• Sistema delle equazioni normali per l'approssimazione polinomiale ai minimi quadrati

Dimostrazione:

Dire che $a \in \mathbb{R}^{m+1}$ è di minimo (assoluto) per $\phi(a)$ equivale a dire che:

$$\phi(a+b) \geq \phi(a) \quad \forall b \in \mathbb{R}^{m+1}$$

ma

$$\begin{aligned}
\phi(a+b) &= (y - V(a+b),\, y - V(a+b)) \\
&= (y - Va - Vb,\, y - Va - Vb) \\
&= (y - Va,\, y - Va) + (y - Va,\, -Vb) + (-Vb,\, y - Va) + (-Vb,\, -Vb) \\
&= \phi(a) + 2(Va - y,\, Vb) + (Vb,\, Vb) \\
&= \phi(a) + 2(V^t(Va - y),\, b) + (Vb,\, Vb)
\end{aligned}$$

dove abbiamo usato le seguenti proprietà del prodotto scalare in $\mathbb{R}^m$ (per chiarezza indicato con $(u,v)_n$; ricordiamo che $(u,v)_n = u^t v$ interpretando i vettori come vettori-colonna):

1. $(u,v)_n = (v,u)_n \quad u,v,w \in \mathbb{R}^n$
2. $(\alpha u, v)_n = \alpha(u,v)_n \quad \alpha \in \mathbb{R}$
3. $(u+v, w)_n = (u,w)_n + (v,w)_n$
4. $(u, Az)_n = (A^t u, z)_k \quad u \in \mathbb{R}^n,\ z \in \mathbb{R}^k,\ A \in \mathbb{R}^{n \times k}$

In particolare la 4) è stata usata per scrivere

$$\underset{(Va-y,\,Vb)_N}{\underset{||}{(Va - y,\, Vb)}} = \underset{(V^t(Va-y),\,b)_{m+1}}{\underset{||}{(V^t(Va - y),\, b)}}$$

- Dimostriamo per prima l'implicazione "$\Leftarrow$": assumendo che $V^tVa = V^ty$ abbiamo che:

$$V^tVa - V^ty = V^t(Va - y) = 0 \quad \text{e} \quad (V^t(Va-y),\, b) = \underset{\text{vettore nullo in } \mathbb{R}^{m+1}}{\underset{||}{(0, b)}} = 0$$

da cui:

$$\phi(a+b) = \phi(a) + \underset{\sum_{i=1}^{N}(Vb)_i^2 \geq 0}{\underset{||}{(Vb,\, Vb)}} \geq \phi(a) \quad b \in \mathbb{R}^{m+1}$$

- Per dimostrare "$\Rightarrow$", assumiamo che

$$\phi(a+b) \geq \phi(a) \quad \forall b \in \mathbb{R}^{m+1}$$

Allora:

$$\phi(a+b) = \phi(a) + 2(V^t(Va-y),\, b) + (Vb, Vb) \geq \phi(a) \quad \forall b$$

Cioè:

$$2(V^t(Va-y),\, b) + (Vb, Vb) \geq 0 \quad \forall b$$

Prendiamo $b = \varepsilon v$, con v versore (cioè vettore di lunghezza 1, $(v, v) = 1$). Si ha:

$$\begin{aligned} &2(V^t(Va-y),\, \varepsilon v) + (V(\varepsilon v),\, V(\varepsilon v)) = \\ &= 2\varepsilon(V^t(Va-y),\, v) + \varepsilon^2(Vv, Vv) \geq 0 \quad \forall \varepsilon \geq 0 \text{ e } \forall v \end{aligned}$$

Dividendo per $\varepsilon > 0$:

$$2(V^t(Va-y),\, v) + \varepsilon(Vv, Vv) \geq 0 \quad \forall \varepsilon \text{ e } \forall v$$

Per $\varepsilon \to 0$ la disuguaglianza viene mantenuta, ottenendo:

$$(V^t(Va-y),\, v) \geq 0 \quad \forall v$$

Ma se vale $\forall$ versore, possiamo prendere $-v$ al posto di v e otteniamo:

$$(V^t(Va-y),\, -v) = -(V^t(Va-y),\, v) \geq 0 \quad \forall v$$

Cioè $(V^t(Va-y),\, v) \leq 0$, $\forall v$. Siccome $(V^t(Va-y),\, v)$ risulta sia ≥ 0 che ≤ 0, allora è 0:

$$(V^t(Va-y),\, v) = 0 \quad \forall v$$

Ma chi è l'unico vettore ortogonale a tutti i vettori? É il vettore nullo, cioè

$$V^t(Va-y) = 0 \iff V^tVa = V^ty$$

Ovvero a è soluzione del sistema delle equazioni normali.