L'ALGORITMO DI GRAM E SCHMIDT PER LA COSTRUZIONE DI UNA BASE ORTOGONALE

(PLACIDO LONGO 16/11/2018)

Questa nota è dedicata ad un'estensione del classico teorema della base il quale, negli spazi di dimensione finita (e non nulla), garantisce l'esistenza di basi estratte dai sistemi di generatori: questi esistono di certo per effetto dell'ipotesi sulla dimensione. Iniziamo col presentare l'enunciato, che verrà dimostrato nel corso della nota.

DEFINIZIONE: una base costituita da vettori a due a due ortogonali è detta **ORTOGONALE**. È detta **ORTONORMALE** se

<u>tutti i suoi membri sono di norma unitaria (versors).</u>

<u>TEOREMA</u> (<u>Gram-Schmidt</u>): "<u>Ogni spazio euclideo di dimensione finita (non nulla) ha una base ortogonale</u>".

Prima di fornire una dimostrazione di questo lapidario enunciato, che verrà ottenuto per induzione sulla dimensione dello spazio, presenteremo qualche semplice osservazione.

"In effetti, la costruzione, seppure esposta in modo "pulito" mediante il principio di induzione finita, è in realtà un algoritmo che permette, partendo da un sistema di generatori dello spazio, di eliminare quelli dipendenti dagli elementi della base già costruiti e di costruirne degli altri a partire da quelli indipendenti, facendo uso solo di operatori di proiezione su un sistema ortogonale. Ciò è meglio illustrato da un esempio.

IDEA DELL' ALGORITMO:

1) Esaminare in sequenza i generatori dello spazio.
2) Scartare i vettori nulli, eventualmente presenti.
3) Sottrarre ad ogni vettore esaminato la sua proiezione su quelli già scelti, scartandolo, se la sua componente ortogonale così ottenuta è nulla, o inserendola come nuovo elemento della base ortogonale, altrimenti.

Osserviamo che le proiezioni richieste al punto 3) sono relative ad un sistema ortogonale, e si calcolano direttamente con le formule di Eulero e Fourier, senza dover risolvere alcun sistema lineare.

ESEMPIO: Costruire una base ortogonale per $\left\langle \begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix} \right\rangle$.

Si sceglie $(1,1,1)$ come primo elemento della base, perché non nullo.

Poi, si calcola la componente di $(1,2,1)$ ortogonale a $(1,1,1)$ e cioè

$$(1,2,1) - (1,2,1)_{(1,1,1)} = (1,2,1) - \frac{4}{3}(1,1,1) = \left(-\frac{1}{3}, \frac{2}{3}, -\frac{1}{3}\right)$$

Il vettore così trovato è ortogonale a $(1,1,1)$ ed appartiene a $\langle(1,1,1),(1,2,1)\rangle$, e così accade anche per il triplo $(-1,2,-1)$, che utilizzeremo per la base ortogonale al posto di quello originale, per alleggerire i calcoli seguenti.

Osserviamo anche che, dall'ultima equazione precedente, risolta rispetto a $(1,2,1)$, segue subito che $(1,2,1) \in \langle(1,1,1),(-1,2,-1)\rangle$ da cui

$$\langle(1,1,1),(1,2,1)\rangle = \langle(1,1,1),(-1,2,1)\rangle.$$

Consideriamo infine il terzo generatore $(0,1,0)$, e calcoliamone la sua proiezione sullo span dei due già individuati

$$(0,1,0)_{\langle(1,1,1),(-1,2,-1)\rangle} = (0,1,0)_{(1,1,1)} + (0,1,0)_{(-1,2,-1)} =$$

$$= \frac{1}{3}(1,1,1) + \frac{2}{6}(-1,2,-1) = (0,1,0)$$

La componente di $(0,1,0)$ ortogonale a $\langle (1,1,1), (-1,2,-1) \rangle$ è:

$$(0,1,0) - (0,1,0)_{\langle (1,1,1),(-1,2,-1) \rangle} = (0,1,0) - (0,1,0) = 0.$$

Ne segue che $(0,1,0) \in \langle (1,1,1), (-1,2,-1) \rangle = \langle (1,1,1), (1,2,1) \rangle$ (in effetti, è $(0,1,0) = (1,2,1) - (1,1,1)$), e può essere eliminato senza alterare lo span. Una base ortogonale per

$$\langle (1,1,1), (1,2,1), (0,1,0) \rangle$$

è dunque

$$\{ (1,1,1), (-1,2,-1) \}.$$

DIMOSTRAZIONE DEL TEOREMA.

Proveremo, per induzione sul numero dei generatori $k \leq n$, la proprietà P_k seguente:

" Esiste in $X = \langle u_1, .., u_n \rangle$ una base ortogonale $\mathcal{B}_k$ di $\langle u_1, .., u_k \rangle$ ".

Passo iniziale: poiché $X = \langle u_1, .., u_n \rangle \neq \{0\}$, qualcuno dei generatori deve essere non nullo. Sia u_i quello di indice minimo. Allora, per il lemma fondamentale, $\langle u_1, .., u_{i-1}, u_i \rangle = \langle u_i \rangle$ e, se si pone $\mathcal{B}_i = \{u_i\}$, segue subito P_i. È possibile che sia $i > 1$.

Passo induttivo ($P_k \Rightarrow P_{k+1}$): verrà provato per $i \leq k < n$. Poiché P_k è vera, siano $v_1, .., v_h$ i vettori della base ortogonale $\mathcal{B}_k$ di $\langle u_1, .., u_k \rangle$. Si pone allora

$$v_{h+1} = u_{k+1} - (u_{k+1})_{\langle v_1, .., v_h \rangle} = u_{k+1} - \sum_{j=1}^{h} (u_{k+1})_{v_j} .$$

Si osservi che da P_k segue $\langle v_1, \ldots, v_h \rangle = \langle u_1, \ldots, u_k \rangle$, e dunque $v_1, v_2, \ldots, v_h, v_{h+1} \in \langle u_1, u_2, \ldots, u_{k+1} \rangle \subseteq X$. Inoltre, per il teorema della proiezione, v_{h+1} è ortogonale a $\langle v_1, \ldots, v_h \rangle$, ma non è detto che sia <u>non nullo</u>, il che è necessario perché $\{v_1, \ldots, v_{h+1}\}$ sia una base ortogonale. Per scegliere $\mathcal{B}_{k+1}$ verificante P_{k+1} si procede allora così:

- $\boxed{\textbf{se } v_{h+1} = 0,}$ allora $u_{k+1} = (u_{k+1})_{\langle v_1, \ldots, v_h \rangle} \in \langle v_1, \ldots, v_h \rangle$ e, per P_k, tale spazio coincide con $\langle u_1, \ldots, u_k \rangle$, da cui infine $u_{k+1} \in \langle u_1, \ldots, u_k \rangle$.

Dunque, per il lemma fondamentale, $\langle u_1, \ldots, u_{k+1} \rangle = \langle u_1, \ldots, u_k \rangle$ e, di conseguenza, se si pone $\mathcal{B}_{k+1} = \mathcal{B}_k$ segue subito P_{k+1}.

<u>In sostanza</u>, u_{k+1} <u>viene totalmente ignorato</u>;

– $\boxed{\text{se } v_{h+1} \neq 0,}$ allora $\{v_1, v_2, \ldots, v_h, v_{h+1}\}$ è un sistema ortogonale. Per provare che genera $\langle u_1, \ldots, u_{k+1}\rangle$, si osservi che da P_k segue

$$u_1, \ldots, u_k \in \langle v_1, \ldots, v_h\rangle \subseteq \langle v_1, \ldots, v_h, v_{h+1}\rangle.$$

Inoltre, dalla definizione di v_{h+1}, si ha anche

$$u_{k+1} = v_{h+1} + (u_{k+1})_{\langle v_1, \ldots, v_h\rangle} \in \langle v_1, \ldots, v_h, v_{h+1}\rangle$$

da cui $\langle u_1, \ldots, u_{k+1}\rangle \subseteq \langle v_1, \ldots, v_{h+1}\rangle$.

Per provare P_{k+1}, basta dunque porre $\mathcal{B}_{k+1} = \{v_1, \ldots, v_{h+1}\}$.

La tesi P_n segue da P_i, provata direttamente, e dalla catena $P_i \Rightarrow P_{i+1} \Rightarrow \ldots \Rightarrow P_{n-1} \Rightarrow P_n$, ottenuta iterando il passo induttivo da $k = i$ a $k = n-1$. $\blacksquare$

NOTA: un caso particolare in cui $V_{n+1} = 0$ è quello, di scarso interesse pratico, che si verifica quando $u_{k+1} = 0$.
Naturalmente, un generatore nullo può essere soppresso senza alcun bisogno di fare calcoli di sorta: è un caso molto semplice, e riconoscibile, di generatore (in ogni caso) combinazione dei precedenti.

NOTA: un caso meno agevole è invece quello generale, nel quale un generatore è non nullo, ma dipendente dai precedenti. In tal caso, i calcoli svolti per determinare la sua componente ortogonale sono "sprecati": è il prezzo da pagare per avere, nello stesso algoritmo, tanto l'estrazione della base quanto la costruzione della base ortogonale. Per non "buttare via nulla", occorre eseguire preventivamente una riduzione a scala per eliminare i generatori dipendenti, il che è comunque un prezzo da pagare!

NOTA: Se risultasse vantaggioso, si possono sostituire i vettori ortogonali $v_1, \ldots, v_k$ con i loro versori, ottenendo così una **BASE ORTONORMALE**. Si è già visto nel precedente esempio come si possano sostituire alle componenti ortogonali loro multipli convenienti, per agevolare i calcoli.

Come applicazione notevole del risultato precedente, proviamo ora un teorema di esistenza e unicità della proiezione ortogonale su un sottospazio (di dimensione finita). La questione della unicità, già affrontata altrove, verrà richiamata per completezza.

TEOREMA (di esistenza e unicità della proiezione): sia $Y = \langle u_1, \ldots, u_k \rangle$, sottospazio di uno spazio euclideo X. Allora, per ogni $x \in X$ esiste ed è unico $\bar{x} \in Y$ tale che:

$$(x - \bar{x}) v = 0 \quad \forall v \in \langle u_1, \ldots, u_k \rangle .$$

DIM. Sia $e_1, .., e_h$ una base ortogonale di $\langle u_1, .., u_k \rangle$, ottenuta applicando ai generatori $u_1, .., u_k$ l'algoritmo di Gram-Schmidt.
Dal teorema delle proiezioni per i sistemi ortogonali, basta porre

$$\bar{x} = \sum_{i=1}^{h} x_{e_i} = \sum_{i=1}^{h} \frac{x e_i}{|e_i|^2} e_i$$

per ottenere subito $(x - \bar{x}) e_i = 0 \quad \forall i = 1..h$.
Poiché $Y = \langle e_1, .., e_h \rangle$ ed il prodotto scalare è lineare rispetto a ciascuno dei suoi fattori, ne segue che $(x - \bar{x}) v = 0$ $\forall v \in Y$.
La costruzione precedente fornisce però un elemento $\bar{x} \in Y$ tale che $x - \bar{x} \perp Y$. Che sia unico deriva dal fatto che, se $\bar{x}, \bar{\bar{x}} \in Y$ verificano $(x - \bar{x}) v = 0 \quad \forall v \in Y$ e $(x - \bar{\bar{x}}) v = 0 \quad \forall v \in Y$, allora

$$(x-\bar{x})v - (x-\bar{\bar{x}})v = 0 \quad \forall v \in Y$$

da cui, per la linearità del prodotto scalare rispetto al primo fattore, segue

$$(\bar{\bar{x}}-\bar{x})v = 0 \quad \forall v \in Y$$

Poiché $\bar{x}$ e $\bar{\bar{x}}$ appartengono a Y, anche $\bar{\bar{x}}-\bar{x}$ vi appartiene e, scelto $v = \bar{\bar{x}}-\bar{x}$, ne segue $(\bar{\bar{x}}-\bar{x})(\bar{\bar{x}}-\bar{x}) = 0$, da cui $\bar{\bar{x}}-\bar{x}=0$.

$\blacksquare$

DEFINIZIONE La proiezione di x su $\langle u_1,..,u_k\rangle$ è il vettore $x_{\langle u_1,..,u_k\rangle} = \sum_{i=1}^{h} \frac{x e_i}{|e_i|^2} e_i$, ove $\{e_1,..,e_h\}$ è una qualunque base ortogonale di $\langle u_1,..,u_k\rangle$.

Per l'unicità della proiezione, il risultato non dipende dalla base: si può adoperare quella fornita dall'algoritmo, applicato a $\{u_1,..,u_k\}$.

NOTA: Il teorema appena dimostrato garantisce che il sistema lineare avente per incognite le "coordinate" α_j della proiezione

$$\left(x - \sum_{i=1}^{k} \alpha_i u_i\right) v_j = 0 \quad \forall j = 1..k$$

ha certamente soluzioni (teorema di esistenza); inoltre, qualunque soluzione $(\bar{\alpha}_1, .., \bar{\alpha}_k)$ si scelga, il vettore $\sum_1^k \bar{\alpha}_i u_i$ risulterà sempre lo stesso (teorema di unicità): la proiezione! Mentre le coordinate $\alpha_1 .. \alpha_k$ di $x_{<u_1,..,u_k>}$ sono uniche se i vettori $u_1, .., u_k$ sono indipendenti, ci sarebbero infinite scelte possibili per i coefficienti $\alpha_1 .. \alpha_k$ che producono lo <u>stesso</u> risultato $\sum_1^k \alpha_i u_i$, se $u_1, .., u_k$ sono dipendenti. Dunque l'unicità della proiezione non dipende dall'unicità della soluzione del sistema lineare precedente, che invece dipende solo dall'indipendenza dei generatori dello spazio. Il "vecchio metodo" resta comunque un'alternativa valida, in quanto il sistema ha sempre soluzione!

NOTA: è possibile provare l'esistenza di soluzioni per il sistema lineare che ha per soluzioni le "coordinate" della proiezione rispetto ai generatori di X senza far riferimento alla teoria di Gram e Schmidt: chi fosse interessato, troverà una dimostrazione che utilizza le proprietà delle basi in un altro contributo.

NOTA: le "stranezze" del prodotto scalare negli spazi euclidei complessi non sono di nessun ostacolo per l'estensione del teorema a tali spazi; bastano solo minimi aggiustamenti, come lo spostare il vettore su cui è necessario adoperare la linearità al primo fattore, rispetto al quale il prodotto hermitiano è indistinguibile da quello reale.

NOTA: con altri minimi aggiustamenti, l'ortogonalizzazione di Gram e Schmidt funziona ottimamente anche in dimensione infinita. Un caso importante è quello in cui vengano ortogonalizzate le potenze $1, t, t^2, \dots, t^n, \dots$ rispetto al prodotto scalare

$$fg = \int_a^b f(t)g(t)\,dt,$$

o altri ad esso collegati. Le basi ortogonali così ottenute definiscono i **POLINOMI ORTOGONALI** su $[a,b]$, con le loro applicazioni importanti, ad esempio, alla teoria dell'integrazione numerica (che funziona egregiamente anche nei casi nei quali sarebbe impossibile il calcolo delle primitive). I riferimenti d'obbligo sono i testi d'Analisi Numerica, e, in particolare, il capitolo sulle "formule di quadratura" (integrazione).

MORALE: "Se $u_1, \dots, u_n$ sono indipendenti, una base ortogonale si costruisce ponendo $v_1 = u_1$ e $v_{k+1} = u_{k+1} - \sum_{i=1}^{k} \frac{u_{k+1} v_i}{|v_i|^2} v_i$. Se sono dipendenti, ignorare gli u_{k+1} per cui $v_{k+1} = 0$".

NOTA BIBLIOGRAFICA

Gli studiosi di Analisi Numerica amano molto poco l'algoritmo di Gram-Schmidt per ragioni ottime, dal loro punto di vista. In sostanza, il fatto che le operazioni su ogni nuovo elemento della base coinvolgano tutti gli elementi precedenti fa sì che gli eventuali errori si propaghino selvaggiamente: un errore sul primo produce effetti sul secondo, che già deve combattere con i propri; gli errori di entrambi intervengono nel calcolo del terzo, che ha già i propri, e così via... Esistono algoritmi che riducono tali effetti (SVD, "Singular Values Decomposition"), ogni approfondimento sui quali va ricercato nei testi di Analisi Numerica come, ad esempio, il "Numerical Recipes in C++" di PRESS, VETTERLING, FLANNERY e TEUCHOLSKIJ, dedicato agli aspetti applicativi.