

Appunti di algebra numerica

Daniele Zago

8 febbraio 2020

Capitolo 1

Fondamentali

Lezione 1: Operazioni e quantità fondamentali

2020-02-05

Sia A una matrice $m \times n$, la trasformazione $x \mapsto Ax$ è **lineare**, ovvero per ogni $x, y \in \mathbb{C}^n$ e per ogni $\alpha \in \mathbb{C}$,

$$A(x + y) = Ax + Ay,$$

$$A(\alpha x) = \alpha Ax.$$

In particolare, una qualunque trasformazione lineare $\mathbb{C}^n \rightarrow \mathbb{C}^m$ si può scrivere in termini di prodotto per una matrice $A_{m \times n}$.

Il prodotto matrice per vettore si può rappresentare nel modo seguente

$$\begin{bmatrix} \left| \begin{array}{c} a_1 \end{array} \right| & \left| \begin{array}{c} a_2 \end{array} \right| & \cdots & \left| \begin{array}{c} a_n \end{array} \right| \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} a_1 \end{bmatrix} + x_2 \begin{bmatrix} a_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_n \end{bmatrix}$$

Figura 1: Prodotto matrice per vettore.

ovvero Ax è una combinazione lineare delle colonne di A . È utile pensare al prodotto Ax come l'azione del vettore x su A , e non il viceversa.

Esempio: Matrice di Vandermonde

Se p, q sono polinomi di grado minore di n , allora soddisfano le seguenti proprietà

$$(p + q)(x) = p(x) + q(x)$$

$$(\alpha p)(x) = \alpha p(x)$$

dunque la mappa che trasforma i coefficienti di un polinomio nel vettore $(p(x_1), p(x_2), \dots, p(x_m))$ è

lineare e rappresentabile con la *matrice di Vandermonde*

$$p(x) = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^{n-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix}.$$

Dunque la matrice

$$A = \begin{pmatrix} 1 & x & x^2 & \dots & x^{n-1} \end{pmatrix}$$

si può pensare come la matrice che dà la combinazione lineare dei monomi.

Lo **span** di una matrice A , indicato $\text{span } A$, è l'insieme dei vettori che si possono scrivere come Ax per qualche x .

Teo. (Span di A) $\text{span } A$ è lo spazio generato dalle colonne di A .

Dim.

Ax è una combinazione lineare delle colonne di A , per cui se $y = Ax$ vuol dire che y si può scrivere come combinazione lineare delle colonne di A , ovvero $y \in \langle a_1, a_2, \dots, a_n \rangle$. \square

Il **rango** di A è la dimensione dello spazio generato dalle colonne di A , che è anche uguale allo spazio generato dalle righe di A .

Una matrice $m \times n$ è di **rango pieno** se il rango è pari a $\min\{m, n\}$.

Teo. (Matrice a rango pieno) Una matrice è a rango pieno se la mappa che definisce è iniettiva, ovvero se non mappa due vettori diversi nello stesso vettore.

Una **matrice invertibile** è una matrice quadrata di rango pieno, che ammette una matrice inversa A^{-1} tale che

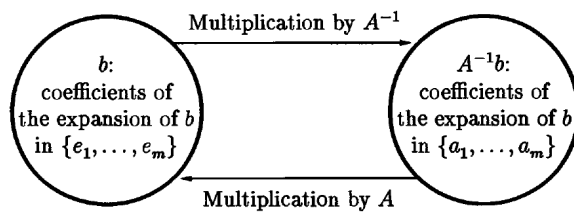
$$A^{-1}A = I.$$

Teo. Sono equivalenti le seguenti condizioni per $A \in \mathbb{C}^{m \times m}$

1. A ammette inversa A^{-1} .
2. $\text{rank } A = m$.
3. $\text{span } A = \mathbb{C}^m$.
4. $\ker A = \{0\}$.
5. 0 non è un autovalore di A .
6. 0 non è un valore singolare di A .
7. $\det A \neq 0$.

La moltiplicazione per A^{-1} è un'operazione di cambio di base, nel senso che $A^{-1}b$ ritorna l'unico vettore x tale che $Ax = b$. In questo senso,

$A^{-1}b$ è il vettore dei coefficienti di b nella base delle colonne di A .



Lezione 2: Vettori e matrici ortogonali

2020-02-05

Data una matrice $A_{m \times n} = (a_{ij})$, si indica con A^* la **coniugata hermitiana** di A data da

$$A^* = \bar{a}_{ji} \quad \text{matrice } n \times m.$$

Se $A = A^*$, la matrice è **hermitiana** e per definizione deve essere quadrata. Se A è reale, la coniugata hermitiana si dice matrice **trasposta** A^T e se $A = A^T$ si dice che A è **simmetrica**.

Il **prodotto scalare** di due vettori $x, y \in \mathbb{C}^m$ è definito come

$$x^* y = \sum_{i=1}^m \bar{x}_i y_i,$$

che induce la **norma euclidea standard**

$$\|x\| = \sqrt{x^* x} = \left(\sum_{i=1}^m |x_i|^2 \right)^{1/2}.$$

Il coseno dell'angolo tra x e y si può calcolare attraverso il prodotto scalare come

$$\cos \vartheta = \frac{x^* y}{\|x\| \cdot \|y\|}.$$

Il prodotto scalare è una forma **bilineare**, in quanto

$$\begin{aligned} (x_1 + x_2)^* y &= x_1^* y + x_2^* y \\ x^* (y_1 + y_2) &= x^* y_1 + x^* y_2 \\ (\alpha x)^* (\beta y) &= \bar{\alpha} \beta x^* y \end{aligned}$$

in particolare, valgono le seguenti proprietà per le matrici

$$\begin{aligned} (AB)^* &= B^* A^* \\ (AB)^{-1} &= B^{-1} A^{-1} \end{aligned}$$

Vettori ortogonali

Due vettori sono **ortogonali** se giacciono ad angolo retto, ovvero se $x^* y = 0$. Si dice che un insieme $\{x_1, x_2, \dots, x_n\}$ è di vettori **ortonormali** se

$$x_i^* x_j = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

Teo. (Indipendenza lineare di vettori ortogonali) *I vettori in un insieme S di vettori ortogonali sono linearmente indipendenti.*

Corollario *Se un insieme $S \in \mathbb{C}^m$ di vettori ortogonali contiene m vettori, allora è una base per \mathbb{C}^m .*

Matrici unitarie

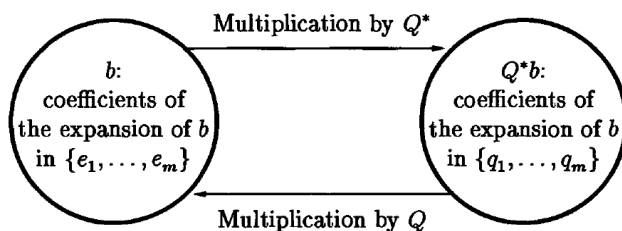
Una matrice $Q_{m \times m}$ è **unitaria** se $Q^* = Q^{-1}$, se è reale si dice che è **ortogonale**. In particolare,

$$Q^*Q = I.$$

In termini delle colonne di Q , vale che $q_i^* q_j = \delta_{ij}$.

Riprendendo quanto detto per il cambio di base di un vettore,

Q^*b è il vettore dei coefficienti di b nella base delle colonne di Q



Questa operazione preserva la geometria euclidea dello spazio, in quanto i prodotti scalari si conservano

$$(Qx)^*(Qy) = x^*(Q^*Q)y = x^*y,$$

in particolare una moltiplicazione per una matrice ortogonale Q corrisponde a

Rotazione rigida $\det Q = 1$

Riflessione $\det Q = -1$

Lezione 3: Norme

2020-02-05

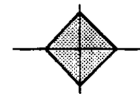
Le norme sono utili per descrivere lunghezze e distanze negli spazi vettoriali, siano essi spazi a dimensione finita o infinita.

Una **norma** è una funzione $\|\cdot\| : \mathbb{C}^m \rightarrow \mathbb{R}$ tale che, se $x, y \in \mathbb{C}^m$ e $\alpha \in \mathbb{C}$,

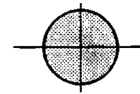
1. $\|x\| \geq 0$ e $\|x\| = 0 \iff x = 0$,
2. $\|\alpha x\| = |\alpha| \|x\|$,
3. $\|x + y\| \leq \|x\| + \|y\|$.

Le norme più importanti sono le cosiddette **p-norme** definite sotto, con rappresentata di fianco la palla di raggio 1 corrispondente $\{x \in \mathbb{C}^m : \|x\| \leq 1\}$.

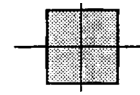
$$\|x\|_1 = \sum_{i=1}^m |x_i|$$



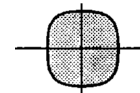
$$\|x\|_2 = \left(\sum_{i=1}^m |x_i|^2 \right)^{1/2}$$



$$\|x\|_\infty = \max_{1 \leq i \leq m} |x_i|$$



$$\|x\|_p = \left(\sum_{i=1}^m |x_i|^p \right)^{1/p} \quad 1 \leq p < \infty$$

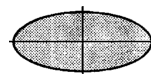


Altre norme interessanti sono le **p-norme pesate**, in cui ogni componente ha un suo peso. Data una norma $\|\cdot\|$, si può scrivere la norma pesata da una matrice W come

$$\|x\|_\Sigma = \|\Sigma x\|,$$

che è una norma per qualunque matrice W non singolare. Se Σ è una matrice diagonale, ad esempio, si può avere la 2-norma pesata

$$\|x\|_\Sigma = \left(\sum_{i=1}^m |\sigma_i x_i|^2 \right)^{1/2}$$



Norme matriciali indotte da norme vettoriali

Una matrice $A_{m \times n}$ si può pensare come un vettore in uno spazio mn -dimensionale, per cui qualunque norma su $\mathbb{C}^{m \cdot n}$ si può utilizzare per misurare la “dimensione” di una matrice.

Alternativamente, si possono considerare le **norme indotte**, che sono definite in base al comportamento della matrice come operatore lineare.

Def. (Norma matriciale indotta) Sia A la matrice di un'applicazione lineare $f : \mathbb{C}^n \rightarrow \mathbb{C}^m$ e siano $\|\cdot\|_{(n)}$ e $\|\cdot\|_{(m)}$ le norme sul dominio e codominio, rispettivamente.

La norma matriciale indotta $\|\cdot\|_{(m,n)}$ è definita come

$$\begin{aligned}\|A\|_{(m,n)} &:= \inf_{C \in \mathbb{R}} : \|Ax\|_{(m)} \leq C\|x\|_{(n)} \quad \text{per ogni } x \in \mathbb{C}^n \\ &= \sup_{x \in \mathbb{C}^n} \frac{\|Ax\|_{(m)}}{\|x\|_{(n)}}\end{aligned}$$

Caratterizzazione equivalente della norma matriciale indotta

Dal momento che $x \mapsto Ax$ è determinata dal comportamento di A sui vettori unitari, vale che

$$\|A\|_{(m,n)} = \sup_{\substack{x \in \mathbb{C}^n \\ \|x\|_{(n)}=1}} \|Ax\|_{(m)},$$

ovvero il valore massimo per cui A “allunga” un vettore unitario.

Esempio : Norma matriciale indotta

Si consideri la matrice

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix},$$

allora le norme matriciali indotte (stessa norma su dominio e codominio) sono ad esempio:

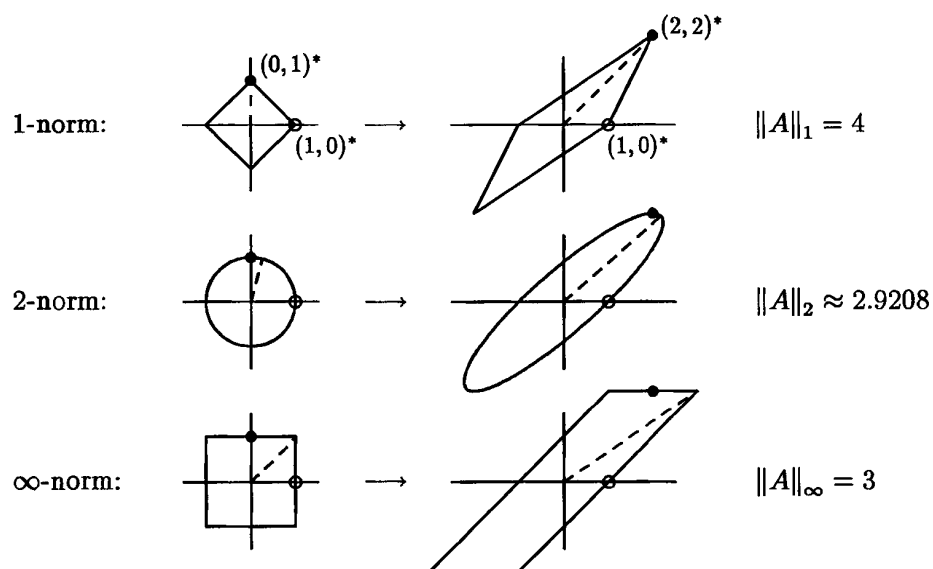


Figura 2: Esempio di norme matriciali indotte.

Esempio : p -norma di una matrice diagonale

Data una matrice diagonale

$$D = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_m \end{pmatrix},$$

allora l'immagine della palla di raggio unitario è un'ellisse, il cui asse maggiore ha lunghezza $\max_i \{|d_i|\}$. Dunque, per ogni p , vale che

$$\|D\|_p = \max_i \{|d_i|\}.$$

Esempio : 1-norma matriciale

Se $A_{m \times n}$ è una matrice generica, allora $\|A\|_1$ è la “massima somma per colonne” di A :

$$\|A\|_1 = \max_{1 \leq j \leq n} \|\vec{a}_j\|_1,$$

infatti

$$\|Ax\|_1 = \left\| \sum_{j=1}^m x_j \vec{a}_j \right\|_1 \leq \sum_{j=1}^n |x_j| \|\vec{a}_j\|_1 \leq \max_{1 \leq j \leq n} \|\vec{a}_j\|_1.$$

Esempio : Norma ∞ di una matrice

Con argomenti analoghi, si mostra che la norma ∞ di una matrice è la “massima somma per righe” di A :

$$\|A\|_\infty = \max_{1 \leq i \leq m} \|\mathbf{a}_i^*\|_1.$$

Disuguaglianze di Cauchy-Schwarz e di Hölder

1. **Disuguaglianza di Cauchy-Schwarz:** per ogni $x, y \in \mathbb{C}^m$,

$$|x^* y| \leq \|x\|_2 \|y\|_2.$$

2. **Disuguaglianza di Hölder:** siano p, q tali che $1/p + 1/q = 1$, con $1 \leq p, q \leq \infty$. Allora,

$$|x^* y| \leq \|x\|_p \|y\|_q.$$

In generale, vale anche una simile disuguaglianza per la norma del prodotto tra matrici. Se $A_{l \times m}$ e $B_{m \times n}$, allora

$$\|AB\|_{(l,n)} \leq \|A\|_{(l,m)} \|B\|_{(m,n)}.$$

In generale però questa uguaglianza NON è un'uguaglianza.

Norme matriciali generali

Altre norme matriciali si possono costruire utilizzando le tre proprietà delle distanze, applicate allo spazio vettoriale mn -dimensionale delle matrici:

1. $\|A\| \geq 0$ e $\|A\| = 0 \implies A = 0$.

2. $\|\alpha A\| = |\alpha| \|A\|$.
3. $\|A + B\| \leq \|A\| + \|B\|$.

La norma matriciale non indotta più importante è la **norma di Frobenius** definita da

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{1/2},$$

che è la norma $\|\cdot\|_2$ applicata al vettore mn -dimensionale. In particolare, vale che

$$\|A\|_F = \sqrt{\text{tr}(A^*A)} = \sqrt{\text{tr}(AA^*)}.$$

Anche in questo caso vale l'upper bound

$$\|AB\|_F \leq \|A\|_F \|B\|_F.$$

Teo. (Invarianza per moltiplicazione unitaria) Per qualunque matrice $A_{m \times n}$ e $Q_{m \times m}$ unitaria, vale che

$$\|QA\|_2 = \|A\|_2$$

$$\|QA\|_F = \|A\|_F$$

Dim.

Per la 2-norma, per ogni y vale che $\|Qy\|_2 = \|y\|_2$, per cui $\sup \|QAy\|_2 = \sup \|Ay\|_2$.

Per la norma di Frobenius, siccome Q è hermitiana,

$$\|QA\|_F = \sqrt{\text{tr}((AQ)(AQ)^*)} = \sqrt{\text{tr}(AQQ^*A^*)} = \sqrt{\text{tr}(AA^*)} = \|A\|_F.$$

□

Il teorema vale anche se $Q_{p \times m}$ è rettangolare ($p > m$) con colonne ortonormali.

Capitolo 2

Ortogonalità e minimi quadrati

Lezione 4: Proiettori

2020-01-26

Def. (Proiettore) Si definisce proiettore una matrice P idempotente, tale che

$$P^2 = P.$$

Osservazioni

1. P è una matrice di proiezione sul sottospazio $\text{range}(P)$, in quanto

$$P(Pv) = P^2v = Pv.$$

2. La direzione della “luce” sul vettore v è quella di $Pv - v$, infatti

$$P(Pv - v) = Pv - Pv = 0,$$

in particolare $Pv - v \in \ker P$.

Def. (Proiettore complementare) Se P è un proiettore, allora la matrice $I - P$ si chiama proiettore complementare.

Osservazioni

1. Anche $I - P$ è idempotente, infatti

$$(I - P)^2 = (I - P)(I - P) = I - P - P + P^2 = I - P.$$

2. $I - P$ proietta su $\ker P$, infatti

$$\begin{cases} v \in \ker P \implies Pv = 0 \implies (I - P)v = v \implies \ker P \subseteq \text{Im}(I - P) \\ (I - P)v = v - Pv \in \ker P \implies \text{Im}(I - P) \subseteq \ker P \end{cases}$$

dunque $\ker P = \text{Im}(I - P)$.

In particolare, poiché $\text{Im } P \cap \text{Im}(I - P) = \text{Im } P \cap \ker P = \{0\}$, una matrice di proiezione separa lo spazio \mathbb{C}^m in due sottospazi S_1, S_2 in somma diretta, ovvero tali che $S_1 \cap S_2 = \{0\}$.

Def. (Proiettore ortogonale) Un proiettore P è ortogonale se i sottospazi S_1 e S_2 sono ortogonali.

Teo. (Proiettore ortogonale) Un proiettore P è un proiettore ortogonale se e solo se

$$P = P^*$$

Dim.

$$\Rightarrow : \langle Pv, (I - P)w \rangle = v^* P^* (I - P)w = v^* (P^* - P^* P)w = v^* (P - P)w = 0.$$

\Leftarrow : più lunga ...

□

Proiettore con base ortogonale

Siccome un proiettore ortogonale ha solo valori singolari uguali a 1 e 0, si può scrivere la decomposizione a valori singolari $P = Q\Sigma Q^*$ e rimuovere le ultime colonne di Q , corrispondenti ai valori singolari nulli:

$$P = \hat{Q}\hat{Q}^*,$$

con \hat{Q} matrice ortonormale.

In generale, un qualunque prodotto $\hat{Q}\hat{Q}^*$ è un proiettore ortogonale su $\text{Im } \hat{Q}$, è solo necessario che le sue colonne siano ortonormali. Analogamente, $I - \hat{Q}\hat{Q}^*$ è il proiettore ortogonale complementare.

Esempio: Proiezione lungo un vettore

Se si prende un vettore arbitrario v , allora

$$P = \frac{vv^*}{v^*v}$$

è la matrice di proiezione ortogonale lungo il vettore v .

Proiettore ortogonale con base arbitraria

Data una qualunque base $\{a_1, a_2, \dots, a_n\}$ di un sottospazio V di \mathbb{C}^m si può costruire una matrice di proiezione ortogonale. La condizione è che per ogni j la proiezione Ax sia tale che

$$a_j^*(Ax - v) = 0 \iff A^*(Ax - v) = 0 \iff A^*Ax = A^*v \iff x = (A^*A)^{-1}A^*v,$$

dunque il proiettore su $\text{Im } A$ è

$$P = A(A^*A)^{-1}A^*.$$

2020-02-02

Lezione 5: Fattorizzazione QR

L'idea della fattorizzazione QR è di costruire una sequenza di vettori ortonormali q_1, q_2, \dots che generino in sequenza gli spazi generati dalle prime k colonne di A

$$\langle a_1 \rangle \subseteq \langle a_1, a_2 \rangle \subseteq \langle a_1, a_2, a_3 \rangle \subseteq \dots,$$

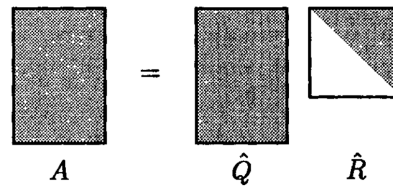
in particolare se $A \in \mathbb{C}^{m \times n}$ è di rango n ,

$$\langle q_1, q_2, \dots, q_j \rangle = \langle a_1, a_2, \dots, a_j \rangle \quad \forall j = 1, \dots, n.$$

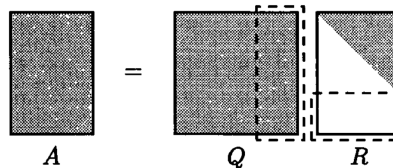
Questo è equivalente a richiedere che a_1, \dots, a_j siano una combinazione lineare di q_1, \dots, q_j per ogni j , ovvero

$$A_{m \times n} = \hat{Q}_{m \times n} \hat{R}_{n \times n}, \quad \hat{R} \text{ triangolare superiore,}$$

e si chiama **fattorizzazione QR ridotta di A** .



La **fattorizzazione QR completa di A** aggiunge $m - n$ colonne ortogonali a \hat{Q} in modo da farla diventare una matrice unitaria ($QQ^* = I$), mentre ad \hat{R} vengono aggiunte $m - n$ righe di zeri.



Osservazione

Le ultime $m - n$ colonne di Q sono costituite da vettori ortogonali alle prime n , per cui formano una base ortonormale di $\ker A$, lo spazio ortogonale a A .

Per ottenere una fattorizzazione QR , si può utilizzare la **decomposizione di Gram-Schmidt**, dal momento che $A = QR \implies Q = AR^{-1}$, ovvero

$$\begin{aligned} q_1 &= \frac{a_1}{r_{11}} \\ q_2 &= \frac{a_2 - r_{12}q_1}{r_{22}} \\ q_3 &= \frac{a_3 - r_{13}q_1 - r_{23}q_2}{r_{33}} \\ &\vdots \end{aligned}$$

dove $r_{ij} = \langle q_i, a_j \rangle$ e $r_{ii} = \langle a_i, a_i \rangle$.

Algorithm 1 Gram-Schmidt (Instabile)

Input: a_1, a_2, \dots, a_n

```

1: for  $j = 1$  to  $n$  do
2:    $v_j = a_j$ 
3:   for  $i = 1$  to  $j - 1$  do
4:      $r_{ij} = q_i^* a_j$ 
5:      $v_j = v_j - r_{ij} q_i$ 
6:   end for
7:    $r_{jj} = \|v_j\|_2$ 
8:    $q_j = v_j / r_{jj}$ 
9: end for

```

Teo. (Esistenza della fattorizzazione QR) Ogni matrice $A \in \mathbb{C}^{m \times n}$, con $m \geq n$, possiede una fattorizzazione QR completa e, quindi, anche ridotta.

Dim.

Se A ha rango pieno, la fattorizzazione QR ridotta esiste in virtù dell'algoritmo di Gram-Schmidt, per cui esiste anche una fattorizzazione completa.

Se A non ha rango pieno, allora un qualche vettore v_j risulta pari a 0, per cui si può scegliere un qualunque vettore q_j ortogonale ai precedenti. Analogamente, si costruisce la fattorizzazione completa. \square

Teo. (Unicità della fattorizzazione QR) La fattorizzazione QR ridotta per una matrice $A \in \mathbb{C}^{m \times n}$ di rango pieno è unica, a meno di cambiare segno di un r_{jj} .

Dim.

Conseguenza immediata dell'algoritmo di Gram-Schmidt, dove non è specificato il segno del fattore di normalizzazione r_{jj} . Se si sceglie $r_{jj} > 0$, la fattorizzazione QR ridotta è univocamente determinata. \square

Basi ortonormali di polinomi

Si consideri lo spazio vettoriale $L^2[-1, 1]$ di polinomi a valori complessi. Si può applicare la fattorizzazione QR con prodotto scalare definito da

$$\langle f, g \rangle = \int_{-1}^1 \overline{f(x)} g(x) dx.$$

La matrice di interesse è

$$A = \begin{pmatrix} 1 & x & x^2 & x^3 & \dots & x^{n-1} \end{pmatrix},$$

da fattorizzare in

$$A = QR = \begin{pmatrix} q_0(x) & q_1(x) & \dots & q_{n-1}(x) \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & r_{nn} \end{pmatrix}$$

con Q tale che

$$\int_{-1}^1 \overline{q_i(x)} q_j(x) dx = \delta_{ij} = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

La soluzione a questo problema permette di scrivere una base ortonormale dei polinomi in $L^2[-1, 1]$, data dai **polinomi di Legendre**

$$q_0(x) = 1$$

$$q_1(x) = x$$

$$q_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$$

$$q_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$$

\vdots

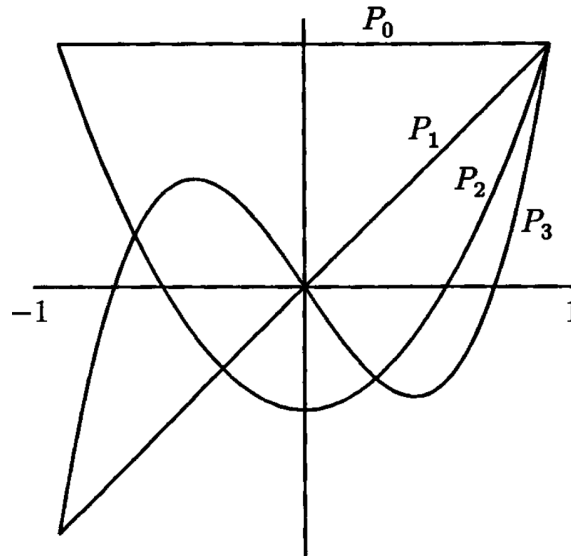


Figura 3: Primi quattro polinomi di Legendre.

La “matrice di proiezione” associata alla matrice \hat{Q} è l’operatore di proiezione che mappa funzioni su $L^2[-1, 1]$ in funzioni su $L^2[-1, 1]$ tramite

$$f(\cdot) \mapsto \sum_{j=0}^{n-1} \underbrace{q_j(\cdot)}_{\text{direzione}} \underbrace{\int_{-1}^1 \overline{q_j(x)} f(x) dx}_{\text{intensità}},$$

con una struttura simile alla proiezione discreta.

Soluzione di un sistema $Ax = b$ tramite QR

Siccome la matrice Q è ortogonale, si può risolvere un sistema di equazioni

$$Ax = b$$

trovando la fattorizzazione QR di A:

$$QRx = b \implies Rx = Q^*b,$$

che è più semplice in quanto R è una matrice triangolare superiore.

Lezione 6: Algoritmo di Gram-Schmidt

2020-02-02

È possibile scrivere l'algoritmo di Gram-Schmidt con l'uso di matrici di proiezione:

$$\begin{aligned} q_1 &= \frac{P_1 a_1}{\|P_1 a_1\|} \\ q_2 &= \frac{P_2 a_2}{\|P_2 a_2\|} \\ &\vdots \\ q_n &= \frac{P_n a_n}{\|P_n a_n\|} \end{aligned}$$

per delle opportune matrici P_1, \dots, P_n . Ciascuna P_j è una matrice $m \times m$ che proietta sulla componente ortogonale a $\langle q_1, \dots, q_{j-1} \rangle$ (se $j = 1$, $P_1 = I$).

Da questa proprietà è chiaro che P_j si può scrivere in termini di \hat{Q}_{j-1} , matrice contenente le prime $j - 1$ colonne di \hat{Q} :

$$\hat{Q}_{j-1} = \begin{pmatrix} q_1 & q_2 & \dots & q_{j-1} \end{pmatrix},$$

allora

$$P_j = I - \hat{Q}_{j-1} \hat{Q}_{j-1}^*.$$

Algoritmo di Gram-Schmidt stabile

L'algoritmo di Gram-Schmidt così discusso non è utilizzato, in quanto la singola proiezione $P_j a_j$ è numericamente instabile. È più stabile se si applica la proiezione sul sottospazio ortogonale a q_j a tutti i vettori nell'istante in cui si conosce q_j :

Algorithm 2 Gram-Schmidt stabile

Input: a_1, a_2, \dots, a_n

Output: v_1, v_2, \dots, v_n ortonormali

```

1: for  $i = 1$  to  $n$  do
2:    $v_i = a_i$ 
3: end for
4: for  $i = 1$  to  $n$  do
5:    $r_{ii} = \|v_i\|$ 
6:    $q_i = v_i / r_{ii}$ 
7:   for  $j = i + 1$  to  $n$  do
8:      $r_{ij} = q_i^* v_j$ 
9:      $v_j = v_j - r_{ij} q_i$ 
10:  end for
11: end for
```

Teo. (Costo computazionale GS) *L'algoritmo di Gram-Schmidt necessita di $O(2np^2)$ operazioni per calcolare la fattorizzazione QR di una matrice $n \times p$.*

Dim.

Le operazioni che dominano il costo computazionale sono nel ciclo FOR annidato:

$$\begin{aligned} r_{ij} &= q_i^* v_j \\ v_j &= v_j - r_{ij} q_i \end{aligned}$$

Il prodotto scalare richiede n moltiplicazioni e $n - 1$ addizioni, mentre la sottrazione richiede n prodotti e n sottrazioni. In totale, ci sono $\approx 4n$ operazioni ad ogni iterazione del ciclo FOR, che in totale ammontano a

$$\sum_{i=1}^p \sum_{j=i+1}^p 4n \sim 4n \cdot \sum_{i=1}^p i \sim 2np^2$$

□

Una dimostrazione visiva si può avere dalla seguente rappresentazione delle operazioni:

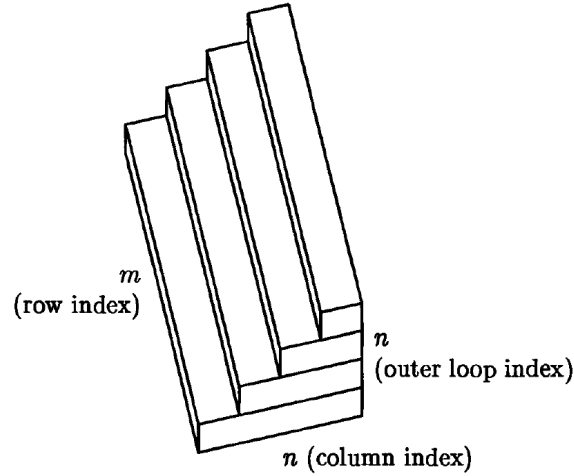


Figura 4: Flops nell'algoritmo di Gram-Schmidt modificato

Il prisma ha volume $np^2/2$, moltiplicato per le 4 operazioni ad ogni ciclo dà il costo totale $2np^2$.

Ogni step dell'algoritmo di Gram-Schmidt modificato cambia tutti i vettori a_1, \dots, a_n della matrice di partenza, moltiplicando per una matrice triangolare superiore:

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix} \begin{pmatrix} \frac{1}{r_{11}} & -\frac{r_{12}}{r_{11}} & -\frac{r_{13}}{r_{11}} & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} q_1 & v_2^{(2)} & \dots & v_n^{(2)} \end{pmatrix}$$

e così via utilizzando successivamente

$$R_2 = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & \frac{1}{r_{22}} & -\frac{r_{23}}{r_{22}} & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad R_3 = \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & \frac{1}{r_{33}} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad \dots$$

Alla fine dell'iterazione, si ottiene

$$A \underbrace{R_1 R_2 \dots R_n}_{\hat{R}^{-1}} = \hat{Q},$$

che mostra come l'algoritmo di Gram-Schmidt sia un metodo di **ortogonalizzazione triangolare**.

Lezione 7: Triangolarizzazione di Householder

2020-02-03

Mentre l'algoritmo di Gram-Schmidt applica una sequenza di triangolarizzazioni ortogonali, il metodo di Householder applica una sequenza di matrici unitarie Q_k in modo che A diventi una matrice triangolare superiore

$$\underbrace{Q_n \dots Q_2 Q_1}_{Q^*} A = R,$$

con R triangolare superiore.

Gram-Schmidt:	$AR_1 \dots R_n = \hat{Q}$	ortogonalizzazione triangolare
Householder:	$Q_n \dots Q_1 A = R$	triangolarizzazione ortogonale

Triangolarizzazione introducendo zeri

Ciascuna matrice ortogonale Q_k introduce zeri sotto il k -esimo elemento della diagonale, preservando gli altri già introdotti:

$$\begin{array}{c} \begin{bmatrix} \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \\ \times & \times & \times \end{bmatrix} \xrightarrow{Q_1} \begin{bmatrix} \times & \times & \times \\ \mathbf{0} & \times & \times \\ \mathbf{0} & \times & \times \\ \mathbf{0} & \times & \times \\ \mathbf{0} & \times & \times \end{bmatrix} \xrightarrow{Q_2} \begin{bmatrix} \times & \times & \times \\ & \times & \times \\ & \mathbf{0} & \times \\ & \mathbf{0} & \times \\ & \mathbf{0} & \times \end{bmatrix} \xrightarrow{Q_3} \begin{bmatrix} \times & \times & \times \\ & \times & \times \\ & & \times \\ & & \mathbf{0} \\ & & \mathbf{0} \end{bmatrix} \\ A \qquad \qquad Q_1 A \qquad \qquad Q_2 Q_1 A \qquad \qquad Q_3 Q_2 Q_1 A \end{array}$$

Figura 5: In grassetto gli elementi modificati, \times indica un elemento generico, 0 e vuoto indicano zeri.

In generale, Q_k agisce sulle righe $i \geq k$ e dopo n step tutte le entrate sotto la diagonale sono state eliminate.

Riflettori di Householder

Le matrici Q_k sono costruite nella forma

$$Q_k = \begin{pmatrix} I_{k-1} & 0 \\ 0 & F \end{pmatrix},$$

dove I_{k-1} è la matrice identità $(k-1) \times (k-1)$ e F è una matrice ortogonale $(m-k+1) \times (m-k+1)$ che introduce zeri nella k -esima colonna. Questa matrice si sceglie in modo che sia un **riflettore di Householder**.

Un riflettore di Householder riflette un punto x rispetto a un sottospazio H nel modo seguente:

1. Il vettore viene proiettato sul sottospazio H ortogonale a un vettore v .
2. La proiezione viene prolungata per la stessa lunghezza, in modo da riflettere il punto rispetto ad H .

Per il punto (1), è sufficiente la proiezione ortogonale sul sottospazio ortogonale a v data da

$$P = I - \frac{vv^*}{v^*v},$$

che va prolungata per ottenere la riflessione

$$F = I - 2\frac{vv^*}{v^*v},$$

e la matrice F così definita è di rango pieno e unitaria.

Nella figura, si riflette il vettore x su $\|x\|e_1$ usando $v = \|x\|e_1 - x$, per inserire gli zeri nel modo corretto sotto la prima componente.

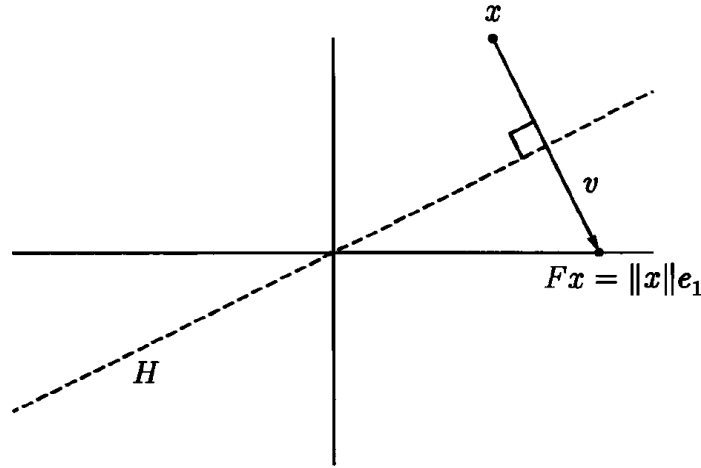


Figura 6: Riflessione di Householder

In realtà, per ragioni di stabilità numerica, si sceglie la componente che fa muovere maggiormente il vettore x , ovvero

$$\begin{aligned} v &= -\operatorname{sgn}(x_1)\|x\|e_1 - x_1 \\ &= \operatorname{sgn}(x_1)\|x\|e_1 + x. \end{aligned}$$

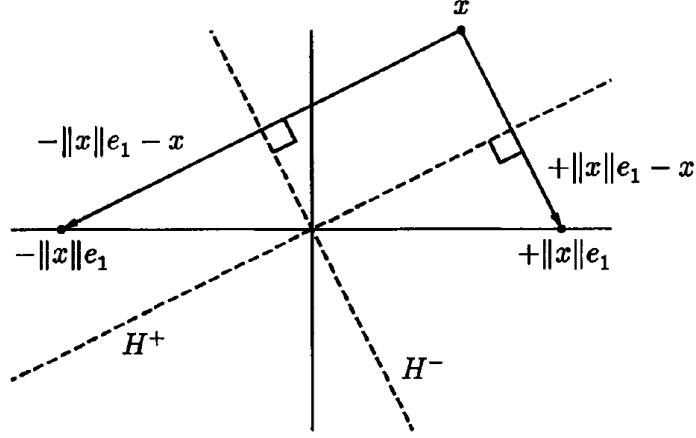


Figura 7: Due possibili riflessioni, per stabilità si sceglie quella che muove x per una distanza maggiore

Notazione

$A_{i:i', j:j'}$	sottomatrice di A con elementi estremi a_{ij} e $a_{i'j'}$
$A_{i, j:j'}, A_{i:i', j}$	sottovettore rispettivamente di una riga o colonna

Algorithm 3 Fattorizzazione QR di Householder

- 1: **for** $k = 1$ to n **do**
 - 2: $x = A_{k:m,k}$ ▷ k -esima colonna, righe da k in giù
 - 3: $v_k = \text{sgn}(x_1)\|x\|e_1 + x$
 - 4: $v_k = v_k / \|v_k\|$
 - 5: $A_{k:m,k:n} = A_{k:m,k:n} - 2v_k v_k^* A_{k:m,k:n}$
 - 6: **end for**
-

Che ha costo computazionale $O(mn^2)$. La matrice Q si può costruire esplicitamente attraverso un altro algoritmo, sfruttando la proprietà

$$Q^* = Q_n \cdots Q_2 Q_1 \implies Q = Q_1 Q_2 \cdots Q_n$$

con costo computazionale $O(mn)$:

Algorithm 4 Calcolo implicito del prodotto Qx

- 1: **for** $k = n$ down to 1 **do**
 - 2: $x_{k:m} = x_{k:m} - 2v_k v_k^* x_{k:m}$
 - 3: **end for**
-

Applicando questo algoritmo a e_1, e_2, \dots, e_m si possono calcolare le colonne di Q .

Per una matrice $A_{n \times p}$, il costo computazionale dell'algoritmo è $\sim 2np^2 - \frac{2}{3}p^3$ operazioni.

Lezione 8: Minimi quadrati

2020-02-04

L'obiettivo è trovare la soluzione x a un'equazione sovradeterminata

$$Ax = b,$$

in modo che minimizzi la norma dei residui

$$\min_{x \in \mathbb{C}^n} \|b - Ax\|_2,$$

che è particolarmente conveniente dal momento che la derivata di una funzione quadratica è lineare.

Teo. (Residui dei minimi quadrati) Un vettore $x \in \mathbb{C}^n$ è tale da minimizzare la norma dei residui

$$\|r\|_2 = \|b - Ax\|_2$$

se e solo se $r \perp \text{span } A$, ovvero

$$A^*Ax = A^*b \quad (\text{Equazioni normali}),$$

che ha soluzione unica se e solo se A è di rango pieno.

Dim.

Libro. □

La soluzione x è data da

$$x = (A^*A)^{-1}A^*b,$$

con la matrice $A^+ = (A^*A)^{-1}A^*$ chiamata **pseudoinversa di A** . Di seguito ci sono tre algoritmi per calcolare i minimi quadrati:

OLS con fattorizzazione di Cholesky

Algorithm 5 Minimi quadrati via equazioni normali

▷ Costo $n \times p$: $\mathcal{O}(np^2 + \frac{1}{3}p^3)$

Scrivere A^*A e A^*b .

Calcolare la decomposizione di Cholesky $A^*A = R^*R$.

Risolvere in w il sistema $R^*w = A^*b$.

Risolvere in x il sistema $Rx = w$.

Osservazione

La procedura è poco stabile per modelli complicati, funziona invece bene per modelli piccoli.

OLS con fattorizzazione QR

I minimi quadrati con fattorizzazione $A = \hat{Q}\hat{R}$ si possono calcolare sfruttando il fatto che $\hat{Q}^*\hat{Q} = I$, per cui

$$A = \hat{Q}\hat{R} \implies A(A^*A)^{-1}A^* = \hat{Q}\hat{Q}^*,$$

con pseudoinversa

$$A^+ = \hat{R}^{-1}\hat{Q}^*.$$

Algorithm 6 Minimi quadrati tramite fattorizzazione QR

▷ Costo $n \times p$: $\mathcal{O}(2np^2 - \frac{2}{3}p^3)$

Calcolare la fattorizzazione $A = \hat{Q}\hat{R}$.

Calcolare il vettore \hat{Q}^*b .

Risolvere in x il sistema triangolare superiore $\hat{R}x = \hat{Q}^*b$.

OLS usando la SVD

I minimi quadrati si possono calcolare attraverso la decomposizione a valori singolari di A

$$A = \hat{U}\hat{\Sigma}V^*,$$

che ha matrice di proiezione $P = \hat{U}\hat{U}^*$ e sistema di equazioni normali

$$\hat{\Sigma}V^*x = \hat{U}^*b.$$

Algorithm 7 Minimi quadrati tramite SVD

▷ Costo: $\mathcal{O}(2np^2 + 11p^3)$

Calcolare la SVD ridotta $A = \hat{U}\hat{\Sigma}V^*$.

Calcolare \hat{U}^*b .

Risolvere in w il sistema $\hat{\Sigma}w = \hat{U}^*b$.

Calcolare $x = Vw$.

Osservazioni

1. Il metodo con fattorizzazione di Cholesky è il più veloce, ma instabile per modelli grandi.
2. Il metodo QR è standard, ma instabile se A ha autovalori molto piccoli.
3. Il metodo basato sulla SVD è più stabile, ma lento se p è grande.

Capitolo 3

Condizionamento e stabilità

Lezione 9: Condizionamento

2020-02-05

In astratto, un **problema** è una funzione $f : X \rightarrow Y$ da uno spazio normato X di dati a uno spazio normato Y di soluzioni. A questo problema è associata un'istanza di **problema**, nel momento in cui si ottengono dei dati x che portano alla soluzione $f(x)$.

Un **problema ben condizionato** è un problema in cui tutte le piccole perturbazioni di x portano a piccoli cambiamenti in $f(x)$.

In un **problema mal condizionato**, invece, alcune piccole perturbazioni di x portano a grandi cambiamenti in $f(x)$.

Def. (Numero di condizionamento) Siano δx una piccola perturbazione di x , $\delta f = f(x + \delta x) - f(x)$ il corrispondente cambiamento di f .

Si chiama numero di condizionamento assoluto $\hat{\kappa} = \hat{\kappa}(x)$ del problema f in x il valore

$$\hat{\kappa} = \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \frac{\|\delta f\|}{\|\delta x\|}$$

Se f è differenziabile, si può usare la derivata di f per calcolare $\hat{\kappa}$: sia $J(x)$ la matrice Jacobiana di f in x , allora per definizione

$$\delta f \approx J(x)\delta x + o(\delta x),$$

da cui si ottiene che

$$\hat{\kappa} = \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \frac{\|\delta f\|}{\|\delta x\|} = \|J(x)\|,$$

con $\|J(x)\|$ la norma di $J(x)$ indotta dalle norme su X e Y .

Def. (Numero di condizionamento relativo) Si definisce numero di condizionamento

relativo $\kappa = \kappa(x)$ come

$$\kappa = \lim_{\delta \rightarrow 0} \sup_{\|\delta x\| \leq \delta} \left(\frac{\|\delta f\|}{\|f(x)\|} \bigg/ \frac{\|\delta x\|}{\|x\|} \right).$$

Di nuovo, se f è differenziabile si può usare lo Jacobiano per definire

$$\kappa = \frac{\|J(x)\|}{\|f(x)\|/\|x\|}.$$

Il numero di condizionamento relativo è generalmente più utile, in quanto l'aritmetica a virgola mobile introduce errori relativi.

Un problema **ben condizionato** ha un numero di condizionamento relativo κ piccolo ($1, 10, 10^2$), mentre un problema **mal condizionato** ha κ elevato ($10^6, 10^{16}$).

Esempio : Calcolo di uno scalare

Il problema $f : x \rightarrow x^2$ ha numero di condizionamento relativo

$$\kappa = \frac{\|J(x)\|}{\|f(x)\|/\|x\|} = \frac{1/2}{(x/2)/x} = 1,$$

per cui è un problema ben condizionato.

Esempio : Calcolo della differenza

Si consideri $f(x_1, x_2) = x_1 - x_2$, utilizzando la norma ∞ . La matrice Jacobiana di f è

$$J = \begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 1 & -1 \end{pmatrix},$$

con $\|J(x)\| = 2$. Il numero di condizionamento, dunque, è

$$\kappa = \frac{\|J\|_\infty}{\|f(x)\|/\|x\|} = \frac{2}{|x_1 - x_2|/\max\{|x_1|, |x_2|\}}.$$

Poichè il denominatore dipende da $|x_1 - x_2|$, per $x_1 \approx x_2$ il numero di condizionamento diventa arbitrariamente grande.

Esempio : Calcolo di $\tan(x)$

La funzione $\tan(x)$ per $x \approx 10^{100}$ ha cambiamenti arbitrariamente grandi per minuscole perturbazioni di x . Quindi, il calcolo di $\tan(10^{100})$ è praticamente impossibile per un normale computer.

Esempio : Calcolo delle radici di un polinomio

Il calcolo delle radici di un polinomio è un problema mal condizionato. Si consideri il polinomio $x^2 - 2xp + 1$, del quale si cercano le radici, con $p \approx 1$. Allora, una delle due soluzioni è

$$x_+(p) = p + \sqrt{p^2 - 1}.$$

Il numero di condizionamento relativo è

$$\begin{aligned}
 \kappa &= \frac{dx_+}{dp} \bigg/ \frac{x_+}{p} \\
 &= \left(1 + \frac{p}{\sqrt{p^2 - 1}} \right) \bigg/ \frac{p + \sqrt{p^2 - 1}}{p} \\
 &= \dots \\
 &= \frac{p}{\sqrt{p^2 - 1}},
 \end{aligned}$$

che per valori di $p \approx \pm 1$ diventa arbitrariamente grande. Il numero di condizionamento relativo, quindi, è $\kappa = +\infty$.

Condizionamento di un prodotto matrice-vettore

Considerato un problema $f : x \rightarrow Ax$, una perturbazione di x porta a

$$\begin{aligned}
 \kappa &= \sup_{\delta x} \left(\frac{\|A(x + \delta x) - Ax\|}{\|Ax\|} \bigg/ \frac{\|\delta x\|}{\|x\|} \right) \\
 &= \left(\sup_{\delta x} \frac{\|A\delta x\|}{\|\delta x\|} \bigg/ \frac{\|Ax\|}{\|x\|} \right) \\
 &= \|A\| \frac{\|x\|}{\|Ax\|},
 \end{aligned}$$

che dipende da A e x . In particolare, poiché per A quadrata e nonsingolare vale che $\|x\|/\|Ax\| \leq \|A^{-1}\|$, si può anche scrivere

$$\kappa \leq \|A\| \cdot \|A^{-1}\|.$$

Se A non è quadrata, si può sostituire A^{-1} con la pseudoinversa $(A^*A)^{-1}A^*$.

Analogamente, se si sostituisce A con A^{-1} , si ha il numero di condizionamento per il problema $b \rightarrow A^{-1}b$.

Teo. (Numero di condizionamento del prodotto) Sia $A_{m \times m}$ non singolare, si consideri il prodotto $Ax = b$. Il problema di calcolare b dato x ha numero di condizionamento

$$\kappa = \|A\| \frac{\|x\|}{\|b\|} \leq \|A\| \cdot \|A^{-1}\|,$$

rispetto alle perturbazioni di x . Il problema di calcolare x dato b ha numero di condizionamento

$$\kappa = \|A^{-1}\| \frac{\|b\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\|,$$

rispetto alle perturbazioni di b . Se $\|\cdot\| = \|\cdot\|_2$, allora la prima è un'uguaglianza se x è multiplo di un vettore singolare destro di A corrispondente al minimo valore singolare σ_m . Analogamente, la seconda è un'uguaglianza se b è multiplo di un vettore singolare sinistro di A corrispondente al massimo valore singolare σ_1 .

Il valore $\kappa(A) = \|A\| \cdot \|A^{-1}\|$ si chiama anche **numero di condizionamento di A** , con $\kappa(A) = \infty$ se A è singolare.

Osservazione

Se $\|\cdot\| = \|\cdot\|_2$, vale che

$$\kappa(A) = \frac{\sigma_1}{\sigma_m},$$

che si può interpretare come l'eccentricità dell'iperellisse immagine della sfera unitaria sotto A .

Per una matrice A rettangolare, si definisce $\kappa(A) = \|A\| \cdot \|A^+\|$, particolarmente utile nel problema dei minimi quadrati. In tal caso, con norma $\|\cdot\|_2$, vale

$$\kappa(A) = \frac{\sigma_1}{\sigma_n}.$$

Teo. (Numero di condizionamento per A quadrata) Se A è quadrata e invertibile, il problema $x = A^{-1}b$ ha numero di condizionamento

$$\kappa = \kappa(A) = \|A\| \cdot \|A^{-1}\|.$$

Questi due teoremi sono fondamentali per calcolare quanto accuratamente si può risolvere un sistema di equazioni.

Lezione 10: Stabilità

2020-02-08

Notazione

Con $h(x) = \mathcal{O}(\varepsilon_{\text{machine}})$ si intende che per un certo limite ($x \rightarrow 0, x \rightarrow \infty, \dots$) vale

$$|h(x)| \leq C\varepsilon_{\text{machine}} \quad \text{uniformemente in } x.$$

Definito un problema come una funzione $f : X \rightarrow Y$, si può definire un **algoritmo** come un'altra funzione $\tilde{f} : X \rightarrow Y$ che, attraverso l'approssimazione del programma, fornisce una soluzione $x \rightarrow \tilde{f}(x)$.

Accuratezza

Un buon algoritmo deve necessariamente approssimare f , ovvero l'**errore relativo**

$$\frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|}$$

deve essere più piccolo possibile.

Stabilità

Un algoritmo deve essere non solo accurato, ma anche **stabile**, in quanto l'approssimazione delle quantità in ingresso è inevitabile. Per questo motivo, a parole

A stable algorithm gives nearly the right answer to nearly the right question.

In formule, per ogni $x \in X$ esiste \tilde{x} tale che

$$\frac{\|\tilde{x} - x\|}{\|x\|} = \mathcal{O}(\varepsilon_{\text{machine}})$$

e si ha che

$$\frac{\|\tilde{f}(x) - f(\tilde{x})\|}{\|f(\tilde{x})\|} = \mathcal{O}(\varepsilon_{\text{machine}}).$$

Stabilità all'indietro

Un algoritmo è **stabile all'indietro** se per ogni $x \in X$,

$$\tilde{f}(x) = f(\tilde{x}) \text{ per qualche } \tilde{x} \text{ tale che } \frac{\|\tilde{x} - x\|}{\|x\|} = \mathcal{O}(\varepsilon_{\text{machine}}).$$

A parole,

A stable algorithm gives exactly the right answer to nearly the right question.

Teo. (Indipendenza dalla norma) *Per il problema f e l'algoritmo \tilde{f} definiti su X e Y finito-dimensionali, le proprietà di accuratezza, stabilità e stabilità all'indietro sono tutte verificate o non verificate indipendentemente dalla scelta di norma in X e in Y .*

Dim.

In uno spazio vettoriale finito-dimensionale, tutte le norme sono equivalenti, nel senso che se $\|\cdot\|$ e $\|\cdot\|'$ sono due norme sullo stesso spazio, allora esistono $C_1, C_2 > 0$ costanti tali che

$$C_1\|x\| \leq \|x\|' \leq C_2\|x\|$$

per ogni x nello spazio. Dunque, modificare la norma porta a un cambiamento nella costante C di $\mathcal{O}(\varepsilon_{\text{machine}})$, ma non nell'esistenza di tale costante. \square