Lezioni di Calcolo Numerico
Lezione 12: Autovalori e valori singolari di matrici

Alberto Tibaldi

8 giugno 2018

Indice

3 Decomposizione ai valori singolari

Moltissimi dei concetti introdotti precedentemente hanno senso solamente quando si ha a che fare con matrici quadrate: si pensi per esempio agli autovalori, al determinante o all'inversa di una matrice. Tuttavia, alcuni di questi concetti possono essere in qualche modo estesi anche a matrici rettangolari $\underline{\underline{A}} \in \mathbb{R}^{m,n}$ aventi m righe e n colonne. Al fine di comprendere queste estensioni, cercheremo in qualche misura di appoggiarci al concetto di autovalore e sfruttando le proprietà di particolari matrici simmetriche definite positive. Prima di tutto, quindi arriveremo a *costruire* lo strumento matematico dietro a queste estensioni, ovvero la **decomposizione ai valori singolari**[1]; poi, proporremo una tecnica, per quanto rudimentale, in grado di **calcolarla**. Infine, studieremo le sue **applicazioni**.

Prima di proseguire con questo programma è doveroso citare la fonte che ha ispirato questa lezione: il magistrale corso *Linear Algebra* tenuto nel 2010 dal prof. Gilbert Strang del Massachussets Institute of Technology[2]. Per dare un'idea di che personaggio sia, il prof. Strang è il primo *MathWorks® Professor of Mathematics*, dove, sì, MathWorks® è proprio **quella** MathWorks®: quella che produce MATLAB®.

3.1 Costruzione della decomposizione ai valori singolari

Data una generica matrice $\underline{\underline{A}} \in \mathbb{R}^{m,n}$, da un punto di vista algebrico si può dire che $\underline{\underline{A}}$ sia un'applicazione lineare che trasforma vettori appartenenti al suo spazio delle righe in vettori appartenenti allo spazio delle colonne. Prendiamo per esempio la matrice

$$\underline{\underline{A}} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix}$$

e il vettore $\underline{x} = [0 \quad -1 \quad 1]^{\mathrm{T}}$. Si può vedere chiaramente che questo appartiene allo spazio delle righe, secondo questa definizione; infatti, il loro prodotto

$$\underline{\underline{A}}\,\underline{x} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} = 0 \times \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} - 1 \times \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + 1 \times \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

[1] per gli amici SVD, acronimo dell'inglese **singular value decomposition**

[2] il materiale, comprensivo di videoregistrazioni, è disponibile all'indirizzo `http://ocw.mit.edu/courses/mathematics/18-06-linear-algebra-spring-2010`

appartiene allo spazio delle colonne[3] di $\underline{\underline{A}}$: è un elemento dell'insieme delle possibili combinazioni lineari delle colonne di $\underline{\underline{A}}$. Questa cosa è particolarmente evidente in questo caso, visto che $\underline{\underline{A}}\,\underline{x}$ è non solo una combinazione lineare, ma addirittura a una delle colonne della matrice ;-). In quest'ottica, il prodotto righe per colonne si può considerare come la scelta di una particolare combinazione lineare dei vettori colonna, dove i coefficienti sono le componenti del vettore $\underline{x}$. Riassumendo, $\underline{x}$ appartiene allo spazio delle righe[4] perché la combinazione lineare che genera al momento di effettuare il prodotto matriciale non è il vettore nullo.

D'altro canto, tutti quei vettori $\underline{x}$ i cui elementi rappresentano i coefficienti delle combinazioni lineari non banali che portano a ottenere il vettore nullo, saranno elementi appartenenti al nucleo di $\underline{\underline{A}}$. Se non è zuppa, è pan bagnato: un vettore $\underline{x} \in \mathbb{R}^n$, in relazione alle righe di una matrice $\underline{\underline{A}} \in \mathbb{R}^{m,n}$, può appartenere o al suo spazio delle righe, o al suo nucleo. Quindi: se moltiplichiamo per esempio la nostra $\underline{\underline{A}}$ per il vettore $\underline{y} = [1 \;\; 1 \;\; -1]^{\mathrm{T}}$, otteniamo

$$\underline{\underline{A}}\,\underline{y} = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 3 \\ 3 & 1 & 4 \\ 4 & 1 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = 1 \times \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + 1 \times \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - 1 \times \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

e quindi scopriamo che $\underline{y}$ appartiene al kernel di $\underline{\underline{A}}$, in quanto la moltiplicazione ha mappato $\underline{y}$ nel vettore nullo, nonostante $\underline{y}$ non fosse identicamente nullo.

Riassumendo: una matrice $\underline{\underline{A}} \in \mathbb{R}^{m,n}$ si può vedere come costituita da m vettori riga messi uno sopra l'altro. Le combinazioni lineari di questi vettori riga costituiscono lo spazio delle righe, il quale a sua volta è un sottospazio di $\mathbb{R}^n$, dal momento che ciascuna riga è un vettore dotato di n componenti. Ma quindi, a partire da questi vettori, è certamente lecito applicare l'algoritmo di Gram-Schmidt al fine di ottenere una base ortonormale dello spazio delle righe; questa sarà costituita da $r \leq m$ elementi, che chiamiamo $\{\underline{v}_i\}$, per $i = 1, 2, \ldots, r$. Qui, r è il rango della matrice $\underline{\underline{A}}$: come noto dall'Algebra Lineare; infatti, il rango coincide con la dimensione dello spazio delle righe, ossia con il minimo numero di vettori linearmente indipendenti le cui combinazioni lineari lo possono generare. Il nucleo di $\underline{\underline{A}}$, anche detto *kernel*, è il complementare a $\mathbb{R}^n$ di questo spazio, e avrà quindi dimensione pari a $n - r$.

Abbiamo parlato tanto delle righe, ma adesso concentriamoci per un momento sulle colonne: se io moltiplico la matrice $\underline{\underline{A}}$ per ciascuno dei vettori della base ortonormale appena introdotta, ottengo, per ciascuno di essi, un vettore $\underline{\underline{A}}\,\underline{v}_i$ che certamente apparterrà allo spazio delle colonne. Tuttavia, da un punto di vista geometrico, moltiplicare la matrice $\underline{\underline{A}}$ per un generico vettore $\underline{v}_i$ provoca una rotazione e/o una dilatazione di quest'ultimo; di conseguenza, nonostante i nostri vettori $\{\underline{v}_i\}$ siano tra loro ortogonali per ipotesi/costruzione, non abbiamo alcuna garanzia in merito al fatto che gli $\{\underline{\underline{A}}\,\underline{v}_i\}$ lo siano: la matrice $\underline{\underline{A}}$ può ruotare e/o dilatare in modo diverso ciascun vettore[5]. Fatta questa introduzione, ora ci poniamo il seguente ambizioso proposito: vogliamo trovare una base dello spazio delle righe $\{\underline{v}_i\}$ tale per cui la applicazione di $\underline{\underline{A}}$ su diversi vettori $\underline{v}_i$ generi vettori colonna $\underline{u}_i$ tra loro ortogonali. Ossia, vogliamo una particolare base ortonormale dello spazio delle righe in cui ciascun elemento si mappa in un elemento di una base ortonormale dello spazio delle colonne! Quindi, in matematichese, vogliamo una cosa del tipo

$$\begin{aligned} \underline{\underline{A}} \left[\underline{v}_1 \quad \underline{v}_2 \quad \underline{v}_3 \quad \cdots \quad \underline{v}_r \right] &= \left[\underline{\underline{A}}\,\underline{v}_1 \quad \underline{\underline{A}}\,\underline{v}_2 \quad \underline{\underline{A}}\,\underline{v}_3 \quad \cdots \quad \underline{\underline{A}}\,\underline{v}_r \right] = \\ &= \left[s_1 \underline{u}_1 \quad s_2 \underline{u}_2 \quad s_3 \underline{u}_3 \quad \cdots \quad s_r \underline{u}_r \right]. \end{aligned} \tag{1}$$

L'equazione matriciale (1) si può scrivere più compattamente come

$$\underline{\underline{A}}\,\underline{\underline{V}}^{(r)} = \underline{\underline{U}}^{(r)} \underline{\underline{S}}^{(r)}, \tag{2}$$

a patto di definire le tre matrici

$$\underline{\underline{V}}^{(r)} = \left[\underline{v}_1 \quad \underline{v}_2 \quad \underline{v}_3 \quad \cdots \quad \underline{v}_r \right],$$

[3] che si può vedere un po' come l'immagine della applicazione lineare associata alla matrice!

[4] che quindi si può un po' vedere come la controimmagine della applicazione lineare associata alla matrice!

[5] si per esempio pensi a un vettore generico, o a un autovettore: un vettore generico subirà una certa rotazione, l'autovettore non ne subirà alcuna!

$$\underline{\underline{U}}^{(r)} = \begin{bmatrix} \underline{u}_1 & \underline{u}_2 & \underline{u}_3 & \cdots & \underline{u}_r \end{bmatrix},$$

e $\underline{\underline{S}}^{(r)}$ come una matrice diagonale $r \times r$ avente come elementi gli s_i. L'equazione (2) ricorda vagamente le espressioni legate alla decomposizione agli autovalori, anche alla luce del fatto che $\underline{\underline{S}}^{(r)}$ è una matrice diagonale; tuttavia, qui la matrice $\underline{\underline{A}}$ è rettangolare, e le matrici $\underline{\underline{V}}^{(r)}$ e $\underline{\underline{U}}^{(r)}$ sono diverse tra loro, quindi sembra che si stia ottenendo una qualche generalizzazione o estensione del concetto di diagonalizzazione di una matrice. Le matrici $\underline{\underline{V}}^{(r)}$ e $\underline{\underline{U}}^{(r)}$, per costruzione, sono certamente ortogonali: le loro r colonne sono vettori per ipotesi ortogonali! Inoltre, dal momento che ciascuna delle colonne di $\underline{\underline{V}}^{(r)}$ deriva dall'applicazione di Gram-Schmidt sui vettori riga di $\underline{\underline{A}}$, allora, sempre per costruzione, $\underline{\underline{V}}^{(r)}$ ha n righe; dualmente, $\underline{\underline{U}}^{(r)}$ avrà m righe. Quindi, $\underline{\underline{V}}^{(r)} \in \mathbb{R}^{n,r}$, mentre $\underline{\underline{U}} \in \mathbb{R}^{m,r}$.

Possiamo fare di più, ed estendere ulteriormente queste definizioni arrivando a ottenere matrici $\underline{\underline{V}}$ e $\underline{\underline{U}}$ quadrate, di dimensione $n \times n$ e $m \times m$, rispettivamente. Per fare questo, è possibile aggiungere a $\underline{\underline{V}}^{(r)}$ e $\underline{\underline{U}}^{(r)}$ nuove colonne, ortogonali alle precedenti, ma quindi facenti parte dei complementi ortogonali a $\mathbb{R}^n$ e $\mathbb{R}^m$ degli spazio delle righe e degli spazi delle colonne, rispettivamente. Per capire meglio questa cosa, pensiamo allo spazio delle righe e a $\mathbb{R}^n$: se la matrice $\underline{\underline{A}}$ ha rango r non massimo, significa che lo spazio delle righe è un sottoinsieme strettamente incluso di $\mathbb{R}^n$, avente dimensione r. Quindi, in aggiunta ai r vettori $\underline{v}_i$, $i = 1, \ldots, r$ definiti come la base ortonormale dello spazio delle righe, è possibile definire altri $n - r$ vettori ortogonali tra di loro e ai precedenti, che però genereranno il complementare dello spazio delle righe: lo spazio nullo della matrice! Si tratterà quindi di vettori aggiuntivi $\underline{v}_i$ tali per cui

$$\underline{\underline{A}}\,\underline{v}_i = \underline{0}, \qquad i = r+1, \ldots, n.$$

Con questa idea applicata su $\underline{\underline{V}}^{(r)}$ e dualmente su $\underline{\underline{U}}^{(r)}$, la (1) può essere estesa includendo, come anticipato, anche i vettori appartenenti ai complementi a $\mathbb{R}^n$ e $\mathbb{R}^m$, arrivando a scrivere

$$\underline{\underline{A}}\underbrace{\begin{bmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_r & \underline{v}_{r+1} & \cdots & \underline{v}_n \end{bmatrix}}_{\underline{\underline{V}}} = \underbrace{\begin{bmatrix} s_1\underline{u}_1 & s_2\underline{u}_2 & \cdots & s_r\underline{u}_r & s_{r+1}\underline{u}_{r+1} & \cdots & s_m\underline{u}_m \end{bmatrix}}_{\underline{\underline{U}}\,\underline{\underline{S}}}. \tag{3}$$

Dal momento che

$$\underline{\underline{A}}\,\underline{v}_{r+1} = \underline{\underline{A}}\,\underline{v}_{r+2} = \ldots = \underline{\underline{A}}\,\underline{v}_n = s_{r+1}\underline{u}_{r+1} + s_{r+2}\underline{u}_{r+2} + \ldots + s_m\underline{u}_m = \underline{0},$$

e poiché non siamo interessati a soluzioni banali[6], vogliamo che tutti i $\underline{u}_i$ siano diversi dal vettore identicamente nullo, quindi a essere nulli saranno necessariamente i coefficienti s_i:

$$s_i = 0, \quad i > r.$$

Riassumiamo il contenuto di (3).

- Indicati in blu troviamo gli elementi della base ortogonale dello spazio delle righe della matrice $\underline{\underline{A}}$.
- Indicati in rosso troviamo gli elementi della base ortogonale del complemento a $\mathbb{R}^n$ dello spazio delle righe della matrice $\underline{\underline{A}}$; di conseguenza, questa sarà una base ortogonale del nucleo della matrice $\underline{\underline{A}}$.
- Indicati in ciano troviamo gli elementi della base ortogonale dello spazio delle colonne, ovvero dell'immagine, della matrice $\underline{\underline{A}}$.
- Indicati in magenta troviamo gli elementi della base ortogonale del complemento a $\mathbb{R}^n$ dello spazio delle colonne della matrice $\underline{\underline{A}}$.

[6] dove per *banali* intendiamo combinazioni lineari che vanno a zero perché costruite con vettori nulli

La (3) si può scrivere in modo sintetico come

$$\underline{\underline{A}}\,\underline{\underline{V}} = \underline{\underline{U}}\,\underline{\underline{S}}, \tag{4}$$

dove $\underline{\underline{V}} \in \mathbb{R}^{n,n}$, $\underline{\underline{U}} \in \mathbb{R}^{m,m}$, e $\underline{\underline{S}} \in \mathbb{R}^{m,n}$ è una matrice rettangolare che ha elementi diversi da zero solo sulla diagonale della matrice quadrata di dimensione massima che si può ritagliare da essa. Gli elementi di $\underline{\underline{S}}$ vengono chiamati **valori singolari**, le colonne di $\underline{\underline{U}}$ vengono dette **vettori singolari sinistri**, e le colonne di $\underline{\underline{V}}$ vengono dette **vettori singolari destri**. L'espressione (4) si può ancora elaborare moltiplicando ambo i membri per $\underline{\underline{V}}$, arrivando a una forma *tipo panino*

$$\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{-1}$$

ma, sapendo che $\underline{\underline{V}}$ è una matrice ortogonale per costruzione, $\underline{\underline{V}}^{-1} = \underline{\underline{V}}^{\mathrm{T}}$ e, quindi, si arriva alla forma che si trova usualmente in letteratura:

$$\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{\mathrm{T}}. \tag{5}$$

Questa è la cosiddetta **decomposizione ai valori singolari**, o **singular value decomposition** (SVD), della matrice $\underline{\underline{A}}$.

3.2 Calcolo della decomposizione ai valori singolari

Fino a questo momento siamo stati molto astratti: abbiamo **costruito** la decomposizione ai valori singolari di una matrice, ovvero, abbiamo *detto come ci piacerebbe che fosse fatta*: abbiamo capito che è una specie di decomposizione dello spazio delle righe e dello spazio delle colonne della matrice, ma non abbiamo alcuna idea di come, operativamente, si possano calcolare le matrici $\underline{\underline{U}}$, $\underline{\underline{S}}$ e $\underline{\underline{V}}$.

Per calcolarle, cercheremo di sfruttare qualcosa che conosciamo già molto bene; infatti, ripartendo da (5), proviamo a studiare il prodotto $\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}}$:

$$\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}} = \left(\underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{\mathrm{T}}\right)^{\mathrm{T}}\left(\underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{\mathrm{T}}\right) = \underline{\underline{V}}\,\underline{\underline{S}}^{\mathrm{T}}\underline{\underline{U}}^{\mathrm{T}}\underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{\mathrm{T}}.$$

Dal momento che la matrice $\underline{\underline{U}}$ è ortogonale, questa espressione si può semplificare; definendo

$$\underline{\underline{S}}^{\mathrm{T}}\underline{\underline{S}} = \underline{\underline{S}}^2,$$

vediamo che essa sarà diagonale, avrà dimensioni $n \times n$, e i suoi elementi saranno i quadrati dei primi n valori singolari. Dunque, abbiamo

$$\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}} = \underline{\underline{V}}\,\underline{\underline{S}}^2\underline{\underline{V}}^{\mathrm{T}}.$$

Dal momento che la matrice $\underline{\underline{V}}$ è per ipotesi ortogonale, questa espressione si può riscrivere ulteriormente come

$$\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}} = \underline{\underline{V}}\,\underline{\underline{S}}^2\underline{\underline{V}}^{-1},$$

che è **esattamente** la decomposizione agli autovalori della matrice $\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}}$. Riassumendo, siamo partiti da una matrice $\underline{\underline{A}}$ che può essere rettangolare, poi considerando $\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}}$, ne abbiamo ottenuta una quadrata e, mettendo in relazione quest'ultima con la decomposizione ai valori singolari della matrice $\underline{\underline{A}}$ (5), abbiamo scoperto che possiamo operativamente calcolare i valori singolari come le radici quadrate degli autovalori di $\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}}$, e la matrice $\underline{\underline{V}}$ come la corrispondente matrice degli autovettori.

Come si può immaginare, se calcoliamo $\underline{\underline{A}}\,\underline{\underline{A}}^{\mathrm{T}}$, otteniamo

$$\underline{\underline{A}}\,\underline{\underline{A}}^{\mathrm{T}} = \left(\underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{\mathrm{T}}\right)\left(\underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{\mathrm{T}}\right)^{\mathrm{T}} = \underline{\underline{U}}\,\underline{\underline{S}}\,\underline{\underline{V}}^{\mathrm{T}}\underline{\underline{V}}\,\underline{\underline{S}}^{\mathrm{T}}\underline{\underline{U}}^{\mathrm{T}} = \underline{\underline{U}}\underline{\underline{S}}^2\underline{\underline{U}}^{\mathrm{T}},$$

dove in questo caso $\underline{\underline{S}}^2$ ha dimensioni $m \times m$. Nel caso avessimo $\underline{\underline{A}} \in \mathbb{R}^{m,n}$ con $m > n$ (più righe che colonne), che gli ultimi $m - n$ elementi diagonali di $\underline{\underline{S}}^2$ ottenuta da $\underline{\underline{A}}\,\underline{\underline{A}}^{\mathrm{T}}$ sarebbero uguali a zero, e viceversa per il caso $n > m$ e la $\underline{\underline{S}}^2$ ottenuta da $\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}}$. Le colonne di $\underline{\underline{U}}$ sono dunque gli autovettori della matrice $\underline{\underline{A}}\,\underline{\underline{A}}^{\mathrm{T}}$, e, proprio come prima, i valori singolari sono le radici quadrate dei suoi autovalori.

3.2.1 Esempio di calcolo della decomposizione ai valori singolari

Una volta trovato un collegamento tra la SVD di una matrice e qualcosa che sappiamo gestire, ovvero gli autovalori, può essere utile svolgere un esercizietto che ci permetta di fissare ulteriormente i concetti. Dunque, calcoliamo la decomposizione ai valori singolari della matrice

$$\underline{\underline{A}} = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix},$$

e discutiamo i risultati.

Prima di tutto, partiamo dal calcolo della matrice $\underline{\underline{V}}$; per ottenerla, usiamo la procedura appena mostrata:

$$\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}} \triangleq \underline{\underline{B}} = \begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix}\begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix} = \begin{bmatrix} 4 \times 4 + 8 \times 8 & 4 \times 3 + 8 \times 6 \\ 3 \times 4 + 6 \times 8 & 3 \times 3 + 6 \times 6 \end{bmatrix} = \begin{bmatrix} 80 & 60 \\ 60 & 45 \end{bmatrix}.$$

«Oh! Ma che sorpresa! È simmetrica! È un caso?»

Ma no: è **sempre** vero che $\underline{\underline{A}}^{\mathrm{T}}\underline{\underline{A}}$ è simmetrica (e definita positiva)!

A parte questo: abbiamo visto che i valori singolari sono uguali alle radici quadrate degli autovalori di questa matrice; per una matrice 2×2, l'equazione caratteristica si può scrivere nella forma traccia-determinante:

$$s^4 - \mathrm{Tr}\{\underline{\underline{B}}\}s^2 + \det\{\underline{\underline{B}}\} = 0,$$

che è biquadratica perché andiamo a cercare direttamente i valori singolari s, piuttosto che gli autovalori λ di $\underline{\underline{B}}$; allora, facendo i conticini,

$$\mathrm{Tr}\{\underline{\underline{B}}\} = 80 + 45 = 125,$$

e

$$\det\{\underline{\underline{B}}\} = 80 \times 45 - 60 \times 60 = 0.$$

L'equazione caratteristica quindi è

$$s^4 - 125s^2 = 0,$$

che ha come radici

$$s_1 = \sqrt{125}$$
$$s_2 = 0.$$

È usanza e convenzione diffusa **e da tener presente** ordinare i valori singolari dal più grande al più piccolo. Abbiamo un valore singolare nullo; siamo sorpresi? Beh, no: se avessimo calcolato il determinante della matrice, ci saremmo accorti subito che $\underline{\underline{A}}$ non è di rango massimo; in effetti, la seconda riga è un multiplo della prima ;-).

Calcoliamo ora, con un po' di furbizia, i vettori singolari destri $\underline{v}_i$. A questo fine, partiamo da $\underline{v}_2$, ossia dal vettore singolare associato al valore $s_2 = 0$; questo richiede di risolvere il sistema omogeneo

$$\begin{bmatrix} 80 & 60 \\ 60 & 45 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Poiché una riga o l'altra sono la stessa cosa, possiamo vedere, dalla prima, che

$$80x_1 + 60x_2 = 0,$$

che ci dice che

$$x_1 = -\frac{6}{8}x_2.$$

Quindi, prendiamo per esempio il vettore

$$\underline{v}_2 = \begin{bmatrix} -\frac{6}{8} \\ 1 \end{bmatrix},$$

e normalizziamolo[7], ottenendo:

$$||\underline{v}_2||_2 = \sqrt{\frac{36}{64} + 1} = \sqrt{\frac{100}{64}} = \frac{10}{8},$$

e, infine,

$$\frac{\underline{v}_2}{||\underline{v}_2||_2} = \begin{bmatrix} -\frac{6}{10} \\ \frac{8}{10} \end{bmatrix} \triangleq \underline{v}_2$$

(ridefiniamo $\underline{v}_2$ come quello normalizzato). **Dal momento che $\underline{v}_2$ è associato al valore singolare $s_2 = 0$, questo vettore singolare destro costituisce una base dello spazio nullo di $\underline{\underline{A}}$.** Infatti, se noi calcolassimo

$$\underline{\underline{A}} \sum_{i=r+1}^{n} \alpha_i \underline{v}_i = \sum_{i=r+1}^{n} \alpha_i \underline{\underline{A}}\, \underline{v}_i = \sum_{i=r+1}^{m} \alpha_i s_i \underline{u}_i,$$

dove r è il rango della matrice, otterremmo sempre 0! Si noti che gli estremi delle sommatorie sono stati scambiati da n a m, perché i valori singolari dopo il r-esimo, come si è detto prima, sono uguali a zero ma le matrici possono avere dimensione diversa.

Questa è una delle applicazioni più interessanti della SVD: determinare la base dello spazio nullo di una matrice! Questa cosa dimostra semplicemente che la nostra costruzione della precedente sezione è andata a buon fine! ;-)

Dal momento che stiamo ragionando su $\underline{\underline{B}} = \underline{\underline{A}}^{\mathrm{T}} \underline{\underline{A}}$, che è simmetrica, i suoi autovettori saranno ortogonali; quindi, invece di rifare tutti questi brutti conti per trovare $\underline{v}_1$, sarà sufficiente dire che $\underline{v}_1$ è l'unico vettore ortogonale a $\underline{v}_2$ di norma pari a 1 e, quindi, per ispezione,

$$\underline{v}_1 = \begin{bmatrix} \frac{8}{10} \\ \frac{6}{10} \end{bmatrix}.$$

Cerchiamo a questo punto di calcolare i vettori singolari sinistri di $\underline{\underline{A}}$. Partiamo quindi dalla matrice

$$\underline{\underline{A}}\,\underline{\underline{A}}^{\mathrm{T}} \triangleq \underline{\underline{C}} = \begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix} \begin{bmatrix} 4 & 8 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 25 & 50 \\ 50 & 100 \end{bmatrix}.$$

A questo punto, se procedessimo beceramente, dovremmo calcolarci gli autovalori. Ma visto che noi siamo giovani astuti e svelti, possiamo ricordarci che gli autovalori sono gli stessi che abbiamo calcolato prima: $s_1^2 = 125$, $s_2^2 = 0$; purtroppo, gli autovettori invece saranno diversi. Interessiamoci per ora solo all'autovettore associato a s_1, e calcoliamolo impostando il sistema lineare omogeneo associato:

$$\begin{bmatrix} 25 - s_1^2 & 50 \\ 50 & 100 - s_1^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -100 & 50 \\ 50 & -25 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Prendendo per esempio la seconda equazione,

$$50x_1 - 25x_2 = 0 \Longrightarrow x_2 = 2x_1,$$

[7] si noti che MATLAB®normalizza sia gli autovettori sia i vettori singolari rispetto alla norma 2!

e, quindi, possiamo definire il vettore singolare sinistro $\underline{u}_1$ come

$$\underline{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Normalizziamolo:

$$\frac{\underline{u}_1}{\|\underline{u}_1\|_2} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \triangleq \underline{u}_1,$$

dove quindi abbiamo ridefinito $\underline{u}_1$ come quello normalizzato.

Al fine di comprendere il significato di un vettore singolare sinistro, che ricordiamo essere, per come abbiamo costruito la SVD, un elemento della base dello spazio delle colonne, proviamo calcolare, per un certo numero di vettori $\underline{x}$ presi a caso,

$$\underline{y} = \underline{\underline{A}}\,\underline{x}.$$

Possiamo per esempio prendere $\underline{x} = [1 \;\; 2]^{\mathrm{T}}$:

$$\begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 10 \\ 20 \end{bmatrix}.$$

Oppure, possiamo provare con $\underline{x} = [4 \;\; 1]^{\mathrm{T}}$:

$$\begin{bmatrix} 4 & 3 \\ 8 & 6 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 19 \\ 38 \end{bmatrix}.$$

> «Ma, ma, porca miseria! È sempre un multiplo di $\underline{u}_1$! Ma perché?»

Beh, perché la matrice è 2×2 lo spazio delle colonne, ovvero l'immagine, ha dimensione 1 (perché solo un valore singolare è non nullo) e, quindi, il generico vettore $\underline{x}$ si può scrivere come

$$\underline{x} = \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2,$$

ma, dal momento che $\underline{v}_2 \in \ker\{\underline{\underline{A}}\}$ per quanto abbiamo detto precedentemente, si ha che

$$\underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}\left(\alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2\right) = \alpha_1 \underline{\underline{A}}\,\underline{v}_1 + \alpha_2 \underline{\underline{A}}\,\underline{v}_2 = \alpha_1 \underline{\underline{A}}\,\underline{v}_1 = \alpha_1 s_1 \underline{u}_1,$$

che guarda caso è quello che ci viene numericamente. Questo ci dimostra che, proprio come volevamo dalla nostra costruzione, i $\{\underline{u}_i\}$, per $i = 1, \ldots, r$, costituiscono una base dell'immagine: moltiplicando la matrice $\underline{\underline{A}}$ per un generico vettore $\underline{x}$, il risultato sarà sempre una combinazione lineare dei vettori singolari sinistri associati a valori singolari non nulli!

Questo, è la SVD: una decomposizione in grado di fornire da un lato una base dell'immagine di un'applicazione lineare, ossia del suo spazio delle colonne, e dall'altro una base del suo nucleo. Per di più, grazie alla nostra costruzione, queste basi sono pure ortonormali, che è il meglio che si possa avere.

Prima di concludere, una domanda di carattere *numerico*.

> «Il calcolo dei valori singolari è affetto da problemi di condizionamento?»

E, signore e signori, la risposta è **no**! Un'intuizione dietro questa buonissima notizia deriva dall'interpretazione dei valori e vettori singolari come autovalori e autovettori di matrici simmetriche che, certamente, sono molto ben condizionati!

Si noti che $\underline{\underline{U}}$ e $\underline{\underline{V}}$ non sono definite univocamente, dal momento che dipendono dalle scelte che effettuiamo come basi degli spazi delle righe e delle colonne; tuttavia, questo non introduce particolari problemi, da un punto di vista dell'accuratezza.

MATLAB®permette di calcolare la decomposizione ai valori singolari mediante il comando `svd`, usato come segue:

```
[U,S,V] = svd(A);
```

il quale ritorna le matrici $\underline{\underline{U}}$, $\underline{\underline{S}}$, e $\underline{\underline{V}}$.

Queste note non volevano spiegare come MATLAB®calcoli la SVD, ma esclusivamente *introdurre per via costruttiva le sue proprietà*. Non è detto che MATLAB®calcoli la SVD risolvendo due problemi agli autovalori; quel metodo, tuttavia, è fondamentale al fine di fornire interpretazioni forti dei vari attori coinvolti.