METODI ITERATIVI PER MATRICI SIMMETRICHE E DEFINITE POSITIVE

$(A^T = A \qquad (A\underline{x}, \underline{x}) > 0 \ \ \forall \underline{x} \in \mathbb{R}^n \ \ \underline{x} \neq 0)$

Risolvere $A\underline{x} = \underline{b} \Rightarrow$

- (P) : Trovare $\underline{x}^* \in \mathbb{R}^n$: $A\underline{x}^* = \underline{b}$

DEF. : Sia F il funzionale: $\mathbb{R}^n \to \mathbb{R}$

$$F(\underline{v}) := \frac{1}{2}(A\underline{v}, \underline{v}) - (\underline{b}, \underline{v}) \qquad \forall \underline{v} \in \mathbb{R}^n$$

Si ha il seguente risultato:

TEOREMA Se A è simmetrica e definita positiva, il problema (P) è equivalente al seguente problema di minimo

$$(PM) \begin{cases} \text{Trovare } \underline{u} \in \mathbb{R}^n : \\ F(\underline{u}) = \displaystyle\min_{\underline{v} \in \mathbb{R}^n} F(\underline{v}) \end{cases}$$

(cioè (PM) ha soluzione unica $\underline{u} \in \mathbb{R}^n$

e $\underline{u} \equiv \underline{x}^*$

$$\boxed{F(\underline{v}) := \frac{1}{2}(A\underline{v},\underline{v}) - (\underline{b},\underline{v})}$$

Dim.: F è un funzionale quadratico. Inoltre

$$\underline{\nabla} F(\underline{v}) = \begin{pmatrix} \frac{\partial F}{\partial v_1} \\ \vdots \\ \frac{\partial F}{\partial v_n} \end{pmatrix} \equiv A\underline{v} - \underline{b} \qquad H(F) = A \text{ def. pos.}$$

L'Hessiana di F ha autovalori positivi

$\Rightarrow$ F convesso (strettamente)

(estensione a $\mathbb{R}^n$ di un paraboloide convesso)

$\Rightarrow$ F ha un minimo unico

Sia $\underline{u} \in \mathbb{R}^n$ il punto di minimo. Allora

$$\underline{\nabla} F(\underline{u}) = 0 \Rightarrow A\underline{u} - \underline{b} = 0$$

$\Rightarrow \underline{u} \equiv \underline{x}^*$ per l'unicità della soluzione di (P)

METODI DI DISCESA

Si cerca la soluzione $\underline{x}^*$ di (P) come punto di minimo di $F(\underline{x})$

Partendo da $\underline{x}^0$ dato, si costruisce una successione $\{\underline{x}^k\}$ t.c. $\underline{x}^k \to \underline{x}^*$

$$\underline{x}^0 \in \mathbb{R}^n \text{ dato } (\forall) \quad (\text{tipicamente } \underline{x}^0 = \underline{0})$$

$$k \geq 0 \quad \underline{x}^{k+1} = \underline{x}^k + \alpha_k \underline{p}^k$$

con:

$$\left.\begin{array}{l} \alpha_k \in \mathbb{R} \ (\text{scalare}) \\ \underline{p}^k \in \mathbb{R}^n \ \text{direzioni di discesa} \end{array}\right\} \text{da scegliere}$$

in modo che

$$F(\underline{x}^{k+1}) < F(\underline{x}^k)$$

Come scegliere α_k, $\underline{p}^k$?

$$\boxed{\underline{x}^{k+1} = \underline{x}^{k} + \alpha_k \underline{p}^{k}}$$ si vuole $F(\underline{x}^{k+1}) < F(\underline{x}^{k})$

Supponiamo scelto $\underline{p}^k$. Cerchiamo α_k tale che:

$$F(\underline{x}^k + \alpha_k \underline{p}^k) < F(\underline{x}^k)$$

Sia $g(\alpha) := F(\underline{x}^k + \alpha\, \underline{p}^k) \qquad \alpha \in \mathbb{R}$

$\left(g(\alpha_k) = F(\underline{x}^{k+1}) \qquad g(0) = F(\underline{x}^k) \right)$

Deve essere: $g(\alpha_k) < g(0)$

si cerca α che minimizza g.

$$g(\alpha) = \frac{1}{2}(A(\underline{x}^k + \alpha\, \underline{p}^k), \underline{x}^k + \alpha\, \underline{p}^k) - (\underline{b}, \underline{x}^k + \alpha\, \underline{p}^k)$$

$$= \underbrace{\frac{1}{2}(A\underline{x}^k, \underline{x}^k) - (\underline{b}, \underline{x}^k)}_{F(\underline{x}^k)} + \alpha(A\underline{x}^k - \underline{b}, \underline{p}^k) + \frac{\alpha^2}{2}(A\underline{p}^k, \underline{p}^k)$$

$$= g(0) + \alpha(A\underline{x}^k - \underline{b}, \underline{p}^k) + \frac{\alpha^2}{2}\underbrace{(A\underline{p}^k, \underline{p}^k)}_{>0}$$

g è una parabola convessa

$g'(\alpha) = 0 \qquad (A\underline{x}^k - \underline{b}, \underline{p}^k) + \alpha(A\underline{p}^k, \underline{p}^k) = 0$

$$\alpha_{opt} = \frac{(\underline{b} - A\underline{x}^k, \underline{p}^k)}{(A\underline{p}^k, \underline{p}^k)} = \frac{(\underline{r}^k, \underline{p}^k)}{(A\underline{p}^k, \underline{p}^k)}$$

Scelte di $\underline{p}^k$ (cenni)

METODO DEL GRADIENTE: $\forall k \quad \underline{p}^k = -\underline{\nabla} F(\underline{x}^k)$

(lento se A è mal condizionata)

METODO DEL GRADIENTE CONIUGATO

Le direzioni di discesa sono A-coniugate:

$$(A\underline{p}^k, \underline{p}^{k-1}) = 0 \quad \forall k$$

Si può dimostrare che il metodo converge in un numero di iterazioni $\leq n$ (n = dimensione del sistema).
È quindi considerato un metodo diretto

SISTEMI SOVRADETERMINATI

(P) $\quad A\underline{x} = \underline{b} \qquad A \in \mathbb{R}^{m\times n} \qquad \underline{b} \in \mathbb{R}^m \qquad \underline{m >> n}$

Vari casi: $\begin{cases} \text{infinite soluzioni} \\ \text{nessuna soluzione} \\ \text{1 soluzione} \end{cases}$

Se A ha rango massimo (n) e il sistema ha soluzioni $\Rightarrow$ la soluzione è unica.

Nel caso in cui (P) ha soluzioni, si chiama SOLUZIONE DI (P) NEL SENSO DEI MINIMI QUADRATI il vettore $\underline{x}^* \in \mathbb{R}^n$ che minimizza il quadrato del residuo:

(PM) $\quad \underline{x}^* \in \mathbb{R}^n: \quad \|\underline{b} - A\underline{x}^*\|^2 = \min_{\underline{x} \in \mathbb{R}^n} \|\underline{b} - A\underline{x}\|^2$

<u>TEOREMA</u> Se A ha rango massimo (PM) ha soluzione unica $\underline{x}^* \in \mathbb{R}^n$, e $\underline{x}^*$ è anche soluzione di

(PN) $\qquad A^T A \underline{x}^* = A^T \underline{b} \qquad$ (eq. normali di (P))

<u>Dim</u> Con le ipotesi fatte, A^TA è simm. e def. positiva

$\Rightarrow$ (PN) ha soluzione unica

Analizziamo $F(\underline{x}) = \| \underline{b} - A\underline{x} \|^2$:

$$F(\underline{x}) = \|\underline{b}\|^2 + \|A\underline{x}\|^2 - 2(\underline{b}, A\underline{x})$$

$$= \|b\|^2 + (A^T A \underline{x}, \underline{x}) - 2(A^T \underline{b}, \underline{x})$$

F è un funzionale quadratico e convesso ($A^T A$ def. positiva)

$\nabla F(\underline{x}) = 2A^T A \underline{x} - 2A^T \underline{b}$ $\qquad$ $H(F) = 2A^T A$

$\Rightarrow$ F ha un unico punto di minimo.

Sia $\underline{u} \in \mathbb{R}^n$ tale punto

$\Rightarrow \nabla F(\underline{u}) = 0 \Rightarrow A^T(A\underline{u} - \underline{b}) = 0$

$\Rightarrow \underline{u} \equiv \underline{x}^*$

Quindi: nel caso in cui il sistema sovradeterminato

$A\underline{x} = \underline{b}$ abbia soluzione, tale soluzione

si cerca risolvendo le "equazioni normali"

$A^T A \underline{x} = A^T \underline{b}$ $\qquad$ (sistema $n \times n$)

METODO DEI MINIMI QUADRATI PER APPROSSIMARE GRANDI QUANTITÀ DI DATI SPERIMENTALI (DATA FITTING)

Richiami: Data $A \in \mathbb{R}^{m \times n}$, rango $A = n$ $(m \geq n)$, $\underline{x}^* \in \mathbb{R}^n$ è soluzione di $A\underline{x} = \underline{b}$ nel senso dei minimi quadrati se $\underline{x}^*$ risolve:

$$(PM) \quad \|\underline{b} - A\underline{x}^*\|^2 = \min_{\underline{x} \in \mathbb{R}^n} \|\underline{b} - A\underline{x}\|^2$$

Se $\underline{x}^*$ è soluzione di (PM), è anche soluzione di

$$(P) \qquad A^T A \underline{x}^* = A^T \underline{b} \qquad \text{(eq. normali)}$$

Cioè: $(P) \Leftrightarrow (PM)$

Uso del risultato per il "data fitting":

Dati m punti $(x_1, y_1), \dots, (x_m, y_m)$, scelto uno spazio vettoriale S_n di dimensione n, si cerca $p(x) \in S_n$ tale che:

$$p(x_i) = y_i \qquad i = 1, 2, \dots, m$$

<u>Problema</u>: S_n spazio vettoriale di dimensione n

$(*)$ $p(x) \in S_n$: $p(x_i) = y_i \quad i = 1, 2, \dots, m$

NOTA: Se $n = m$ e $S_n = \mathbb{P}_{n-1} \Rightarrow (*)$ ha soluzione in senso classico (polinomio interpolatore)

- Se $n < m$, $(*)$ sovradeterminato $\rightarrow$ minimi quadrati

ESEMPIO IMPORTANTE : $n = 2$ $S_n = \mathbb{P}_1$

$\dim \mathbb{P}_1 = 2$ Base $= \{1, x\}$

Cerco $p \in \mathbb{P}_1$ $(p(x) = a_1 \cdot 1 + a_2 \cdot x)$ t.c. $p(x_i) = y_i$ $i = 1, .., m$

Cioè cerco a_1, a_2 t.c.

$$F(a_1, a_2) := \sum_{i=1}^{m} (y_i - (a_1 + a_2 x_i))^2 = \text{minima}$$

F (funzione delle variabili a_1, a_2) è un paraboloide convesso $\rightarrow$ minimo unico

Per trovare il punto di minimo ($\underline{\nabla} F = \underline{0}$):

$$\frac{\partial F}{\partial a_1} = 0 \qquad \frac{\partial F}{\partial a_2} = 0 \qquad \Rightarrow \begin{cases} 2 \sum_{i=1}^{m} (y_i - (a_1 + a_2 x_i))(-1) = 0 \\ 2 \sum_{i=1}^{m} (y_i - (a_1 + a_2 x_i))(-x_i) = 0 \end{cases}$$

Sistema di 2 equazioni nelle 2 incognite a_1, a_2

$$\begin{pmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{pmatrix}$$

Sia $\bar{a}_1, \bar{a}_2$ la soluzione:

$$\boxed{p(x) = \bar{a}_1 + \bar{a}_2 x \quad \text{retta di regressione}}$$

Esercizio: verifico che $\min \|\underline{b} - A\underline{a}\|^2 \Leftrightarrow A^T A \underline{a} = A^T \underline{b}$

Pb. di partenza: trovare $p \in P_1$ t.c. $p(x_i) = y_i \ \forall i$

$p(x) = a_1 + a_2 x$: $a_1 + a_2 x_i = y_i \quad i = 1, .., m$

$$\underbrace{\begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \\ 1 & x_m \end{pmatrix}}_{A} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}}_{\underline{b}} \qquad A\underline{a} = \underline{b}$$

$$A^T = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_m \end{pmatrix}$$

$$A^T A = \begin{pmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{pmatrix} \qquad A^T \underline{b} = \begin{pmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{pmatrix}$$

Altro esempio:

$S_n = \mathbb{P}_2 \quad \dim \mathbb{P}_2 = 3 \quad \mathbb{P}_2 = \text{span}\{1, x, x^2\}$

cerco $p(x) = a_1 + a_2 x + a_3 x^2 : \; p(x_i) = y_i \quad i = 1, 2, \ldots, m$

$$\underbrace{\begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & & \\ 1 & x_m & x_m^2 \end{pmatrix}}_{A} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}}_{\underline{b}} \qquad (A\underline{a} = \underline{b})$$

$$\begin{pmatrix} m & \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \\ & \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i^3 \\ \text{sym} & & \sum_{i=1}^m x_i^4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \sum_i y_i \\ \sum_i x_i y_i \\ \sum_i x_i^2 y_i \end{pmatrix} \qquad (A^T A \underline{a} = A^T \underline{b})$$

Più in generale: Sia $S_n = \text{span}\{\varphi_1, \varphi_2, \ldots, \varphi_n\}$

Si cerca $g(x) \in S_n : \; g(x_i) = y_i \quad i = 1, \ldots, m \quad (m \gg n)$

Equivalentemente: si cerca $g \in S_n$:

$$\sum_{i=1}^m (y_i - g(x_i))^2 \text{ sia minima}$$

$g \in S_n \Rightarrow g(x) = \sum_{i=1}^n a_i \varphi_i(x) \Rightarrow$ Si cerca $(a_1, a_2, \ldots, a_n)$:

$$F(\underline{a}) := \sum_{i=1}^m \left(y_i - \sum_{j=1}^n a_j \varphi(x_i)\right)^2 \text{ sia minima}$$

Deve essere: $\dfrac{\partial F}{\partial a_k} = 0 \quad k = 1, \dots, n$

$$(*) \quad 2 \sum_{i=1}^{m} \left(y_i - \sum_{j=1}^{n} a_j \varphi_j(x_i) \right) \left(-\varphi_k(x_i) \right) = 0 \quad k = 1, \dots, n$$

(*) Sistema di n equazioni nelle n incognite $a_1, \dots, a_n$

$$\begin{pmatrix} \sum_{i=1}^{m} \varphi_1^2(x_i) & \sum_{i=1}^{m} \varphi_1(x_i)\varphi_2(x_i) & \cdots & \sum_{i=1}^{m} \varphi_1(x_i)\varphi_n(x_i) \\ & \sum_{i=1}^{m} \varphi_2^2(x_i) & \cdots & \sum_{i=1}^{m} \varphi_2(x_i)\varphi_n(x_i) \\ & & \ddots & \\ \text{sym} & & & \sum_{i=1}^{m} \varphi_n^2(x_i) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{m} y_i \varphi_1(x_i) \\ \sum_{i=1}^{m} y_i \varphi_2(x_i) \\ \vdots \\ \sum_{i=1}^{m} y_i \varphi_n(x_i) \end{pmatrix}$$

Stesso risultato (ovviamente!) scrivendo le equazioni normali per il problema di partenza

$$\boxed{\sum_{j=1}^{n} a_j \varphi_j(x_i) = y_i \quad (i = 1, \dots, m)} \Rightarrow \underbrace{\begin{pmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2) \\ \vdots & & & \\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m) \end{pmatrix}}_{A} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}}_{\underline{b}}$$

$$(g(x_i) = y_i)$$