 = \|f(\mathbf{v})\|_2$), quindi sono due isometrie*

Dimostrazione. Verifichiamo che φ_λ conserva i prodotti interni; analoga dimostrazione si ha per ψ_λ. Siano $\mathbf{u}, \mathbf{w} \in E_{\mathbf{AA}^H}$ si ha

$$\varphi_\lambda(\mathbf{u})^H \varphi_\lambda(\mathbf{w}) = \left(\frac{1}{\sigma} \mathbf{A}^H \mathbf{u} \right)^H \frac{1}{\sigma} \mathbf{A}^H \mathbf{w} \overset{(1)}{=} \frac{1}{\lambda} \mathbf{u}^H \mathbf{AA}^H \mathbf{w} \overset{(2)}{=} \frac{1}{\lambda} \mathbf{u}^H \lambda \mathbf{w} = \mathbf{u}^H \mathbf{w}$$

dove in (2) si è tenuto conto che $\mathbf{AA}^H\mathbf{w} = \lambda \mathbf{w}$ e in (1) non capisco perché l'H-trasposizione non riguarda la costante $1/\sigma$ mediante coniugio . □ **TODO**: fixme

Osservazione 364. La conseguenza rilevante della proposizione 6.4.2 consiste nel fatto che le due applicazioni φ_λ e ψ_λ scambiano basi ortonormali degli autospazi $E_{\mathbf{AA}^H}(\lambda)$ e $E_{\mathbf{A}^H\mathbf{A}}(\lambda)$. Questo perché:

- sono funzioni lineari tra spazi vettoriali e sono biettive, quindi vale proposizione 2.4.9
- il modulo dei vettori rimane lo stesso, quindi se si parte da una base ortonormale si avrà una base ortonormale (dato che le dimensioni coincidono)

Osservazione 365. Si osservi che le due precedenti proposizioni non dicono nulla sugli spazi nulli delle matrici coinvolte, che infatti non hanno nessuna relazione tra loro.

Definizione 6.4.1 (Decomposizione a valori singolari della matrice $\mathbf{A}$)**.** Data una arbitraria matrice complessa $m \times n$ di rango $\operatorname{rk} \mathbf{A} = k$ si chiama decomposizione in valori singolari di $\mathbf{A}$ una sua fattorizzazione del tipo

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$$

dove $\mathbf{U}$ e $\mathbf{V}$ sono matrici unitarie di ordini rispettivi m e n e $\boldsymbol{\Sigma}$ è una matrice $m \times n$ pseudo-diagonale (cioè con elementi non diagonali nulli) i cui elementi diagonali σ_i sono reali non negativi.

Osservazione 366. Qui chiamiamo elementi diagonali di una matrice rettangolare gli elementi di posto (i, i), estendendo la definizione del caso quadrato.
Per comodità si richiede usualmente che:

$$\sigma_1 \geq \ldots \geq \sigma_k > 0, \qquad \sigma_{k+1} = \ldots = \sigma_n = 0$$

Definizione 6.4.2 (Valori singolari)**.** Gli elementi positivi $\sigma_1, \ldots, \sigma_k$ sulla diagonale $\boldsymbol{\Sigma}$, che sono tanti quanto è il rango di $\boldsymbol{\Sigma}$ e quindi di $\mathbf{A}$, si chiamano *valori singolari* di $\mathbf{A}$.

Osservazione 367. In queste notazioni vale il seguente fatto.

Lemma 6.4.3. *I valori singolari di $\mathbf{A}$ sono univocamente individuati dalla matrice $\mathbf{A}$.*

Dimostrazione. La matrice $\mathbf{A}$ individua $\mathbf{A}\mathbf{A}^H$ e si ha che:

$$\mathbf{A}\mathbf{A}^H = \mathbf{U}\boldsymbol{\Sigma}\underbrace{\mathbf{V}^H\mathbf{V}}_{\mathbf{I}}\boldsymbol{\Sigma}^H\mathbf{U}^H = \mathbf{U}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\mathbf{U}^H$$

quindi $\mathbf{A}\mathbf{A}^H$ individua univocamente, a meno dell'ordine, gli elementi σ_i^2 (sulla diagonale di $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H$) che sono i suoi autovalori. Essendo $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H$ una matrice diagonale con elementi diagonali reali non negativi, resta univocamente determinata la matrice $\boldsymbol{\Sigma}$, per l'ipotesi di non negatività dei suoi elementi diagonali (ossia tengo le radici positive degli elementi sulla diagonale di $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H$). □

Lemma 6.4.4. *$\mathbf{A}$ ha lo stesso rango di $\mathbf{A}\mathbf{A}^H$*

Dimostrazione. Questo in quanto $\boldsymbol{\Sigma}$ ha lo stesso numero di colonne non nulle di $\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H$ □

Osservazione 368. Ora un risultato che ci serve per il prosieguo. Eventualmente metterlo da qualche altra parte

TODO: fixme

Proposizione 6.4.5. *Matrici simili hanno lo stesso rango*

Dimostrazione. Sia $\mathbf{A} = \mathbf{X}\mathbf{B}\mathbf{X}^{-1}$ e proviamo che $\operatorname{rk} \mathbf{A} = \operatorname{rk} \mathbf{B}$. Dobbiamo premettere 2 risultati:

1. proviamo innanzitutto che se $\mathbf{X}$ è invertibile $\mathrm{rk}(\mathbf{XB}) = \mathrm{rk}(\mathbf{B})$.
Dato un $\mathbf{u} \in \mathrm{N}(\mathbf{XB})$, si ha che $\mathbf{Bu} = \mathbf{0}$ e pre-moltiplicando per $\mathbf{X}$ si ha $\mathbf{XBu} = \mathbf{X0} = \mathbf{0}$ ossia $\mathbf{u} \in \mathrm{N}(\mathbf{XB})$, quindi $\mathrm{N}(\mathbf{B}) \subseteq \mathrm{N}(\mathbf{XB})$.
Allo stesso modo se $\mathbf{u} \in \mathrm{N}(\mathbf{XB})$ si ha che $\mathbf{XBu} = \mathbf{0}$, ma essendo $\mathbf{X}$ invertibile possiamo pre-moltiplicare per $\mathbf{X}^{-1}$ ossia $\mathbf{Bu} = \mathbf{X}^{-1}\mathbf{0} = \mathbf{0}$ quindi $\mathbf{u} \in \mathrm{N}(\mathbf{B})$ ossia $\mathrm{N}(\mathbf{XB}) \leq \mathrm{N}(\mathbf{B})$. Pertanto $\mathrm{N}(\mathbf{XB}) = \mathrm{N}(\mathbf{B})$ e dato che $\mathbf{A}$ e $\mathbf{B}$ hanno la stessa dimensione deve essere $\mathrm{rk}(\mathbf{AB}) = \mathrm{rk}(\mathbf{B})$

2. mostriamo ora che se $\mathbf{X}$ è invertibile si ha che $\mathrm{rk}(\mathbf{AX}) = \mathrm{rk}(\mathbf{A})$. Segue da

$$\mathrm{rk}(\mathbf{AX}) = \mathrm{rk}((\mathbf{AX})^T) = \mathrm{rk}(\mathbf{X}^T\mathbf{A}^T) \overset{(1)}{=} \mathrm{rk}(\mathbf{A}^T) = \mathrm{rk}(\mathbf{A})$$

con (1) giustificato da quanto visto al punto precedente

Proviamo ora che se $\mathbf{A} = \mathbf{XBX}^{-1}$, allora $\mathrm{rk}\,\mathbf{A} = \mathrm{rk}\,\mathbf{B}$. Post moltiplicando per $\mathbf{X}$ possiamo riscrivere la similitudine come $\mathbf{AX} = \mathbf{XB}$ passando ai ranghi si ha che $\mathrm{rk}(\mathbf{AX}) = \mathrm{rk}(\mathbf{XB})$ e per i punti 1 e 2 prima esposti si conclude che $\mathrm{rk}\,\mathbf{A} = \mathrm{rk}\,\mathbf{B}$. □

Osservazione 369. Naturalmente si pone il problema dell'esistenza di una decomposizione in valori singolari di una generica matrice $\mathbf{A}$. Ciò è assicurato dal seguente risultato.

Teorema 6.4.6. *Ogni matrice complessa $m \times n$ $\mathbf{A}$ ha una decomposizione in valori singolari.*

Dimostrazione. Sia $\mathbf{AA}^H = \mathbf{U\Lambda U}^H$ una diagonalizzazione unitaria di $\mathbf{AA}^H$, esistente per il teorema spettrale (essendo $\mathbf{AA}^H$ hermitiana e dunque normale). Possiamo supporre che

TODO: Rivedere questa dimostrazione

$$\mathbf{\Lambda} = \mathbf{Diag}(\lambda_1, \ldots, \lambda_k, 0, \ldots, 0)$$

con k uguale al rango di $\mathbf{AA}^H$ (dato che $\mathbf{AA}^H$ e $\mathbf{\Lambda}$ sono simili, per proposizione 6.4.5, e visto ce il rango di $\mathbf{\Delta}$è palesemente k). k coincide con il rango di $\mathbf{A}$: questo credo per 6.4.4, tuttavia ciò presuppone che $\mathbf{A}$ abbia una decomposizione in valori singolari che è quello che vogliamo dimostrare quindi dubbi su come si giustifichi questa affermazione .

TODO: fixme

I λ_i sono numeri reali positivi (usualmente si scelgono i λ_i in ordine non crescente) dato che $\mathbf{AA}^H$ è semi-definita positiva, per teorema 6.3.9 (5 $\implies$ 1).
Sia $\mathbf{U} = [\mathbf{u}_1 \ \ldots \ \mathbf{u}_m]$. Allora una base ortonormale dello spazio nullo $\mathrm{N}\left(\mathbf{AA}^H\right)$, coincidente con l'autospazio di $\mathbf{AA}^H$ relativo all'autovalore 0, è data da

$$[\mathbf{u}_{k+1} \ \ldots \ \mathbf{u}_m]$$

ossia servono k vettori per una base di $\mathrm{C}\left(\mathbf{AA}^H\right)$ e i restanti $m-k$ per lo spazio nullo poiché

$$m = \underbrace{\dim \mathrm{C}\left(\mathbf{AA}^H\right)}_{=\mathrm{rk}\,\mathbf{AA}^H = k} + \dim \mathrm{N}\left(AA^H\right)$$

La proposizione 6.4.2 e il teorema 6.1.4 (punto 3) assicurano che, se $\sigma_i^2 = \lambda_i$ per ogni $i \leq k$ ($\sigma_i = \sqrt{\lambda_i}$, i σ_i sono autovalori di $\mathbf{A}$), posto:

$$\varphi : E_{\mathbf{AA}^H}(\lambda) \to E_{\mathbf{A}^H\mathbf{A}}(\lambda), \qquad \varphi_\lambda(\mathbf{u}) = \mathbf{v}_i = \frac{1}{\sigma_i}\mathbf{A}^H\mathbf{u}_i$$

i vettori $\mathbf{v}_1, \dots, \mathbf{v}_k$ formano una base ortonormale del complemento ortogonale dello spazio nullo $\mathrm{N}\left(\mathbf{A}^H\mathbf{A}\right)$ di $\mathbf{A}^H\mathbf{A}$. Qualora risulti $k < n$, fissiamo una base ortonormale $\{\mathbf{v}_{k+1}, \dots, \mathbf{v}_n\}$ di $\mathrm{N}\left(\mathbf{A}^H\mathbf{A}\right)$ per cui $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ costituisce una base ortonormale di $\mathbb{C}^n$ formata da autovettori di $\mathbf{A}^H\mathbf{A}$ perché generata con φ che produce autovettori in $E_{\mathbf{A}\mathbf{A}^H}$. Di $\mathbf{A}^H\mathbf{A}$ esiste una diagonalizzazione unitaria per il teorema spettrale. Ma allora, posto $\mathbf{V} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]$ matrice unitaria, risulta $\mathbf{A}^H\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H$ dove la matrice diagonale degli autovalori $\mathbf{\Lambda}'$ ha come primi elementi diagonali diversi da zero ordinatamente gli elementi $\lambda_1, \dots, \lambda_k$. Infatti per $i \leq k$ risulta

$$\mathbf{A}^H\mathbf{A}\mathbf{v}_i = \mathbf{A}^H\mathbf{A}\frac{1}{\sigma_i}\mathbf{A}^H\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}^H\lambda_i\mathbf{u}_i = \lambda_i\mathbf{v}_i$$

quindi gli λ_i sono autovalori di $\mathbf{A}^H\mathbf{A}$ e finiscono sulla sua diagonalizzazione.
Sia $\mathbf{\Sigma}$ la matrice pseudo-diagonale $m \times n$ i cui primi k elementi diagonali sono ordinatamente $\sigma_1, \dots, \sigma_k$ e con gli altri elementi diagonali nulli. Proviamo che risulta $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ (occhio che qui è cambiato il $\mathbf{V}$ finale). Ciò è equivalente a provare che $\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}$ ($\mathbf{V}$ è unitaria, ossia $\mathbf{V}^H = \mathbf{V}^{-1}$, ed abbiamo posto moltiplicato per $\mathbf{V}$), ovvero che considerando le colonne separatamente per le prime k colonne si abbia

$$\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i \quad \text{per } i \leq k$$

mentre per le rimanenti

$$\mathbf{A}\mathbf{v}_i = \mathbf{0} \quad \text{per } i > k$$

Verifichiamo le uguaglianze per $i \leq k$:

$$\mathbf{A}\mathbf{v}_i = \mathbf{A}\frac{1}{\sigma_i}\mathbf{A}^H\mathbf{u}_i \overset{(1)}{=} \frac{1}{\sigma_i}\lambda_i\mathbf{u}_i = \sigma_i\mathbf{u}_i$$

come si voleva dimostrare (il passaggio (1) dato che $\mathbf{A}\mathbf{A}^H\mathbf{u}_i = \lambda_i\mathbf{u}_i$, con λ_i autovalori e $\mathbf{u}_i$ autovettori di $\mathbf{A}\mathbf{A}^H$. Infatti se $\mathbf{A}\mathbf{A}^H = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H$, con $\mathbf{U}$ unitaria ossia $\mathbf{U}^H = \mathbf{U}^{-1}$, post moltiplicando per $\mathbf{A}$ si ottiene $\mathbf{A}\mathbf{A}^H\mathbf{U} = \mathbf{U}\mathbf{\Lambda}$, e vedendo colonna per colonna deve essere $\mathbf{A}\mathbf{A}^H\mathbf{u}_i = \lambda_i\mathbf{u}_i$, quindi $\mathbf{u}_i$ sono autovettori).

La verifica delle uguaglianze per $i > k$ è banale, perché i vettori $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ sono una base che abbiamo precedentemente fissato di $\mathrm{N}\left(\mathbf{A}^H\mathbf{A}\right)$ e $\mathrm{N}\left(\mathbf{A}^H\mathbf{A}\right)$ coincide con $\mathrm{N}\left(\mathbf{A}\right)$ (si veda Esercizio 6.22) □

TODO: fixme

Osservazione 370. La strategia usata nel teorema 6.4.6 per trovare una decomposizione in valori singolari $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ della matrice $\mathbf{A}$ è consistita nel diagonalizzare unitariamente la matrice $\mathbf{A}\mathbf{A}^H$, determinando in tal modo la matrice unitaria $\mathbf{U}$, gli autovalori reali positivi di $\mathbf{A}\mathbf{A}^H$ e quindi i valori singolari di $\mathbf{A}$, che sono le loro radici quadrate positive. Si è così determinata anche la matrice $\mathbf{\Sigma}$. La parte centrale della dimostrazione è consistita nel definire opportunamente la matrice unitaria $\mathbf{V}$ a partire dalla matrice $\mathbf{U}$.
Si poteva però partire da una diagonalizzazione unitaria della matrice $\mathbf{A}^H\mathbf{A}$, determinando prima $\mathbf{V}$ e poi definendo di conseguenza $\mathbf{U}$.
Di fatto, usualmente si sceglie la strada che comporta minor fatica, quindi si sceglie la prima via se $m \leq n$, la seconda in caso contrario.

Osservazione 371. Se la matrice $\mathbf{A}$ è reale, le matrici $\mathbf{A}\mathbf{A}^H$ e $\mathbf{A}^H\mathbf{A}$ sono simmetriche reali e possono quindi essere ortogonalmente diagonalizzate in ambito reale. Ne consegue che $\mathbf{A}$ ha una decomposizione in valori singolari $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ con entrambi le matrici $\mathbf{U}$ e $\mathbf{V}$ ortogonali reali.

Esempio 6.4.1. Si consideri la matrice incontrata nell'esempio 1.6.4:

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}$$

Risulta

$$\mathbf{A}\mathbf{A}^H = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 2 & 1 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 56 & 28 \\ 28 & 14 \end{bmatrix}$$

Un facile calcolo mostra che gli autovalori di $\mathbf{A}\mathbf{A}^H$ sono $\lambda_1 = 70$ e $\lambda_2 = 0$: infatti, procediamo determinando le radici del polinomio caratteristico:

$$\begin{aligned} \det(\mathbf{A}\mathbf{A}^H - \mathbf{I}X) = \begin{bmatrix} 56 - X & 28 \\ 28 & 14 - X \end{bmatrix} &= (56 - X)(14 - X) - 28^2 \\ &= 784 - 56X - 14X + X^2 - 784 = X^2 - 70X \end{aligned}$$

Si ha che $X^2 - 70X = 0 \iff X = 0 \vee X = 70$ pertanto $\lambda_1 = 70, \lambda_1 = 0$ (poniamoli già in ordine decrescente).
Ora troviamo gli autovettori di lunghezza unitaria per λ_1, λ_2; per il primo partiamo dal determinare

$$\mathrm{N}\left(\mathbf{A}\mathbf{A}^H - 0 \cdot \mathbf{I}\right) = \mathrm{N}\left(\mathbf{A}\mathbf{A}^H\right) = \mathrm{N}\left(\begin{bmatrix} 56 & 28 \\ 28 & 14 \end{bmatrix}\right)$$

ossia

$$\begin{bmatrix} 56 & 28 \\ 28 & 14 \end{bmatrix} \begin{bmatrix} 56 & 28 \\ 56 & 28 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix} \to \begin{cases} 2x_1 + h = 0 \\ x_2 = h \end{cases} \quad \begin{cases} x_1 = -\frac{h}{2} \\ x_2 = h \end{cases} \quad \to h \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}$$

Passando al secondo si ha

$$\mathrm{N}\left(\mathbf{A}\mathbf{A}^H - 70 \cdot \mathbf{I}\right) = \mathrm{N}\left(\begin{bmatrix} 56 - 70 & 28 \\ 28 & 14 - 70 \end{bmatrix}\right) = \mathrm{N}\left(\begin{bmatrix} -14 & 28 \\ 28 & -56 \end{bmatrix}\right)$$

ossia

$$\begin{bmatrix} -14 & 28 \\ 28 & -56 \end{bmatrix} \begin{bmatrix} 28 & -56 \\ 28 & -56 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 0 \end{bmatrix} \to \begin{cases} x_2 = h \\ x_1 - 2h = 0 \end{cases} \quad \begin{cases} x_2 = h \\ x_1 = 2h \end{cases} \quad \to h \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Ora normalizziamo gli autovettori trovati dividendoli per i rispettivi moduli:

$$\begin{aligned} &\left\| \begin{bmatrix} -1/2 \\ 1 \end{bmatrix} \right\| = \sqrt{\frac{1}{4} + 1} = \frac{\sqrt{5}}{2} && \to \frac{2}{\sqrt{5}} \begin{bmatrix} -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \\ &\left\| \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\| = \sqrt{5} && \to \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \end{aligned}$$

La diagonalizzazione unitaria (una delle possibili) la facciamo tenendo gli autovalori in ordine decrescente sulla diagonale delle matrice diagonale e ordinatamente in accordo le colonne di $\mathbf{U}$ con i corrispondenti autovettori

$$\mathbf{A}\mathbf{A}^H = \underbrace{\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} 70 & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}}_{\mathbf{U}^{-1}=\mathbf{U}^H=\mathbf{U}^T}$$

Per verifica che effettivamente la decomposizione sia corretta effettuiamo il calcolo raccogliendo il doppio $1/\sqrt{5}$; si ha:

$$\frac{1}{5}\begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 70 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 56 & 28 \\ 28 & 14 \end{bmatrix} = \mathbf{A}\mathbf{A}^H$$

La matrice $\mathbf{U}$ è quindi individuata e la matrice $\mathbf{\Sigma}$ si ricava dal valore singolare $\sigma_1 = \lambda_1^{1/2} = \sqrt{70}$:

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{70} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Ora dobbiamo determinare $\mathbf{V}$ per completare la SVD. In accordo con la dimostrazione del teorema 6.4.6, la prima colonna della matrice unitaria $\mathbf{V}$ che diagonalizza $\mathbf{A}^H\mathbf{A}$ è data da:

$$\mathbf{v}_1 = \frac{1}{\sigma_1}\mathbf{A}^H\mathbf{u}_1 = \frac{1}{\sqrt{70}}\begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 2 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \frac{1}{\sqrt{70}\sqrt{5}}\begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 2 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{70}\sqrt{5}}\begin{bmatrix} 10 \\ 0 \\ 5 \\ 15 \end{bmatrix} = \frac{5}{\sqrt{70\cdot 7}}\begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix} = \frac{5}{5\sqrt{14}}\begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix} = \sqrt{\frac{1}{14}}\begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix}$$

A questo punto ci fermiamo poiché gli altri autovalori sono nulli. Per le altre colonne di $\mathbf{V}$, dato che risulta $k = 1 < n = 4$ troviamo una base ortonormale di $\mathrm{N}\left(\mathbf{A}^H\mathbf{A}\right)$ che è il complemento ortogonale di $\langle \mathbf{v}_1 \rangle$. Si ha che

$$\mathbf{A}^H\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 2 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 20 & 0 & 10 & 30 \\ 0 & 0 & 0 & 0 \\ 10 & 0 & 5 & 15 \\ 30 & 0 & 15 & 45 \end{bmatrix}$$

Procediamo per trovare $\mathrm{N}\left(\mathbf{A}^H\mathbf{A}\right)$:

$$\begin{bmatrix} 20 & 0 & 10 & 30 \\ 0 & 0 & 0 & 0 \\ 10 & 0 & 5 & 15 \\ 30 & 0 & 15 & 45 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 3 \\ 2 & 0 & 1 & 3 \end{bmatrix}\begin{bmatrix} 2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \to \begin{cases} x_4 = h_4 \\ x_3 = h_3 \\ x_2 = h_2 \\ 2x_1 + h_3 + 3h_4 = 0 \end{cases}$$

$$\begin{cases} x_1 = -\frac{h_3+3h_4}{2} \\ x_2 = h_2 \\ x_3 = h_3 \\ x_4 = h_4 \end{cases} \to \begin{bmatrix} -\frac{h_3+3h_4}{2} \\ h_2 \\ h_3 \\ h_4 \end{bmatrix}$$

Ora vedendo i vari casi

$$h_2 = 1, h_3 = 0, h_4 = 0 \rightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad h_2 = 0, h_3 = 1, h_4 = 0 \rightarrow \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad h_2 = 0, h_3 = 0, h_4 = 1 \rightarrow \begin{bmatrix} -3/2 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Procediamo ad ortogonalizzare la base con l'algoritmo di Gram Schmidt:

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} 1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2 = \begin{bmatrix} -3/2 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} -3/2 \\ 0 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -3/2 \\ 0 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix}^T \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -3/2 \\ 0 \\ 0 \\ 1 \end{bmatrix} - 0 - \frac{3/4}{1/4 + 1} \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 0 \\ 0 \\ 1 \end{bmatrix} - \frac{3}{4} \cdot \frac{4}{5} \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 0 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 3/10 \\ 0 \\ -3/5 \\ 0 \end{bmatrix} = \begin{bmatrix} -6 \\ 0 \\ -3 \\ 5 \end{bmatrix}$$

Quindi la base ortogonale è

$$\left\langle \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -1/2 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} -6 \\ 0 \\ -3 \\ 5 \end{bmatrix} \right\rangle = \left\langle \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \end{bmatrix} \begin{bmatrix} -6 \\ 0 \\ -3 \\ 5 \end{bmatrix} \right\rangle$$

Normalizziamo ora dividendo per il modulo

$$\left\| \begin{bmatrix} -1 & 0 & 2 & 0 \end{bmatrix}^T \right\| = \sqrt{5}$$

$$\left\| \begin{bmatrix} -6 & 0 & -3 & 5 \end{bmatrix}^T \right\| = \sqrt{36 + 9 + 25} = \sqrt{72} = 6\sqrt{2}$$

La base ortonormale ricercata è dunque

$$\left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{5} \\ 0 \\ 2/\sqrt{5} \\ 0 \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ -1/2\sqrt{2} \\ 5/6\sqrt{2} \end{bmatrix} \right\}$$

e va a formare $\mathbf{V}$ assieme a $\mathbf{v}_1$ individuato in precedenza

$$\mathbf{V} = \begin{bmatrix} 2\sqrt{1/14} & 0 & -1/\sqrt{5} & -1/\sqrt{2} \\ 0 & 1 & 0 & 0 \\ \sqrt{1/14} & 0 & 2/\sqrt{5} & -1/(2\sqrt{2}) \\ 3\sqrt{1/14} & 0 & 0 & 5/(6\sqrt{2}) \end{bmatrix}$$

Per verifica effettivamente risulta che:

$$\begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} \sqrt{70} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2\sqrt{1/14} & 0 & \sqrt{1/14} & 3\sqrt{1/14} \\ 0 & 1 & 0 & 0 \\ -1/\sqrt{5} & 0 & 2/\sqrt{5} & 0 \\ -1/\sqrt{2} & 0 & -1/2\sqrt{2} & 5/6\sqrt{2} \end{bmatrix} = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix} = \mathbf{A}$$

Questa è una delle decomposizioni possibili (ad esempio il libro giunge ad un'altra).

6.4.1 Applicazioni

6.4.1.1 Forma polare di una matrice

Osservazione 372. Ogni numero complesso z si può scrivere in forma trigonometrica come:

$$z = \rho(\cos\alpha + i\sin\alpha)$$

dove $\rho \in \mathbb{R}, \rho \geq 0$ coincide con $|z|$, mentre α è l'argomento; ρ e α sono le coordinate polari (il polo è l'origine), e per questo la forma trigonometrica è detta anche forma polare.

Osservazione 373. Per le matrici complesse $n \times n$ e più generalmente $m \times n$ con $m \leq n$ esiste una decomposizione analoga.

Definizione 6.4.3 (Forma polare di matrice complessa)**.** Si chiama *forma polare* di una matrice complessa $m \times n$ $\mathbf{A}$, con $m \leq n$, una fattorizzazione del tipo

$$\mathbf{A} = \mathbf{PQ}$$

dove $\mathbf{P}$ è hermitiana semi-definita positiva e $\mathbf{Q}$ è $m \times n$ a righe ortonormali (cioè $\mathbf{QQ}^H = \mathbf{I}_m$).

Teorema 6.4.7. *Ogni matrice complessa $m \times n$ $\mathbf{A}$ con $m \leq n$, ha una decomposizione polare $\mathbf{A} = \mathbf{PQ}$ con $\mathbf{P}$ univocamente individuata, e $\mathbf{Q}$ univocamente individuata se $\mathbf{A}$ ha rango m.*

Dimostrazione. Sia $A = \mathbf{U\Sigma V}^H$ una decomposizione in valori singolari di $\mathbf{A}$. Sia $\mathbf{\Sigma}_0$ la matrice diagonale $m \times m$ avente la stessa diagonale di $\mathbf{\Sigma}$ (ossia è $\mathbf{\Sigma}$ tagliata sino ad avere un quadrato).
Posto $\mathbf{J} = [\mathbf{I}_m\ \mathbf{O}]$, dove il blocco nullo di $\mathbf{J}$ ha dimensioni $m \times (n-m)$ (e dunque $\mathbf{J}$ è $m \times n$), è immediato verificare che $\mathbf{\Sigma} = \mathbf{\Sigma}_0\mathbf{J}$ (si tratta di $\mathbf{\Sigma_0}$ (per $\mathbf{I}$) cui si aggiungono delle colonne di 0 alla fine) e che $\mathbf{JJ}^H = \mathbf{I}_m$. Si ha di conseguenza:

$$\mathbf{A} = \mathbf{U\Sigma V}^H = \mathbf{U\Sigma_0 JV}^H = \mathbf{U\Sigma_0}\underbrace{\mathbf{I}}_{=\mathbf{U}^H\mathbf{U}}\mathbf{JV}^H = \underbrace{\mathbf{U\Sigma_0U}^H}_{\mathbf{P}}\underbrace{\mathbf{UJV}^H}_{\mathbf{Q}} = \mathbf{PQ}$$

Ponendo come si è fatto $\mathbf{P} = \mathbf{U}\mathbf{\Sigma}_0\mathbf{U}^H$ e $\mathbf{Q} = \mathbf{U}\mathbf{J}\mathbf{V}^H$, si ha che $\mathbf{P}$ è semi-definita positiva per il teorema 6.3.9 (è unitariamente diagonalizzata con una matrice di autovalori positivi o nulli); inoltre si ha che $\mathbf{Q}$ è a righe ortonormali in quanto

$$\mathbf{Q}\mathbf{Q}^H = \mathbf{U}\mathbf{J}\underbrace{\mathbf{V}\mathbf{V}^H}_{=\mathbf{I}}\mathbf{J}^H\mathbf{U}^H = \mathbf{U}\underbrace{\mathbf{J}\mathbf{J}^H}_{\mathbf{I}}\mathbf{U}^H = \mathbf{U}\mathbf{U}^H = \mathbf{I}_m$$

ossia nei passaggi si è usato il fatto che $\mathbf{V}$, $\mathbf{J}$, $\mathbf{U}$ sono unitarie. Quanto all'unicità di $\mathbf{P}$, essa segue dalla proposizione 6.3.10 ($\mathbf{P}$ è la radice quadrata di $\mathbf{A}\mathbf{A}^H$) in quanto ora mostriamo che $\mathbf{P}^2 = \mathbf{A}\mathbf{A}^H$; si ha che

$$\mathbf{P}^2 = \mathbf{U}\mathbf{\Sigma}_0\mathbf{U}^H\mathbf{U}\mathbf{\Sigma}_0\mathbf{U}^H = \mathbf{U}\mathbf{\Sigma}_0^2\mathbf{U}^H$$

mentre

$$\mathbf{A}\mathbf{A}^H = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{V}\mathbf{\Sigma}^H\mathbf{U} = \mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^H$$

e l'uguaglianza $\mathbf{P}^2 = \mathbf{A}\mathbf{A}^H$ è garantita dal fatto che $\mathbf{\Sigma}^2 = \mathbf{\Sigma}_0^2$.
Infine, se $\mathbf{A}$ ha rango m, allora anche $\mathbf{P}$ ha rango m, come tentiamo di andare a mostrare : si è visto che $\mathbf{A}\mathbf{A}^H$ ha lo stesso rango di $\mathbf{A}$ per quanto visto nella dimostrazione di teorema 6.4.4. Mostriamo che $\mathbf{P}$ ha lo stesso rango di $\mathbf{A}\mathbf{A}^H$: si ha che **TODO**: check

$$\mathbf{A}\mathbf{A}^H = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H\mathbf{V}\mathbf{\Sigma}^H\mathbf{U}^H = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^H\mathbf{U}^H$$

ma Σ_0 (che diagonalizza $\mathbf{P}$) ha lo stesso numero di colonne non nulle di $\mathbf{\Sigma}\mathbf{\Sigma}^H$ (che diagonalizza $\mathbf{A}\mathbf{A}^H$).
Avendo rango m, pertanto, $\mathbf{P}$ è invertibile (prodotto di quadrate ordine m, quindi ha ordine m, coincidente col suo rango) e da $\mathbf{A} = \mathbf{P}\mathbf{Q}$ pre-moltiplicando per $\mathbf{P}^{-1}$ risulta $\mathbf{Q} = \mathbf{P}^{-1}\mathbf{A}$ ($\mathbf{Q}$ univocamente individuata). □

Osservazione 374. Si osservi che nella costruzione eseguita nel teorema 6.4.7 la matrice $\mathbf{P}$ ha il medesimo rango di $\mathbf{A}$; inoltre, se $\mathbf{A}$ è reale, la decomposizione in valori singolari $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ può essere scelta, per la osservazione 371 con entrambe le matrici $\mathbf{U}$ e $\mathbf{V}$ ortogonali reali. Ne consegue che $\mathbf{A}$ ha una decomposizione polare $\mathbf{A} = \mathbf{P}\mathbf{Q}$ con $\mathbf{P}$ e $\mathbf{Q}$ entrambe matrici reali.

Esempio 6.4.2. Per avere una decomposizione polare a partire dalla decomposizione in valori singolari ottenuta in esercizio 6.4.1, ossia $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ con

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}, \quad \mathbf{\Sigma} = \begin{bmatrix} \sqrt{70} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{V}^H = \begin{bmatrix} 2\sqrt{1/14} & 0 & \sqrt{1/14} & 3\sqrt{1/14} \\ 0 & 1 & 0 & 0 \\ -1/\sqrt{5} & 0 & 2/\sqrt{5} & 0 \\ -1/\sqrt{2} & 0 & -1/2\sqrt{2} & 5/6\sqrt{2} \end{bmatrix}$$

Per ottenere la forma polare $\mathbf{A} = \mathbf{P}\mathbf{Q}$, per esteso

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H = (\mathbf{U}\mathbf{\Sigma}_0\mathbf{U}^H)(\mathbf{U}\mathbf{J}\mathbf{V}^H) = \mathbf{P}\mathbf{Q}$$

seguendo la dimostrazione di teorema 6.4.7 abbiamo

$$\mathbf{\Sigma}_0 = \begin{bmatrix} \sqrt{70} & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{J} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

con $\mathbf{J}$ di $m \times n$ come $\mathbf{A}$ e costruita affiancando a $\mathbf{I}$ degli 0 sino a trovare la forma/dimensione desiderata.
Si ha che

$$\mathbf{P} = \mathbf{U}\boldsymbol{\Sigma}_0\mathbf{U}^H = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} \sqrt{70} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$
$$= \frac{\sqrt{70}}{5} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \frac{\sqrt{70}}{5} \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \frac{\sqrt{70}}{5} \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

mentre

$$\mathbf{Q} = \mathbf{U}\mathbf{J}\mathbf{V}^H = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{V}^H$$
$$= \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} & 0 & 0 \\ 1/\sqrt{5} & 2/\sqrt{5} & 0 & 0 \end{bmatrix} \begin{bmatrix} 2\sqrt{1/14} & 0 & \sqrt{1/14} & 3\sqrt{1/14} \\ 0 & 1 & 0 & 0 \\ -1/\sqrt{5} & 0 & 2/\sqrt{5} & 0 \\ -1/\sqrt{2} & 0 & -1/2\sqrt{2} & 5/6\sqrt{2} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{4\sqrt{1/14}}{\sqrt{5}} & -1/\sqrt{5} & \frac{2\sqrt{1/14}}{\sqrt{5}} & \frac{6\sqrt{1/14}}{\sqrt{5}} \\ \frac{2\sqrt{1/14}}{\sqrt{5}} & \frac{2}{\sqrt{5}} & \frac{\sqrt{1/14}}{\sqrt{5}} & \frac{3\sqrt{1/14}}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} 4 \cdot \frac{1}{\sqrt{14}} \cdot \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & 2 \cdot \frac{1}{\sqrt{14}} \cdot \frac{1}{\sqrt{5}} & 6 \cdot \frac{1}{\sqrt{14}} \cdot \frac{1}{\sqrt{5}} \\ 2 \cdot \frac{1}{\sqrt{14}} \cdot \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{14}} \cdot \frac{1}{\sqrt{5}} & 3 \cdot \frac{1}{\sqrt{14}} \cdot \frac{1}{\sqrt{5}} \end{bmatrix}$$
$$= \frac{1}{\sqrt{14}\sqrt{5}} \begin{bmatrix} 4 & -\sqrt{14} & 2 & 6 \\ 2 & 2\sqrt{14} & 1 & 3 \end{bmatrix}$$

6.4.1.2 Decomposizione in valori singolari e pseudo inversa

Osservazione 375. L'esistenza della matrice pseudo-inversa di una generica matrice complessa $m \times n$ $\mathbf{A}$ di rango k è stata ricavata nel paragrafo 1.6 tramite una decomposizione a rango pieno di $\mathbf{A}$ (ossia una fattorizzazione del tipo $\mathbf{A} = \mathbf{BC}$, con $\mathbf{B}$ $m \times n$ e $\mathbf{C}$ $k \times n$).

Osservazione 376. Vediamo qui una prova dell'esistenza della matrice pseudo-inversa alternativa a quella fornita dal teorema 1.6.16, e nello specifico come l'esistenza di decomposizioni in valori singolari di $\mathbf{A}$ produce una dimostrazione alternativa dell'esistenza della matrice pseudo-inversa.

Proposizione 6.4.8. *Data una decomposizione in valori singolari* $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$ *della matrice* $m \times n$ $\mathbf{A}$ *con valori singolari* $\sigma_1, \ldots, \sigma_k$ *si ha che* $\mathbf{A}^+ = \mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^H$, *dove* $\boldsymbol{\Sigma}^+$ *è la pseudo-diagonale* $n \times m$ *con sulla diagonale ordinatamente gli elementi* $\sigma_1^{-1}, \ldots, \sigma_k^{-1}, 0, \ldots, 0$

TODO: fixme

Dimostrazione. Lasciamo al lettore la facile verifica che la matrice $\mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^H$ sopra definita soddisfa alle quattro condizioni richieste per essere la pseudo-inversa $\mathbf{A}^+$. È poi altrettanto facile verificare che la matrice pseudo-diagonale con elementi diagonali $\sigma_1^{-1}, \ldots, \sigma_k^{-1}, 0, \ldots, 0$ è la pseudo-inversa di $\boldsymbol{\Sigma}$. □

Osservazione 377. Un terzo modo per ottenere la pseudo-inversa della matrice $\mathbf{A}$ è quello di fare uso di una decomposizione polare $\mathbf{A} = \mathbf{PQ}$.

Proposizione 6.4.9. *Si ha che* $\mathbf{A}^+ = \mathbf{Q}^H\mathbf{P}^H$, *dove* $\mathbf{P}^+$ *si ricava da una diagonalizzazione unitaria* $\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$ *di* $\mathbf{P}$ *con* $\boldsymbol{\Lambda} = \mathbf{Diag}(\lambda_1, \ldots, \lambda_k, 0, \ldots, 0)$ *e i numeri* λ_i *reali positivi:*

$$\mathbf{P}^+ = \mathbf{U}\boldsymbol{\Lambda}^+\mathbf{U}^H, \qquad \boldsymbol{\Lambda}^+ = \mathbf{Diag}\big(\lambda_1^{-1}, \ldots, \lambda_k^{-1}, 0, \ldots, 0\big)$$

Dimostrazione. Esercizio 6.27 □ **TODO**: fixme

Esempio 6.4.3. Si consideri la matrice

$$\mathbf{A} = \begin{bmatrix} 4 & 0 & 2 & 6 \\ 2 & 0 & 1 & 3 \end{bmatrix}$$

si è visto nell'esercizio 1.6.4 che la sua pseudo-inversa è

$$\mathbf{A}^+ = \frac{1}{70}\mathbf{A}^T$$

Calcoliamo $\mathbf{A}^+$ tramite la proposizione 6.4.8. Una decomposizione in valori singolari $\mathbf{A} = \mathbf{U\Sigma V}^H$ è stata calcolata nell'esempio 6.4.1; dalla proposizione 6.4.8 sappiamo che $\mathbf{A}^+ = \mathbf{V\Sigma}^+\mathbf{U}^H$, quindi risulta:

$$\begin{aligned}\mathbf{A}^+ &= \begin{bmatrix} 2/\sqrt{14} & 0 & 0 & 5/\sqrt{35} \\ 0 & 1 & 0 & 0 \\ 1/\sqrt{14} & 0 & 3/\sqrt{10} & -1/\sqrt{35} \\ 3/\sqrt{14} & 0 & -1/\sqrt{10} & -3/\sqrt{35} \end{bmatrix} \begin{bmatrix} 1/\sqrt{70} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \\ &= \begin{bmatrix} 2/\sqrt{14 \cdot 70} & 0 \\ 0 & 0 \\ 1/\sqrt{14 \cdot 70} & 0 \\ 3/\sqrt{14 \cdot 70} & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4/70 & 2/70 \\ 0 & 0 \\ 2/70 & 1/70 \\ 6/70 & 3/70 \end{bmatrix}\end{aligned}$$

in accordo con i risultati dell'esempio esercizio 1.6.4.

6.5 Esercizi riassuntivi: autovalori, autovettori e matrici normali

6.5.1 Fatti

Esempio 6.5.1. Si consideri la matrice dipendente dal numero complesso $z = a + ib$

$$\mathbf{A}(z) = \begin{bmatrix} 0 & z \\ \overline{z} & i \end{bmatrix} = \begin{bmatrix} 0 & a+ib \\ a-ib & i \end{bmatrix}$$

- Si dica per quali valori di z la matrice $\mathbf{A}(z)$ è unitariamente diagonalizzabile
- Si dica per quali valori di z la matrice $\mathbf{A}(z)$ è diagonalizzabile
- Si trovi una diagonalizzazione unitaria della matrice $\mathbf{A}(i/2)$

Primo punto: per il teorema spettrale la matrice è unitariamente diagonalizzaile se e solo se è normale, ossia se commuta con la sua h-trasposta, ossia $\mathbf{AA}^H = \mathbf{A}^H\mathbf{A}$; si ha che

$$\begin{aligned}\mathbf{AA}^H &= \begin{bmatrix} 0 & z \\ \overline{z} & i \end{bmatrix} \begin{bmatrix} 0 & z \\ \overline{z} & -i \end{bmatrix} = \begin{bmatrix} z\overline{z} & -iz \\ i\overline{z} & z\overline{z}+1 \end{bmatrix} \\ \mathbf{A}^H\mathbf{A} &= \begin{bmatrix} 0 & z \\ \overline{z} & -i \end{bmatrix} \begin{bmatrix} 0 & z \\ \overline{z} & i \end{bmatrix} = \begin{bmatrix} z\overline{z} & zi \\ -i\overline{z} & z\overline{z}+1 \end{bmatrix}\end{aligned}$$

Si ha che $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ se

$$\begin{cases} zi = -iz \\ -i\overline{z} = i\overline{z} \end{cases}$$

ossia $z = 0$. Se dunque $z = 0$ la matrice è unitariamente diagonalizzabile.

Venendo al secondo punto la matrice è diagonalizzabile se ha tutti autovalori distinti o alternativamente le molteplicità geometriche dei vari autovalori coincidono con quelle algebriche. Calcoliamo gli autovettori per $\mathbf{A}(z)$

$$p_{A(z)}(X) = \det(\mathbf{A} - \mathbf{I}X) = (0 - X)(i - X) - |z|^2 = X^2 - iX - |z|^2$$

Si ha dunque

$$\lambda_{1,2} = \frac{i \pm \sqrt{i^2 + 4|z|^2}}{2}$$

Si hanno radici differenti se $i^2+4|z|^2 \neq 0$ (ricordandoci che siamo in $\mathbb{C}$) ossia per i complessi z tali che $|z| \neq \frac{1}{2}$ e in tal caso la matrice è certamente diagonalizzabile. Se invece $|z| = \frac{1}{2}$ si ha un unico autovalore $\lambda = \frac{i}{2}$ con molteplicità algebrica 2; se anche la molteplicità geometrica di λ è 2 allora la matrice è diagonalizzabile anche in questo caso, altrimenti per $|z| = \frac{1}{2}$ non lo è. Tale molteplicità geometrica coincide con $\dim(\mathrm{N}\left(\mathbf{A} - \frac{i}{2}\mathbf{I}\right))$; ricordando che per il teorema nullità + rango si ha che

$$2 = \dim \mathrm{C}\left(\mathbf{A} - \frac{i}{2}\mathbf{I}\right) + \dim \mathrm{N}\left(\mathbf{A} - \frac{i}{2}\mathbf{I}\right)$$

possiamo calcolare la nullità come differenza di $2 - \mathrm{rk}(\mathbf{A} - \frac{i}{2}\mathbf{I})$. Procediamo dunque a ridurre $\mathbf{A} - \frac{i}{2}\mathbf{I}$ (ricordando che coniugare una intera riga non è una operazione ammessa):

$$\begin{bmatrix} -\frac{1}{2} & \overline{z} \\ \overline{z} & \frac{i}{2} \end{bmatrix} \begin{bmatrix} -i & 2z \\ 2\overline{z} & i \end{bmatrix} \begin{bmatrix} 1 & 2zi \\ 2\overline{z} & i \end{bmatrix} \begin{bmatrix} 2\overline{z} & 4z\overline{z}i \\ 2\overline{z} & i \end{bmatrix} \begin{bmatrix} 2\overline{z} & 4z\overline{z}i \\ 0 & i - 4z\overline{z}i \end{bmatrix}$$

Pertanto il sistema ha due colonne dominanti, $\dim \mathrm{C}\left(\mathbf{A} - \frac{i}{2}\mathbf{I}\right) = 2$ e la molteplicità geometrica $\dim \mathrm{N}\left(\mathbf{A} - \frac{i}{2}\mathbf{I}\right) = 2 - 2 = 0 \neq 2$ (molteplicità algebrica), pertanto la matrice non è diagonalizzabile.

Vedendo al terzo unto per la triangolarizzazione unitaria di

$$\mathbf{A}(\frac{i}{2}) = \begin{bmatrix} 0 & \frac{i}{2} \\ -\frac{i}{2} & i \end{bmatrix}$$

Iniziamo trovando gli autovalori di $\mathbf{A}(i/2)$; dato che $\left|\frac{i}{2}\right| = \frac{1}{2}$ sostituendo nell'espressione del polinomio caratteristico trovata in precedenza si ha

$$p_{\mathbf{A}}(X) = X^2 - iX - \frac{1}{4}$$

Pertanto

$$\lambda_{1,2} = \frac{+i \pm \sqrt{i^2 + 1}}{2} = \frac{i}{2}$$

Si ha un unico autovalore. Troviamo un autovettore $\mathbf{v}_i \in E_{\mathbf{A}}(i/2)$ ossia sfruttando che $v_i \in \mathrm{N}\left(A - \frac{i}{2}\mathbf{I}\right)$. A questo punto abbiamo una strada sicura è quella di trovare una base di tale spazio nullo e generare un vettore a piacere

$$\begin{bmatrix} -i/2 & i/2 \\ -i/2 & i/2 \end{bmatrix} \begin{bmatrix} -i/2 & i/2 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{cases} x_2 = h \\ -(i/2)x_1 + (i/2)h = 0 \end{cases} \qquad \begin{cases} x_2 = h \\ x_1 = h \end{cases} \quad \rightarrow \left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle$$

Scegliamo proprio $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ come vettore da usare nella triangolarizzazione; dobbiamo ora completare ad una base ortogonale (espandere a base ed applicare al secondo vettore l'algoritmo di Gram-Schmidt) e poi normalizzare. Partiamo dall'espandere a base l'autovettore

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Nell'ultima matrice le prime due colonne sono le dominanti quindi aggiungiamo $\mathbf{e}_1$ alla base. Procediamo ora a ortogonalizzazione mediante l'algoritmo di Gram-Schmidt

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}^T = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \end{bmatrix}$$

questo secondo vettore moltiplichiamolo per 2, tenendo $\begin{bmatrix} 1 & -1 \end{bmatrix}$ per comodità. Procediamo dunque a normalizzare dividendo per le norme (in entrambi i casi $\sqrt{2}$) e arriviamo finalmente a

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Ora, dato che $A = UTU^{-1}$ si ha che $T = U^{-1}AU$; ci mancano U^{-1} e T per completare la triangolarizzazione. La prima, se non abbiamo commesso errori di calcolo in $\mathbf{U}$ si ottiene prendendo l'H-trasposta di $\mathbf{U}$, dato che quest'ultima è unitaria; nel dubbio si può verificare che $\mathbf{U}\mathbf{U}^H = \mathbf{I}$ o alternativamente i questi casi 2×2 calcolare l'inversa e verificare che torni $\mathbf{U}^{-1} = \mathbf{U}^H$. Seguendo la seconda strada si ha

$$\mathbf{U}^{-1} = \frac{1}{\det \mathbf{U}} \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{-\frac{1}{2} - \frac{1}{2}} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & -1 \\ -1 & 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \mathbf{U}^H (= \mathbf{U})$$

Si ha ora (facendo i calcoli per esteso) che

$$\mathbf{T} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & i/2 \\ -i/2 & i \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} i/2 & -i \\ 0 & i/2 \end{bmatrix}$$

ci torna che sulla diagonale ci siano gli autovalori individuati in precedenza; inoltre verificando (con maxima) effettivamente torna che $\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{U}^{-1}$.

Esempio 6.5.2. Data la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 8 \end{bmatrix}$$

Si trovi:

- il polinomio caratteristico $p_{\mathbf{A}}(X)$ di $\mathbf{A}$
- la decomposizione spettrale moltiplicativa di $\mathbf{A}$
- una matrice simmetrica reale $\mathbf{W}$ tale che $\mathbf{A} = \mathbf{W}\mathbf{W}^T$
- la matrice di proiezione sullo spazio nullo $\mathrm{N}(\mathbf{A})$

Rispettivamente

1. Il polinomio caratteristico è

$$\begin{aligned}
\det \begin{bmatrix} 1-X & 0 & 2 \\ 0 & 1-X & 2 \\ 2 & 2 & 8-X \end{bmatrix} &= (1-X)\det \begin{bmatrix} 1-X & 2 \\ 2 & 8-X \end{bmatrix} + 2\det \begin{bmatrix} 0 & 2 \\ 1-X & 2 \end{bmatrix} \\
&= (1-X)\left[(1-X)(8-X)-4\right] + 2(-2(1-X)) \\
&= (1-X)\left[8 - X - 8X + X^2 - 4\right] - 4(1-X) \\
&= (1-X)\left[4 - 9X + X^2 - 4\right] \\
&= (1-X)(X^2 - 9X) \\
&= (1-X)(X)(X-9)
\end{aligned}$$

   Per cui $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 9$

2. per *decomposizione spettrale moltiplicativa* si intende la *diagonalizzazione unitaria*: essa è fattibile perché la matrice è normale, ossia $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ (verificato con maxima). Ora, dato che i 3 autovalori sono differenti, gli autovettori saranno di spazi indipendenti, quindi ortogonali. Ci basta prendere un autovettore per ciascun autovalore, normalizziamo a norma unitaria, assembliamo in matrice, calcoliamo l'inversa che (dato che unitaria, è poi la trasposta e giungiamo a fattorizzazione.
   Qui dato che le molteplicità geometriche sono sempre 1, i rango delle varie $\mathbf{A} - \lambda_i\mathbf{I}$ sono sempre 2; quindi non si può risolvere direttamente 1 equazione a scelta del sistema nullo ma bisogna sceglierne 2 e risolvere il sistema derivante (la rimanente equazione è combinazione lineare delle 2 scelte e sarà certamente soddisfatta dalle soluzioni trovate).
   In merito a $\lambda_1 = 0$ si ha che $\mathrm{N}(\mathbf{A})$ si ottiene

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 8 \end{bmatrix} \to \begin{cases} x + 2z = 0 \\ y + 2z = 0 \\ z = h \end{cases} \quad \begin{cases} x = -2h \\ y = -2h \\ z = h \end{cases} \to \left\langle \begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix} \right\rangle$$

   dove teniamo $-1\begin{bmatrix} -2 & -2 & 1 \end{bmatrix}^T = \begin{bmatrix} 2 & 2 & -1 \end{bmatrix}^T$ per ridurre i segni negativi.

In merito a $\lambda_2 = 1$, $\mathrm{N}(\mathbf{A} - \mathbf{I})$ è determinato come

$$\begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 2 \\ 2 & 2 & 7 \end{bmatrix} \begin{cases} 2x_3 = 0 \\ x_2 = h \\ 2x_1 + 2h + 7 \cdot 0 = 0 \end{cases} \quad \begin{cases} x_3 = 0 \\ x_2 = h \\ x_1 = -h \end{cases} \quad \to \left\langle \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\rangle$$

Teniamo proprio $\begin{bmatrix} -1 & 1 & 0 \end{bmatrix}^T$.
Infine per $\lambda_3 = 9$ si ha

$$\begin{bmatrix} -8 & 0 & 2 \\ 0 & -8 & 2 \\ 2 & 2 & -1 \end{bmatrix} \to \begin{cases} 4x - z = 0 \\ 4y - z = 0 \\ z = h \end{cases} \quad \begin{cases} x = h/4 \\ y = h/4 \\ z = h \end{cases} \quad to \left\langle \begin{bmatrix} 1/4 \\ 1/4 \\ 1 \end{bmatrix} \right\rangle$$

Per questa teniamo $4 \cdot \begin{bmatrix} 1/4 & 1/4 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 1 & 4 \end{bmatrix}$.
Andiamo a normalizzare i tre vettori, calcolandone le norme

$$\left\| \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix} \right\| = \sqrt{4+4+1} = 3 \to \begin{bmatrix} 2/3 \\ 2/3 \\ -1/3 \end{bmatrix}$$

$$\left\| \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\| = \sqrt{2} \to \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

$$\left\| \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} \right\| = \sqrt{18} = 3\sqrt{2} \to \begin{bmatrix} 1/3\sqrt{2} \\ 1/3\sqrt{2} \\ 4/3\sqrt{2} \end{bmatrix}$$

Pertanto una diagonalizzazione possibile vede

$$\mathbf{S} = \begin{bmatrix} 2/3 & -1/\sqrt{2} & 1/3\sqrt{2} \\ 2-3 & 1/\sqrt{2} & 1/3\sqrt{2} \\ -1/3 & 0 & 4/3\sqrt{2} \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 9 \end{bmatrix}$$

$$\mathbf{S}^{-1} = \mathbf{S}^H = \begin{bmatrix} 2/3 & 2/3 & -1/3 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/3\sqrt{2} & 1/3\sqrt{2} & 4/3\sqrt{2} \end{bmatrix}$$

3. si noti che $\mathbf{A}$ è semi-definita positiva dato che $\lambda_i \geq 0$, quindi la matrice richiesta $\mathbf{W}$ tale che $\mathbf{A} = \mathbf{W}\mathbf{W}^H$ esiste; qui la si vuole anche simmetrica. In attesa di trovare un eventuale riferimento algoritmico nella teoria, piuttosto che affidarsi a intuizioni geniali, si può notare come la diagonalizzazione appena fatta possa tornare comoda in quanto, se si pone $\mathbf{W} = \mathbf{S}\sqrt{\mathbf{D}}\mathbf{S}^T$ con $\sqrt{\mathbf{D}} = \mathbf{Diag}\left(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}\right)$ si ha che, effettivamente, essa è simmetrica in quanto

$$\mathbf{W}^T = (\mathbf{S}\sqrt{\mathbf{D}}\mathbf{S}^T)^T = \mathbf{S}\sqrt{\mathbf{D}}^T\mathbf{S}^T = \mathbf{S}\sqrt{\mathbf{D}}\mathbf{S}^T = \mathbf{W}$$

e inoltre, dato che $\mathbf{S}^T = \mathbf{S}^{-1}$

$$\mathbf{W}\mathbf{W}^T = \mathbf{S}\sqrt{\mathbf{D}}\mathbf{S}^T\mathbf{S}\sqrt{\mathbf{D}}\mathbf{S}^T = \mathbf{S}\sqrt{\mathbf{D}}\sqrt{\mathbf{D}}\mathbf{S}^T = \mathbf{S}\mathbf{D}\mathbf{S}^T = \mathbf{A}$$

Pertanto

$$\begin{aligned}\mathbf{W} &= \begin{bmatrix} 2/3 & -1/\sqrt{2} & 1/3\sqrt{2} \\ 2-3 & 1/\sqrt{2} & 1/3\sqrt{2} \\ -1/3 & 0 & 4/3\sqrt{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2/3 & 2/3 & -1/3 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/3\sqrt{2} & 1/3\sqrt{2} & 4/3\sqrt{2} \end{bmatrix} \\ &= \ldots maxima \ldots = \begin{bmatrix} 2/3 & -1/3 & 2/3 \\ -1/3 & 2/3 & 2/3 \\ 2/3 & 2/3 & 8/3 \end{bmatrix}\end{aligned}$$

e si verifica che effettivamente (sempre con maxima) che $\mathbf{A} = \mathbf{W}\mathbf{W}^T$

4. infine, ottenuta una base $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ ortonormale di $\mathrm{N}(\mathbf{A})$ calcoliamo $\mathbf{P}$ come

$$\mathbf{P} = \sum_i \mathbf{u}_i\mathbf{u}_i^T$$

La base ortonormale di $\mathrm{N}(\mathbf{A})$ l'avevamo calcolata in precedenza e risultava

$$\begin{bmatrix} 2/3 \\ 2/3 \\ -1/3 \end{bmatrix}$$

Pertanto

$$\mathbf{P} = \begin{bmatrix} 2/3 \\ 2/3 \\ -1/3 \end{bmatrix} \begin{bmatrix} 2/3 & 2/3 & -1/3 \end{bmatrix} = \begin{bmatrix} 4/9 & 4/9 & -2/9 \\ 4/9 & 4/9 & -2/9 \\ -2/9 & -2/9 & 1/9 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 4 & 4 & -2 \\ 4 & 4 & -2 \\ -2 & -2 & 1 \end{bmatrix}$$

Esempio 6.5.3. Si provi che la matrice

$$\mathbf{A} = \begin{bmatrix} 2 & -4 & 0 \\ -4 & 9 & -1 \\ 0 & -1 & 4 \end{bmatrix}$$

è definita positiva e si trovi una matrice $\mathbf{W}$ tale che $\mathbf{A} = \mathbf{W}\mathbf{W}^T$.
$\mathbf{A}$ è definita positiva se ha tutti gli autovalori > 0; o li si calcola fattorizzando il polinomio caratteristico o si calcola il determinante delle sottomatrici principali k-esime e verificando che siano tutti positivi; seguiamo questa seconda strada:

$$\begin{aligned}\det[2] &= 2 > 0 \\ \det\begin{bmatrix} 2 & -4 \\ -4 & 9 \end{bmatrix} &= 18 - 16 = 2 > 0 \\ \det\begin{bmatrix} 2 & -4 & 0 \\ -4 & 9 & -1 \\ 0 & -1 & 4 \end{bmatrix} &= 2\det\begin{bmatrix} 9 & -1 \\ -1 & 4 \end{bmatrix} - 4(-1)^{1+2}\det\begin{bmatrix} -4 & -1 \\ 0 & 4 \end{bmatrix} \\ &= 2(36-1) + 4(-16) = 70 - 64 = 6 > 0\end{aligned}$$

Quindi sì, la matrice è definita positiva.
Per trovare la matrice $\mathbf{W}$ richiesta il primo passo (suggerito dalla dimostrazione

della caratterizzazione di queste matrici) è fattorizzare $\mathbf{A} = \mathbf{L}_0\mathbf{U}_0$, ossia trovare le matrici $E_{ij}(\alpha)$ nella riduzione di Gauss che rendono $\mathbf{A}$ in forma ridotta $\mathbf{U}_0$: $\mathbf{L}_0$ è la moltiplicazione di queste (in ordine inverso) e $\mathbf{U}_0$ è la matrice in forma ridotta (ma con pivot ancora non unitari); dopodiché splittare $\mathbf{U}_0$ in 2 matrici

$$\mathbf{U}_0 = \mathbf{Diag}(d_1, \ldots, d_n)\mathbf{U}$$

pre-moltiplicandola per la diagonale con i pivot di $\mathbf{U}_0$, dividendo poi le righe di $\mathbf{U}_0$ per tali pivot, al fine di avere i pivot unitari in $\mathbf{U}$. Trovati $\mathbf{L}_0$ e $\mathbf{D}$ si ha poi che $\mathbf{W} = \mathbf{L}_0\mathbf{D}^{1/2}$.
Partiamo col tener traccia delle matrici impiegate nella riduzione. Qui siamo in pre-moltiplicazione quindi $E_{ij}(\alpha)$, ricordiamo, somma alla riga i la riga j moltiplicata per α

$$\begin{bmatrix} 2 & -4 & 0 \\ -4 & 9 & -1 \\ 0 & -1 & 4 \end{bmatrix} \xrightarrow{E_{21}(2)} \begin{bmatrix} 2 & -4 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 4 \end{bmatrix} \xrightarrow{E_{32}(1)} \underbrace{\begin{bmatrix} 2 & -4 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 3 \end{bmatrix}}_{\mathbf{U}_0}$$

Per cui si ha che

$$E_{32}(1)E_{21}(2)\mathbf{A} = \mathbf{U}_0$$

e dato che

$$(E_{32}(1)E_{21}(2))^{-1} = E_{21}(2)^{-1}E_{32}(1)^{-1} = E_{21}(-2)E_{32}(-1)$$

Si ha che

$$\mathbf{A} = \underbrace{E_{21}(-2)E_{32}(-1)}_{\mathbf{L}_0}\mathbf{U}_0$$

Pertanto

$$\mathbf{L}_0 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

Ora fattorizziamo $\mathbf{U}_0$ come

$$\mathbf{U}_0 = \underbrace{\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{U}}$$

Pertanto

$$\mathbf{W} = \mathbf{L}_0\mathbf{D}^{1/2} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{3} \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ -2\sqrt{2} & 1 & 0 \\ 0 & -1 & \sqrt{3} \end{bmatrix}$$

Esempio 6.5.4. • Si dica per quali valori del numero reale r la matrice

$$\mathbf{A}(r) = \begin{bmatrix} 9 & 3 & 3 \\ 3 & 1 & 1 \\ 3 & 1 & r \end{bmatrix}$$

è semi-definita positiva e per quali r è definita positiva

- si trovi un vettore $\mathbf{v}$ di $\mathbb{R}^3$ tale che $\mathbf{A}(1) = \mathbf{v}\mathbf{v}^T$ e si trovino gli autovalori di $\mathbf{A}(1)$

Allora qui si è provato con il polinomio caratteristico e ricerca delle radici ma non è cosa banale. Sfruttiamo allora la caratterizzazione di queste matrici: $\mathbf{A}$ è definita positiva se tutte le sottomatrici principali k-esime hanno determinante positivo; si ha che

$$\det[9] = 9 > 0$$
$$\det \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix} = 9 - 9 = 0$$

Pertanto la matrice non è definita positiva; può essere semidefinita positiva se ogni sottomatrice principale ha determinante ≥ 0. Partiamo dal calcolo più lungo

$$\begin{aligned} \det(\mathbf{A}) &= 3(-1)^{2+1}[3 \cdot r - 3] + 1(-1)^{2+2}[9 \cdot r - 9] + 1(-1)^{2+3}[9 \cdot 1 - 9] \\ &= -3(3r - 3) + 9r - 9 \\ &= -9r + 9 + 9r - 9 = 0 \end{aligned}$$

Quindi qui non ci sono problemi: vediamo ora le sottomatrici principali rimanenti

$$\begin{aligned} \det[1] &= 1 \\ \det[r] &= r \implies r \geq 0 \\ \det \begin{bmatrix} 9 & 3 \\ 3 & r \end{bmatrix} &= 9r - 9 \implies r \geq 1 \\ \det \begin{bmatrix} 1 & 1 \\ 1 & r \end{bmatrix} & r - 1 \implies r \neq 1 \end{aligned}$$

Pertanto affinché la matrice sia semidefinita positiva deve essere $r > 1$.
Il punto 2

TODO: da fare

Esempio 6.5.5. Si consideri la matrice

$$A = \begin{bmatrix} \sqrt{2} & 1+i & 0 \\ 1-i & 2 & i/2 \\ 0 & -i/2 & 1 \end{bmatrix}$$

1. Si determinino tramite i loro centri ed i loro raggi i tre cerchi di Gerschgorin C_1, C_2, C_3 associati alla matrice $\mathbf{A}$
2. dall'esame di C_1, C_2, C_3 e dal teorema di Gerschgorin si deduca che $\mathbf{A}$ è una matrice semi-definita positiva

I 3 cerchi sono:

$$\begin{aligned} C_1(\mathbf{A}) &= \left\{ z \in \mathbb{C} : \left| z - \sqrt{2} \right| \leq |1 + i| \right\} \\ C_2(\mathbf{A}) &= \{ z \in \mathbb{C} : |z - 2| \leq |1 - i| + |i/2| \} \\ C_3(\mathbf{A}) &= \left\{ z \in \mathbb{C} : |z - 1| \leq \left| -\frac{i}{2} \right| \right\} \end{aligned}$$

Pertanto il primo ha centro in $\sqrt{2}$ e raggio $\sqrt{2}$, il secondo centro in 2 e raggio $\sqrt{2}+\frac{1}{2} \approxeq 1.9$, mentre il terzo centro in $+1$ e raggio $1/2$.
Si noti che $\mathbf{A}$ è hermitiana quindi gli autovalori di $\mathbf{A}$ sono reali; inoltre sono comunque racchiusi nei 3 cerchi.
I 3 cerchi incrociano l'asse reale in zone positive (C_2, C_3) o al più nulle (C_1) pertanto gli autovalori sono positivi o al più nulli, ma mai negativi. Quindi la matrice è sicuramente semi definita positiva (potrebbe essere anche definita positiva, eventualmente, se tutti gli autovalori fossero positivi).

Esempio 6.5.6. È data la matrice dipendente dal parametro *complesso* z:

$$\mathbf{A}(z) = \begin{bmatrix} -2 & z \\ 1 & 3 \end{bmatrix}$$

1. Dire, giustificandolo, per quali valori di z la matrice è unitariamente diagonalizzabile
2. Si trovi una diagonalizzazione di $\mathbf{A}(0)$
3. Si dica, giustificandolo, per quali valori *reali* di z la matrice è ortogonalmente triangolarizzabile in $\mathbb{R}$
4. Si trovi una triangolarizzazione ortogonale di $A(-25/4)$

Rispettivamente:

1. la matrice è unitariamente diagonalizzabile solo se normale, ossia $\mathbf{A}^H\mathbf{A} = \mathbf{A}\mathbf{A}^H$; si ha che

$$\mathbf{A}\mathbf{A}^H = \begin{bmatrix} -2 & z \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ \overline{z} & 3 \end{bmatrix} = \begin{bmatrix} 4+z\overline{z} & -2+3z \\ -2+3\overline{z} & 10 \end{bmatrix}$$

$$\mathbf{A}^H\mathbf{A} = \begin{bmatrix} -2 & 1 \\ \overline{z} & 3 \end{bmatrix} \begin{bmatrix} -2 & z \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 5 & -2z+3 \\ -2\overline{z}+3 & z\overline{z}+9 \end{bmatrix}$$

   Si ha che $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$ se

$$\begin{cases} 5 = 4 + z\overline{z} \\ -2+3z = -2z+3 \\ -2+3\overline{z} = -2\overline{z}+3 \\ 10 = z\overline{z}+9 \end{cases} \qquad \begin{cases} |z| = 1 \\ z = 1 \\ \overline{z} = 1 \\ |z| = 1 \end{cases}$$

   ossia se $z = 1$

2. per trovare una diagonalizzazione di

$$\mathbf{A}(0) = \begin{bmatrix} -2 & 0 \\ 1 & 3 \end{bmatrix}$$

   notiamo che $\mathbf{A}(0)$ è diagonalizzabile poiché ha autovalori distinti:

$$p_{\mathbf{A}}(X) = (-2-X)(3-X) = -6+2X-3X+X^2 = X^2-X-6$$

   da cui

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1+24}}{2} = \frac{+1 \pm 5}{2} \implies \lambda_1 = 3, \lambda_2 = -2$$

Ora troviamo un autovettore per ciascun autovalore; essi appartengono a $\mathrm{N}(\mathbf{A}-\lambda_i\mathbf{I})$. Per $\lambda_1 = 3$

$$\begin{bmatrix} -2-3 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -5 & 0 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 5 & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix} \to 5x + 0y = 0 \iff x = 0 \to \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Per $\lambda_2 = -2$

$$\begin{bmatrix} 0 & 0 \\ 1 & 5 \end{bmatrix} \to 1x + 5y = 0 \iff \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

Quindi

$$\mathbf{S} = \begin{bmatrix} 0 & 5 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}, \quad \mathbf{S} = \frac{1}{-5} \begin{bmatrix} -1 & -5 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 1/5 & 1 \\ 1/5 & 0 \end{bmatrix}$$

e verificando effettivamente $\mathbf{SDS}^{-1} = \mathbf{A}$

3. $\mathbf{A}$ è ortogonalmente triangolarizzabile se reale e se tutti i suoi autovalori sono reali; dato che z è per ipotesi reale la matrice è allora reale. Troviamo i suoi autovalori

$$\begin{aligned} p_{\mathbf{A}(z)}(X) &= (-2-X)(3-X) - z = -6 + 2X - 3X + X^2 - z \\ &= X^2 - X - (z-6) \end{aligned}$$

da cui

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1 + 4z + 24}}{2}$$

Pertanto è ortogonalmente triangolarizzabile se $4z + 25 \geq 0$ ossia $z \geq -25/4$

4. per trovare una triangolarizzazione unitaria di

$$\mathbf{A}(-25/4) = \begin{bmatrix} -2 & -25/4 \\ 1 & 3 \end{bmatrix}$$

partiamo dal calcolo degli autovalori

$$\begin{aligned} p_{\mathbf{A}}(X) &= (-2-X)(3-X) + \frac{25}{4} = -6 + 2X - 3X + X^2 + \frac{25}{4} \\ &= X^2 - X + \frac{1}{4} \end{aligned}$$

da cui

$$\lambda_{1,2} = \frac{1 \pm \sqrt{1-1}}{2} = \frac{1}{2}$$

Troviamo un autovettore di $\lambda = 1/2$

$$\begin{bmatrix} -2-1/2 & -25/4 \\ 1 & 3-1/2 \end{bmatrix} \begin{bmatrix} -5/2 & -25/4 \\ 1 & 5/2 \end{bmatrix} \begin{bmatrix} 5/2 & 25/4 \\ 5/2 & 25/4 \end{bmatrix} \to \frac{5}{2}x + \frac{25}{4}y = 0 \iff 2x + 5y = 0 \to \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$

Completamento a base si può fare mediante $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ e per normalizzare entrambi i vettori hanno modulo $\sqrt{29}$, dunque

$$\mathbf{U} = \frac{1}{\sqrt{29}} \begin{bmatrix} 5 & 2 \\ -2 & 5 \end{bmatrix}$$

$$\mathbf{U}^{-1} = \frac{1}{25/29 + 4/29} \cdot \frac{1}{\sqrt{29}} \begin{bmatrix} 5 & -2 \\ 2 & \\ 5 & \end{bmatrix} = \frac{1}{\sqrt{29}} \begin{bmatrix} 5 & -2 \\ 2 & \\ 5 & \end{bmatrix}$$

e si verifica che $\mathbf{U}$ è ortogonale in quanto $U^{-1} = \mathbf{U}^T$. Infine in calcolo della triangolare:

$$\mathbf{T} = \mathbf{U}^{-1}\mathbf{A}\mathbf{U} = \frac{1}{29}\begin{bmatrix} 5 & -2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} -2 & -\frac{25}{4} \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 5 & 2 \\ -2 & 5 \end{bmatrix} = maxima = \begin{bmatrix} \frac{1}{2} & -29/4 \\ 0 & 1/2 \end{bmatrix}$$

Esempio 6.5.7. È data la matrice

$$\mathbf{A} = \begin{bmatrix} (1+\sqrt{2})/2 & 1 \\ -3/4 & (-1+\sqrt{2})/2 \end{bmatrix}$$

Trovare:

1. il polinomio caratteristico di $\mathbf{A}$ e i suoi autovalori
2. una diagonalizzazione $\mathbf{A} = \mathbf{S}\mathbf{D}\mathbf{S}^{-1}$

Rispettivamente:

1. il polinomio caratteristico è

$$\begin{aligned} p_{\mathbf{A}}(X) = \det(\mathbf{A} - \mathbf{I}X) = \det & \begin{bmatrix} \frac{1+\sqrt{2}}{2} - X & 1 \\ -3/4 & \frac{-1+\sqrt{2}}{2} - X \end{bmatrix} \\ &= \left(\frac{1+\sqrt{2}-2X}{2}\right)\left(\frac{-1+\sqrt{2}-2X}{2}\right) + \frac{3}{4} \\ &= \frac{-1+\sqrt{2}-2X-\sqrt{2}+2-2\sqrt{2}X+2X-2\sqrt{2}X+4X^2}{4} + \frac{3}{4} \\ &= \frac{4-4\sqrt{2}X+4X^2}{4} \\ &= X^2 - \sqrt{2} + 1 \end{aligned}$$

Gli autovalori sono dunque

$$\lambda_{1,2} = \frac{\sqrt{2} \pm \sqrt{2-4}}{2} = \frac{\sqrt{2} \pm i\sqrt{2}}{2} = \frac{1 \pm i}{\sqrt{2}}$$

2. gli autovalori sono diversi quindi la matrice è diagonalizzabile; dato che la molteplicità geometrica di ciascun autovalore è 1 basterà risolvere una equazione del sistema alla base di $\mathrm{N}(A - \lambda_i \mathbf{I})$ per trovare un autovettore (l'altra è combinazione lineare dell'equazione risolta e può essere ignorata). Per l'autovalore $(1+i)/\sqrt{2}$ si ha che gli autovettori risiedono in $\mathrm{N}\left(\mathbf{A} - \frac{1+i}{\sqrt{2}}\mathbf{I}\right)$

$$\begin{bmatrix} \frac{1+\sqrt{2}}{2} - \frac{1+i}{\sqrt{2}} & 1 \\ -3/4 & \frac{-1+\sqrt{2}}{2} - \frac{1+i}{\sqrt{2}} \end{bmatrix}$$

Possiamo risolverne solo una di equazione, dato che la dimensione dello spazio nullo è 1 e dunque il rango 1. Scegliendo la prima si ha

$$\begin{aligned}
&\frac{1+\sqrt{2}}{2} - \frac{1+i}{\sqrt{2}}x + 1y = 0\\
&\frac{1+\sqrt{2}-\sqrt{2}-\sqrt{2}i}{2}x + y = 0\\
&(1-\sqrt{2}i)x + 2y = 0\\
&\begin{bmatrix} -2 \\ 1-i\sqrt{2} \end{bmatrix}
\end{aligned}$$

Per quanto riguarda $\lambda_2 = \frac{1-i}{\sqrt{2}}$ si ha

$$\mathbf{A} - \lambda_2\mathbf{I} = \begin{bmatrix} (1+\sqrt{2})/2 - (1-i)/\sqrt{2} & 1 \\ -3/4 & (-1+\sqrt{2})/2 - (1-i)/\sqrt{2} \end{bmatrix}$$

da cui scegliendo la prima equazione

$$\begin{aligned}
&\frac{1+\sqrt{2}-\sqrt{2}+\sqrt{2}i}{2}x + 1y = 0\\
&(1+\sqrt{2}i)x + 2y = 0 \begin{bmatrix} -2 \\ 1+\sqrt{2}i \end{bmatrix}
\end{aligned}$$

Pertanto

$$\mathbf{S} = \begin{bmatrix} -2 & -2 \\ 1-i\sqrt{2} & 1+i\sqrt{2} \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} \frac{1+i}{\sqrt{2}} & 0 \\ 0 & \frac{1-i}{\sqrt{2}} \end{bmatrix}$$

e infine

$$\begin{aligned}
\mathbf{S}^{-1} &= \frac{1}{(-2)(1+i\sqrt{2}) + 2(1-i\sqrt{2})} \begin{bmatrix} 1+i\sqrt{2} & 2 \\ -1+i\sqrt{2} & -2 \end{bmatrix}\\
&= \frac{1}{-2-2i\sqrt{2}+2-2\sqrt{2}i} \begin{bmatrix} 1+i\sqrt{2} & 2 \\ -1+i\sqrt{2} & -2 \end{bmatrix}\\
&= -\frac{1}{4\sqrt{2}i} \begin{bmatrix} 1+i\sqrt{2} & 2 \\ -1+i\sqrt{2} & -2 \end{bmatrix}
\end{aligned}$$

Esempio 6.5.8. È data la matrice

$$\mathbf{A} = \begin{bmatrix} 0 & 4 \\ -1 & 4 \end{bmatrix}$$

1. Si trovi il polinomio caratteristico $p_{\mathbf{A}}(X)$ di $\mathbf{A}$
2. Si dica, giustificandolo, se $\mathbf{A}$ è o non è diagonalizzabile in $M_2(\mathbb{C})$
3. Si trovi una triangolarizzazione unitaria $\mathbf{A} = \mathbf{QTQ}^T$ con $\mathbf{Q}$ matrice reale ortogonale e $\mathbf{T}$ matrice triangolare superiore

Rispettivamente:

1. si ha che

   $$p_{\mathbf{A}}(X) = (-X)(4 - X) + 4 = -4X + X^2 + 4 = X^2 - 4X + 4 = (X - 2)^2$$

2. la matrice è diagonalizzabile in $\mathbb{C}$ se ha autovalori diversi o con medesima molteplicità (geometrica = algebrica). Gli autovalori sono

   $$\lambda_{1,2} = \lambda = 2$$

   Dobbiamo verificare che la molteplicità geometrica di $\lambda = 2$ sia pari a 2 (alla molteplicità algebrica). Dato che

   $$2 = \dim \mathrm{C}\left(\mathbf{A} - 2\mathbf{I}\right) + \dim \mathrm{N}\left(\mathbf{A} - 2\mathbf{I}\right)$$

   , affinché $\dim \mathrm{N}\left(\mathbf{A} - 2\mathbf{I}\right) = 2$ dovrebbe essere

   $$\dim \mathrm{C}\left(\mathbf{A} - 2\mathbf{I}\right) = \mathrm{rk}\, \mathrm{C}\left(\mathbf{A} - 2\mathbf{I}\right) = 0$$

   (ma quale matrice ha rango 0?). Comunque, facciamo la verifica

   $$\begin{bmatrix} 0-2 & 4 \\ -1 & 4-2 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ -2 & 4 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ 0 & 0 \end{bmatrix}$$

   Si ha che il rango è 1, molteplicità geometrica $1 \neq 2$ e la matrice non è diagonalizzabile

un'altra volta

Esempio 6.5.9. È data la matrice dipendente dal parametro reale a:

$$\mathbf{A}(a) = \begin{bmatrix} 2 & 0 & -1 \\ 0 & -1 & 1/2 \\ 0 & a & 2 \end{bmatrix}$$

1. Si dica per quali valori di a la matrice è diagonalizzabile
2. Si trovi una diagonalizzazione esplicita di $\mathbf{A} = \mathbf{A}(-4)$ (cioè tre matrici $\mathbf{S}$, $\mathbf{D}$ e $\mathbf{S}^{-1}$ tali che $\mathbf{A} = \mathbf{SDS}^{-1}$)

Rispettivamente

1. procediamo al calcolo gli autovettori determinando il polinomio caratteristico

   $$\begin{aligned} p_{\mathbf{A}}(X) &= \det \begin{bmatrix} 2-X & 0 & -1 \\ 0 & -1-X & 1/2 \\ 0 & a & 2-X \end{bmatrix} = (2-X)\left[(-1-X)(2-X) - \frac{a}{2}\right] \\ &= (2-X)\left[X^2 - X - 2 - \frac{a}{2}\right] \\ &= 2X^2 - 2X - 4 - a - X^3 + X^2 + 2X + \frac{aX}{2} \\ &= -X^3 + 3x^2 + \frac{aX}{2} - 4 - a \end{aligned}$$

Ora vediamo che

$$p_{\mathbf{A}}(2) = -8 + 12 + a - 4 - a = 0$$

Pertanto $X = 2$ è una radice del polinomio, ossia un autovettore e possiamo fattorizzarlo per trovare più agevolmente i rimanenti. Applicando Ruffini

$$\begin{array}{c|cccc} & -1 & 3 & a/2 & -4-a \\ 2 & & -2 & 2 & a+a \\ \hline & -1 & 1 & (a+4)/2 & 0 \end{array}$$

Si ha dunque che

$$p_{\mathbf{A}}(X) = (X-2)(-X^2 + X + \frac{a+4}{2})$$

Troviamo le rimanenti radici

$$\lambda_{2,3} = \frac{-1 \pm \sqrt{1 + 2(a+4)}}{-2} = \frac{1 \pm \sqrt{2a+9}}{-2}$$

Le radici λ_2 e λ_3 sono distinte se $a \neq -\frac{9}{2}$; viceversa se $a \neq -\frac{9}{2}$ però può essere che $\lambda_1 = \lambda_3 = 2$ oppure $\lambda_1 = \lambda_2 = 2$. Vediamo allora per quali valori di a ciò accade:

$$2 = \frac{-1 \pm \sqrt{2a+9}}{-2} \iff -4 = -1 \pm \sqrt{2a+9} \iff -3 = \pm\sqrt{2a+9}$$

da cui il ramo $-3 = \sqrt{2a+9}$ è impossibile e non viene considerato, ma si ha che $-3 = -\sqrt{2a+9}$ se $a = 0$. Pertanto anche se $a = 0$ si ha che due autovalori coincidono e in tal caso può essere che la matrice non sia diagonalizzabile (e occorre pertanto passare all'analisi delle molteplicità per concludere lo studio).
Sintetizzando i casi di interesse sono $a = 0 \vee a = -\frac{9}{2}$; in tutti gli altri casi la matrice è diagonalizzabile.
Se $a = 0$ si ha

$$\mathbf{A}(0) = \begin{bmatrix} 2 & 0 & -1 \\ 0 & -1 & 1/2 \\ 0 & 0 & 2 \end{bmatrix}$$

e $\lambda_1 = 2$, $\lambda_2 = \frac{-1+3}{-2} = -1$, $\lambda_3 = \frac{-1-3}{-2} = 2$. L'autovalore -1 non pone problemi (la sua molteplicità geometrica coinciderà con quella aritmetica dato che vi è sempre inferiore; bisognerà invece controllare la molteplicità algebrica di 2 verificando che sia 2, ossia che $\dim \mathrm{N}(\mathbf{A} - 2\mathbf{I}) = 2$. Ricordando dal teorema nullità+rango che

$$3 = \dim \mathrm{C}(\mathbf{A} - 2\mathbf{I}) + \dim \mathrm{N}(\mathbf{A} - 2\mathbf{I})$$

controlliamo equivalentemente che $\dim \mathrm{C}(\mathbf{A} - 2\mathbf{I}) = \mathrm{rk}(\mathbf{A} - 2\mathbf{I}) = 1$ (se fosse così la matrice sarebbe diagonalizzabile

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 2-2 & 0 & -1 \\ 0 & -1-2 & 1/2 \\ 0 & 0 & 2-2 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & -3 & 1/2 \\ 0 & 0 & 0 \end{bmatrix}$$

Effettuando la riduzione ci accorgiamo che la matrice ha rango 2, pertanto per $a = 0$ la matrice non è diagonalizzabile.
Nel caso $a = -\frac{9}{2}$ si ha

$$\mathbf{A}(-9/2) = \begin{bmatrix} 2 & 0 & -1 \\ 0 & -1 & 1/2 \\ 0 & -9/2 & 2 \end{bmatrix}$$

e $\lambda_1 = 2$, $\lambda_2 = \lambda_3 = 1/2$, quindi qui la matrice è diagonalizzabile se $\dim \mathrm{N}(\mathbf{A} - \mathbf{I}/2) = 2$ o equivalentemente $\dim \mathrm{C}(\mathbf{A} - \mathbf{I}/2) = 1$. Si ha

$$\mathbf{A} - \frac{\mathbf{I}}{2} = \begin{bmatrix} 2-\frac{1}{2} & 0 & -1 \\ 0 & -1-1/2 & 1/2 \\ 0 & -9/2 & 3/2 \end{bmatrix} \begin{bmatrix} 3/2 & 0 & -1 \\ 0 & -3/2 & 1/2 \\ 0 & -9/2 & 3/2 \end{bmatrix} \begin{bmatrix} 3 & 0 & -2 \\ 0 & -3 & 1 \\ 0 & -9 & 3 \end{bmatrix} \begin{bmatrix} 3 & 0 & -2 \\ 0 & -3 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Il rango è dunque 2 e la nullità 1; pertanto la matrice non è diagonalizzabile per $a = -9/2$.
In sintesi la matrice è diagonalizzabile se $a \neq 0$, $a \neq -9/2$

2. per la diagonalizzazione di

$$\mathbf{A}(-4) = \begin{bmatrix} 2 & 0 & -1 \\ 0 & -1 & 1/2 \\ 0 & -4 & 2 \end{bmatrix}$$

gli autovalori sono $\lambda_1 = 2$, $\lambda_2 = \frac{-1 \pm \sqrt{-8+9}}{-2} = 0$ e $\lambda_3 = \frac{-1-1}{-2} = 1$.
Autovettore per $\lambda_1 = 2$ appartiene a $\mathrm{N}(\mathbf{A} - 2\mathbf{I})$; si ha che

$$\mathbf{A} - 2\mathbf{I} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & -3 & 1/2 \\ 0 & -4 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -1/6 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1/6 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{cases} x_1 = h \\ x_2 = 0 \\ x_3 = 0 \end{cases} \to \begin{bmatrix} h \\ 0 \\ 0 \end{bmatrix}$$

e teniamo $\mathbf{e}_1$.
Troviamo ora un auovettore per $\lambda_2 = 0$ che $\in \mathrm{N}(\mathbf{A})$:

$$\begin{bmatrix} 2 & 0 & -1 \\ 0 & -1 & 1/2 \\ 0 & -4 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -1/2 \\ 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 2 & -1 \\ 0 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1/2 \\ 0 & 1 & -1/2 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{cases} x_1 - \frac{1}{2}h_3 = 0 \\ x_2 - \frac{1}{2}h_3 = 0 \\ x_3 = h_3 \end{cases}$$

$$\begin{cases} x_1 = \frac{1}{2}h_3 \\ x_2 = \frac{1}{2}h_3 \\ x_3 = h_3 \end{cases} \to \begin{bmatrix} 1/2 \\ 1/2 \\ 1 \end{bmatrix} \to \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

Infine un autovettore per $\lambda_3 = 1$ è in $\mathrm{N}(\mathbf{A} - \mathbf{I})$

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & -2 & 1/2 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -4 & 1 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & -4 & 1 \\ 0 & 0 & 0 \end{bmatrix} \to \begin{cases} x_3 = h \\ -4x_2 + h = 0 \\ x_1 = 0 \end{cases} \quad \begin{cases} x_1 = h \\ x_2 = h/4 \\ x_3 = h \end{cases} \to \begin{bmatrix} 1 \\ 1/4 \\ 1 \end{bmatrix} \to \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix}$$

Quindi innanzitutto

$$\mathbf{S} = \begin{bmatrix} 1 & 1 & 4 \\ 0 & 1 & 1 \\ 0 & 2 & 4 \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Poi procediamo ad invertire $\mathbf{S}$

$$\begin{bmatrix} 1 & 1 & 4 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 4 & 0 & 0 & 1 \end{bmatrix} \dots \begin{bmatrix} 1 & 0 & 0 & 1 & 2 & -3/2 \\ 0 & 1 & 0 & 0 & 2 & -1/2 \\ 0 & 0 & 1 & 0 & -1 & 1/2 \end{bmatrix}$$

Per cui

$$\mathbf{S}^{-1} = \begin{bmatrix} 1 & 2 & -3/2 \\ 0 & 2 & -1/2 \\ 0 & -1 & 1/2 \end{bmatrix}$$

Esempio 6.5.10. È data la matrice dipendente dal parametro reale a

$$\mathbf{A}(a) = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 4 & 0 \\ a & 0 & 8 \end{bmatrix}$$

- Si dica per quali valori di a la matrice è diagonalizzabile
- Si trovi una triangolarizzazione unitaria di $\mathbf{A}(-4)$

Facciamo solo il secondo punto: una triangolarizzazione unitaria di

$$\mathbf{A}(-4) = \begin{bmatrix} 0 & 0 & 4 \\ 0 & 4 & 0 \\ -4 & 0 & 8 \end{bmatrix}$$

Determiniamo gli autovalori della matrice:

$$\begin{aligned} \det(\mathbf{A} - \mathbf{I}X) &= \det \begin{bmatrix} -X & 0 & 4 \\ 0 & 4-X & 0 \\ -4 & 0 & 8-X \end{bmatrix} \\ &= (4-X)(-1)^{2+2} \det \begin{bmatrix} -X & 4 \\ -4 & 8-X \end{bmatrix} \\ &= (4-X)(-X(8-X)+16) = (X-4)(X(8-X)-16) \\ &= (X-4)(8X - X^2 - 16) \end{aligned}$$

Quindi $\lambda_1 = 4$; inoltre

$$\lambda_{2,3} = \frac{-8 \pm \sqrt{64-64}}{-2} = 4$$

quindi $\lambda_1 = \lambda_2 = \lambda_3 = \lambda = 4$; un autovettore per λ si ha $\in \mathrm{N}(\mathbf{A} - 4\mathbf{I})$

$$\begin{bmatrix} -4 & 0 & 4 \\ 0 & 0 & 0 \\ -4 & 0 & 4 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \to -x + z = 0 \to \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Completiamo a base di $\mathbb{C}^3$

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

Sono dominanti le prime tre quindi prendiamo per base

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Ortogonalizziamo e poi normalizziamo

$$\mathbf{u}_1 = \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \alpha_{12}\mathbf{u}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \ldots = \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \alpha_{13}\mathbf{u}_1 - \alpha_{23}\mathbf{u}_2$$

$$= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}{\cdots} - \frac{\begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}^T \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}{\cdots} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Quindi la base ortogonale è

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1/2 \\ 0 \\ -1/2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Che normalizziamo costruendo la matrice

$$\mathbf{S} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}$$

Procediamo a calcolare $\mathbf{S}^{-1}$ mediante la trasposta

$$\mathbf{S}^{-1} = \mathbf{S}^T = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix}$$

Infine

$$\begin{aligned} \mathbf{T} &= \mathbf{S}^{-1}\mathbf{A}\mathbf{S} \\ &= \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 4 \\ 0 & 4 & 0 \\ -4 & 0 & 8 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \\ &= \ldots = \begin{bmatrix} 4 & 8 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} \end{aligned}$$

E una verifica mediante maxima conferma che $\mathbf{A} = \mathbf{S}\mathbf{T}\mathbf{S}^{-1}$.

Esempio 6.5.11. È data la matrice dipendente dal parametro reale α:

$$\mathbf{A}(\alpha) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & \alpha & 1 \\ -4 & 0 & 1 \end{bmatrix}$$

1. Si dica per quali valori di α la matrice è unitariamente diagonalizzabile

2. Si trovino i valori di α per cui la matrice è diagonalizzabile

3. Si trovi una diagonalizzazione di $\mathbf{A}(3)$

Rispettivamente:

1. $\mathbf{A}(\alpha)$ è unitariamente diagonalizzaible se normale ossia se $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}$. In questi casi non conviene sviluppare tutto un prodotto e poi tutto l'altro poiché ci si può fermare molto prima con un po' di fortuna. In questo caso già al primo passaggio in quanto

$$\mathbf{A}\mathbf{A}^H = \begin{bmatrix} 1 & 0 & -1 \\ 2 & \alpha & 1 \\ -4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 4 \\ 0 & \alpha & 0 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}$$
$$\mathbf{A}^H\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 \\ 0 & \alpha & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 2 & \alpha & 1 \\ -4 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 21 & \dots & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots \end{bmatrix}$$

   E dato che $2 \neq 21$ velocemente si conclude che nessun α può rendere $\mathbf{A}$ normale. Dunque $\mathbf{A}$ non è unitariamente diagonalizzabile

2. partiamo calcolando il polinomio caratteristico

$$p_{\mathbf{A}}(X) = \det \begin{bmatrix} 1-X & 0 & -1 \\ 2 & \alpha - X & 1 \\ -4 & 0 & 1-X \end{bmatrix}$$
$$= (\alpha - X)\left[(1-X)^2 - 4\right] = (\alpha - X)(1 + X^2 - 2X - 4)$$
$$= (\alpha - X)(X^2 - 2X - 3) = (X-3)(X+1)(\alpha - X)$$

   quindi $\lambda_1 = 3$, $\lambda_2 = -1$, $\lambda_3 = \alpha$. Ora se $\alpha \neq 3 \vee \alpha \neq -1$ la matrice è diagonalizzabile. Se invece:

   - $\alpha = -1$ la matrice è diagonalizzabile se $\dim \mathrm{N}(\mathbf{A} + \mathbf{I}) = 2$ ossia se $\dim \mathrm{C}(\mathbf{A} + \mathbf{I}) = 1$
   - $\alpha = 3$ la matrice è diagonalizzabile se $\dim \mathrm{N}(\mathbf{A} - 3\mathbf{I}) = 2$ ossia se $\dim \mathrm{C}(\mathbf{A} - 3\mathbf{I}) = 1$

   Nel caso $\alpha = -1$

$$\begin{bmatrix} 2 & 0 & -1 \\ 2 & 0 & 1 \\ -4 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ 2 & 0 & 1 \\ 2 & 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ 2 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Il rango è 2 quindi la dimensione dello spazio nullo è 1 (dovrebbe esser 2) e la matrice non è diagonalizzabile.
Nel caso $\alpha = 3$

$$\mathbf{A} - 3\mathbf{I} = \begin{bmatrix} -2 & 0 & -1 \\ 2 & 0 & 1 \\ -4 & 0 & -2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 2 & 0 & 1 \\ 2 & 0 & 1 \end{bmatrix}$$

Il rango è 1, quindi spazio nullo di dimensione 2 e per $\alpha = 3$ effettivamente la matrice è diagonalizzabile.

3. procediamo a diagonalizzare

$$\mathbf{A}(3) = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & 1 \\ -4 & 0 & 1 \end{bmatrix}$$

avente autovalori $\lambda_1 = 3, \lambda_2 = -1, \lambda_3 = \alpha = 3$. Costruiamo $\mathbf{D}$ ponendo gli autovalori in ordine crescente

$$\mathbf{D} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Costruiamo $\mathbf{S} = [\mathbf{s}_1\, \mathbf{s}_2\, \mathbf{s}_3]$ con $\mathbf{s}_1$ autovettore per $\lambda_2 = -1$ e $\mathbf{s}_2, \mathbf{s}_3$ due vettori indipendenti dell'autovalore 3. Procediamo con $\mathbf{s}_1 \in N(\mathbf{A} + \mathbf{I})$

$$\begin{bmatrix} 2 & 0 & -1 \\ 2 & 4 & 1 \\ -4 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ 2 & 4 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ 0 & 4 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{cases} x_3 = h \\ 2x_2 + h = 0 \\ 2x_1 - h = 0 \end{cases} \begin{cases} x_1 = h/2 \\ x_2 = -h/2 \\ x_3 = h \end{cases} \rightarrow \left\langle \begin{bmatrix} 1/2 \\ -1/2 \\ 2 \end{bmatrix} \right\rangle$$

e per questo prendiamo $\mathbf{s}_1 = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$. Ora troviamo due vettori indipendenti dell'autovalore 3. Tali vettori $\in \mathrm{N}(\mathbf{A} - 3\mathbf{I})$; se ne costruiamo una base, a 2 elementi, possiamo scegliere tali elementi per la nostra matrice

$$\begin{bmatrix} -2 & 0 & -1 \\ 2 & 0 & 1 \\ -4 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{cases} 2x_1 + h_3 = 0 \\ x_2 = h_2 \\ x_3 = h_3 \end{cases} \begin{cases} x_1 = -h_3/2 \\ x_2 = h_2 \\ x_3 = h_3 \end{cases} \rightarrow \begin{bmatrix} -h_3/2 \\ h_2 \\ h_3 \end{bmatrix} = h_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + h_3 \begin{bmatrix} -1/2 \\ 0 \\ 1 \end{bmatrix}$$

Dunque teniamo

$$\mathbf{s}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{s}_3 = \begin{bmatrix} -1 \\ 0 \\ 2 \end{bmatrix}$$

Costruiamo dunque

$$\mathbf{S} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix}$$

E invertiamola per aver $\mathbf{S}^{-1}$

$$\begin{bmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 2 & 0 & 2 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 4 & -2 & 0 & 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1/2 & 0 & 1/4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 & 1/2 & 0 & 1/4 \\ 0 & 1 & 0 & 1/2 & 1 & 1/4 \\ 0 & 0 & 1 & -1/2 & 0 & 1/4 \end{bmatrix}$$

Dunque

$$\mathbf{S}^{-1} = \begin{bmatrix} 1/2 & 0 & 1/4 \\ 1/2 & 1 & 1/4 \\ -1/2 & 0 & 1/4 \end{bmatrix}$$

e facendo le verifiche con maxima i calcoli tornano.

6.5.2 Fatti ma eventualmente da ricopiare

Esempio 6.5.12. È data la matrice

$$\mathbf{B} = \begin{bmatrix} 0 & 1+i \\ -1+i & 2 \end{bmatrix}$$

1. si trovino il polinomio caratteristico di $\mathbf{B}$ ed i suoi autovalori
2. si trovi una diagonalizzazione $\mathbf{B} = \mathbf{SDS}^{-1}$

Esempio 6.5.13. Si consideri la matrice dipendente dal parametro complesso β

$$\mathbf{B}(\beta) = \begin{bmatrix} 2 & \beta \\ 1 & 3 \end{bmatrix}$$

- Si trovi l'unico valore β_0 di β per cui $\mathbf{B}(\beta)$ è unitariamente diagonalizzabile
- Si determinino i valori di $\beta \in \mathbb{C}$ per cui $\mathbf{B}(\beta)$ è diagonalizzabile
- si trovi una triangolarizzazione unitaria di $\mathbf{B} = \mathbf{B}(-1/4)$

6.5.3 Da fare

Esempio 6.5.14. Si consideri la matrice dipendente dal parametro complesso α

$$\mathbf{A}(\alpha) = \begin{bmatrix} i & 1 \\ \alpha & 1+i \end{bmatrix}$$

- Si trovi l'unico valore α_0 di α per cui $\mathbf{A}(\alpha)$ è unitariamente diagonalizzabile
- Si determinino i valori $\alpha \in \mathbb{C}$ per cui $\mathbf{A}(\alpha)$ è diagonalizzabile
- Si trovi una triangolarizzazione unitaria di $\mathbf{A} = \mathbf{A}(-1/4)$

Esempio 6.5.15. È data la matrice dipendente dal parametro complesso

$$\mathbf{B}(z) = \begin{bmatrix} 1 & 5 \\ z & 2 \end{bmatrix}$$

1. Dire, giustificandolo, per quali valori di z la matrice è unitariamente diagonalizzabile
2. Si trovi una diagonalizzazione di $\mathbf{B}(0)$
3. Si dica, giustificandolo, per quali valori *reali* di z la matrice è ortogonalmente triangolarizzabile in $\mathbb{R}$
4. si trovi una triangolarizzazione ortogonale di $\mathbf{B}(-1/20)$

Esempio 6.5.16. È data la matrice reale dipendente dal parametro reale a

$$\mathbf{A}(a) = \begin{bmatrix} 0 & a & -1 \\ 1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix}$$

Si dica:

1. per quali valori di a la matrice è unitariamente diagonalizzabile
2. giustificando completamente, per quali valori di a la matrice è diagonalizzazione

Esempio 6.5.17. Data la matrice dipendente dal parametro reale β

$$\mathbf{B}(\beta) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & \beta & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

- Si dica per quali valori di β la matrice è unitariamente diagonalizzabile
- Si trovino i valori di β per cui la matrice è diagonalizzabile
- Si trovi una diagonalizzazione di $\mathbf{B}(1)$

Esempio 6.5.18.
1. Si trovino le uniche due matrici normali complesse 3×3 $\mathbf{A}$ e $\mathbf{B}$ che soddisfano i seguenti requisiti:
   - gli autovalori distinti sono 1 e i
   - un autospazio è generato dai due vettori $[1\,1\,1]^T$ e $[1\,0\,0]^T$
2. Si provi che $\mathbf{AB} = \mathbf{BA} = i\mathbf{I}_3$, $\mathbf{A} + \mathbf{B} = (1+i)\mathbf{I}$, $\mathbf{A}^2 + \mathbf{B}^2 = \mathbb{O}$

Esempio 6.5.19.
1. Si trovi l'unica matrice *normale* complessa $\mathbf{A}$ 3×3 con
   - due autovalori $\lambda_1 = 1$ di molteplicità algebrica 2 e $\lambda_2 = i$ di molteplicità algebrica 1, e con l'autospazio $E_{\mathbf{A}}(\lambda_1)$ generato dai due vettori
   $$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} i \\ 0 \\ 1 \end{bmatrix}$$
   (suggerimento: usare il teorema spettrale, nella versione moltiplicativa o in quella additiva)
   - Si scrivano espressamente sia la diagonalizzazione unitaria di $\mathbf{A}$ che la sua decomposizione spettrale additiva

Esempio 6.5.20. Sia $\mathbf{P}$ una matrice $n \times n$ complessa di proiezione, tale cioè che $\mathbf{P} = \mathbf{P}^2 = \mathbf{P}^H$:

1. si privi che gli unici possibili autovalori di $\mathbf{P}$ sono 0 e 1
2. usando il fatto che $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$ per ogni coppia di matrici $\mathbf{A}$ e $\mathbf{B}$ ed utilizzando il teorema spettrale per $\mathbf{P}$ si provi che $\mathrm{Tr}(\mathbf{P}) = \mathrm{rk}\,\mathbf{P}$

Esempio 6.5.21. Trovare la matrice normale $\mathbf{A}$ 3×3 complessa avente come autovalori $\lambda_1 = 4$, di molteplicità 1, e $\lambda_2 = -2$ di molteplicità 2 e tale che l'autospazio $E_{\mathbf{A}}(\lambda_2)$ è generato da due vettori

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Esempio 6.5.22. Trovare la matrice normale $\mathbf{B}$ 3×3 complessa avente come autovalori $\lambda_1 = 3$ di molteplicità 2, e $\lambda_2 = -3$ di molteplicità 1, e tale che l'autospazio $E_{\mathbf{B}}(\lambda_1)$ è generato dai due vettori

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ i \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} i \\ 1 \\ 0 \end{bmatrix}$$

Esempio 6.5.23.

1. Si trovi la matrice normale $\mathbf{A}$ 3×3 a coefficienti reali avente come autovalori $\lambda_1 = 2$ con relativo autospazio $E_{\mathbf{A}}(\lambda_1)$ generato dai vettori
$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$
e $\lambda_2 = 0$
2. si trovino le matrici di proiezione sugli autospazi $E_{\mathbf{A}}(\lambda_1)$ e $E_{\mathbf{A}}(\lambda_2)$

Esempio 6.5.24. Si provi che la matrice

$$\mathbf{B} = \begin{bmatrix} 4 & 0 & -1 \\ 0 & 2 & -4 \\ -1 & -4 & 9 \end{bmatrix}$$

è definita positiva e si trovi una matrice $\mathbf{V}$ tale che $\mathbf{B} = \mathbf{V}\mathbf{V}^T$

Esempio 6.5.25.

- Si dica per quali valori del numero reale r la matrice
$$\mathbf{B}(r) = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & r \end{bmatrix}$$
è semi-definita positiva e per quali r è definita positiva
- si trovi un vettore $\mathbf{w}$ di $\mathbb{R}^3$ tale che $\mathbf{B}(9) = \mathbf{w}\mathbf{w}^T$ e si trovino gli autovalori di $\mathbf{B}$

Esempio 6.5.26. Data la matrice

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 5 & 5 \\ 3 & 5 & 14 \end{bmatrix}$$

- Si provi che $\mathbf{A}$ è semidefinita positiva
- Si dica in quali dei tre cerchi di Gerschgorin C_1, C_2, C_3 stanno i tre autovalori di A

Esempio 6.5.27. 1. Si dica se la matrice realie simmetrica

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$$

è definita positiva o semi-definita positiva

2. Si trovi una matrice $\mathbf{W}$ tale che $\mathbf{A} = \mathbf{W}\mathbf{W}^T$

Esempio 6.5.28. • Siano $\mathbf{A}$ e $\mathbf{B}$ due matrici normali. Si provi che, se $\mathbf{A}$ e $\mathbf{B}^H$ commutano, allora il prodotto $\mathbf{AB}$ è normale

- Si provi che in generale non è vero il contrario (suggerimento: si trovino una matrice ortogonale reale 2×2 $\mathbf{A}$ e una matrice diagonale reale $\mathbf{B}$ di tipo particolare tali che $\mathbf{AB}$ è normale ma $\mathbf{AB}^T \neq \mathbf{B}^T\mathbf{A}$

Esempio 6.5.29. Sia $\mathbf{A}$ una matrice complessa $n \times n$ tale che $\mathbf{A} = \alpha \mathbf{A}^2$ ($\alpha \in \mathbb{C}$):

- si provi che gli autovalori di $\mathbf{A}$ possono assumere al più due valori
- si provi che se $\alpha \operatorname{Tr}(\mathbf{A}) = n$, allora $\mathbf{A} = \alpha^{-1}\mathbf{I}_n$

Esempio 6.5.30. Sia $\mathbf{A}$ una matrice $n \times n$ complessa normale. Si provi che $\mathbf{A}$ è sia unitaria che anti-hermitiana se e solo se $\mathbf{A}^2 = -\mathbf{I}_n$

Esempio 6.5.31. Sia $\mathbf{B}$ una matrice $n \times n$ complessa tale che $B^2 = -\mathbf{B}^4$:

- si provi che, se $\mathbf{B}$ è normale, allora è anche anti-hermitiana e $i\mathbf{B}$ è la differenza di due matrici di proiezione a prodotto nullo.
- se $\mathbf{B}$ è 3×3 e $\det \mathbf{B} = -i$, si provi che $\mathbf{B} = i\mathbf{I}_3$ oppure $\operatorname{Tr}(\mathbf{B}) = -i$

Esempio 6.5.32. Sia $\mathbf{A}$ una matrice $n \times n$ complessa tale che $A = A^3$:

- si provi che se λ è un autovalore di $\mathbf{A}$, allora $\lambda \in \{0, 1, -1\}$
- si provi che se $\mathbf{A}$ è normale, allora è anche hermitiana ed è la differenza di due matrici di proiezione a prodotto nullo

Esempio 6.5.33. Si provi che una matrice $n \times n$ $\mathbf{A}$ a coefficienti complessi è sia hermitiana che unitaria se e solo se $\mathbf{A} = \mathbf{P}_1 - \mathbf{P}_2$ con $\mathbf{P}_1, \mathbf{P}_2$ matrici di proiezione (di cui una possibilmente nulla) tali che $\mathbf{P}_1\mathbf{P}_2 = \mathbb{O} = \mathbf{P}_2\mathbf{P}_1$ e $\mathbf{P}_1 + \mathbf{P}_2 = \mathbf{I}_n$ (Suggerimento: usare il teorema spettrale additivo)

Esempio 6.5.34. Sia $\mathbf{A}$ una matrice $n\times n$ complessa normale. Sia $\mathbf{P}$ la matrice di proiezione su $\mathrm{N}(\mathbf{A})$. Si provi che esiste una matrice $\mathbf{B}$ tale che $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n - \mathbf{P}$ (suggerimento: usare una decomposizione spettrale additiva di $\mathbf{A}$)

Esempio 6.5.35. Sia $\mathbf{A}$ una matrice normale complessa con esattamente due autovalori distinti. Si provi che $\mathbf{A} = \alpha\mathbf{P} + \beta\mathbf{I}_n$ dove $\alpha \neq 0$ e $\mathbf{P}$ è una matrice di proiezione

Esempio 6.5.36.

- Data una matrice complessa $n \times n$ $\mathbf{A}$, si ponga $\mathbf{B} = (\mathbf{A} + \mathbf{A}^H)/2$ e $\mathbf{C} = (\mathbf{A} - \mathbf{A}^H)/2$. Si provi che $\mathbf{A}$ è normale se e solo se $\mathbf{B}$ e $\mathbf{C}$ commutano, cioè $\mathbf{BC} = \mathbf{CB}$
- Se $\mathbf{A}$ è normale con autovalori $\lambda_1 = a_1 + ib_1, \dots, \lambda_n = a_n + ib_n$ (dove a_j e b_j sono numeri reali per ogni j), si trovino gli autovalori di $\mathbf{B}$ e $\mathbf{C}$

Esempio 6.5.37. Sia $\mathbf{U}$ una matrice complessa unitaria $n \times n$:

- si spieghi perché, se $\mathbf{U} = \mathbf{QR}$ è una decomposizione QR normalizzata di $\mathbf{U}$ allora necessariamente $\mathbf{U} = \mathbf{Q}$ e $\mathbf{R} = \mathbf{I}$
- provare che se $\mathbf{U} = \mathbf{VT}$ con $\mathbf{V}$ pure unitaria e $\mathbf{T} = [t_{ij}]$ triangolare superiore, allora $\mathbf{T}$ è diagonale e $|t_{ii}| = 1$ per ogni $i \leq n$

Esempio 6.5.38. Siano $\mathbf{A}$ e $\mathbf{B}$ matrici complesse $n \times n$ tali che $\mathbf{B}$ è invertibile e $\mathbf{AB} = \mathbf{BA}$; sia λ un autovalore di $\mathbf{A}$ con relativo autovettore $\mathbf{v}$:

- si provi che $\mathbf{Bv}$ è pure un autovettore di $\mathbf{A}$ relativo a λ
- si provi che, se la molteplicità geometrica di λ è 1, allora $\mathbf{v}$ è un autovettore di $\mathbf{B}$

Esempio 6.5.39. Sia $\mathbf{A}$ una matrice complessa $n \times n$:

- si provi che se $\mathbf{v}$ è un autovettore di $\mathbf{AA}^H$ relativo all'autovalore λ e $\mathbf{A}^H\mathbf{v} \neq \mathbf{0}$, allora $\mathbf{A}^H\mathbf{v}$ è autovettore di $\mathbf{A}^H\mathbf{A}$ relativo all'autovalore λ
- si provi che, se $\lambda \neq 0$, allora $\mathbf{A}^H\mathbf{v} \neq \mathbf{0}$

Concludere che $\mathbf{AA}^H$ ed $\mathbf{A}^H\mathbf{A}$ hanno gli stessi autovalori non nulli

Esempio 6.5.40. Sia $\mathbf{B}$ una matrice complessa 3×3 con polinomio caratteristico

$$p_{\mathbf{B}}(X) = X^3 + X^2 + X + 1$$

- si provi che $\mathbf{B}$ è invertibile
- dire per quali valori del numero complesso λ la matrice $\mathbf{B} - \lambda\mathbf{I}_3$ è invertibile

Esempio 6.5.41. Sia $\mathbf{A}$ una matrice complessa 4×4 con polinomio caratteristico

$$p_{\mathbf{A}}(X) = X^4 - X^2 - 1$$

- si provi che $\mathbf{A}$ è invertibile
- si provi che $\mathbf{A}$ è diagonalizzabile e non può essere ne hermitiana ne unitaria

Esempio 6.5.42. Sia $\mathbf{Q}$ una matrice reale ortogonale $n \times n$. Si provi che:

1. $\mathbf{A} = \mathbf{Q} + \mathbf{Q}^{-1}$ ha tutti i suoi autovalori reali
2. detto $\mathbf{v}$ un autovettore reale di $\mathbf{A}$ relativo ad un autovalore fissato λ, si provi che il sottospazio vettoriale $V = \langle \mathbf{v}, \mathbf{Qv} \rangle$ di $\mathbb{R}^n$ è Q-invariante, cioè $\mathbf{Qx}$ appartiene a V per ogni vettore $\mathbf{x} \in V$