Salvatore Cherubino

Risoluzione senza determinanti dei sistemi lineari di equazioni.

Bollettino dell'Unione Matematica Italiana, Serie 3, Vol. **7** (1952), n.1, p. 54–59.

Zanichelli

<http://www.bdim.eu/item?id=BUMI_1952_3_7_1_54_0>

SEZIONE STORICO-DIDATTICA

Risoluzione senza determinanti dei sistemi lineari di equazioni.

Nota di SALVATORE CHERUBINO (a Pisa).

Sunto. - *Sistemazione della teoria generale e risoluzione dei sistemi di equazioni lineari ottenuta col simbolismo elementare delle matrici.*

In due note apparse su questo Bollettino, G. AQUARO [1] ed F. SBRANA [2] hanno dimostrato, senza determinanti, la risolubilità di un sistema lineare di n equazioni ad n incognite e termini noti arbitrari, quando la matrice dei coefficienti abbia le righe oppure le colonne linearmente indipendenti.

Poichè entrambi gli Autori si appoggiano sopra un teorema di dimostrazione piuttosto faticosa e su un processo di ortonormalizzazione non poco laborioso, mi sembra utile mostrare come il simbolismo delle matrici (non il calcolo di matrici) consenta di dare insieme la sistemazione della teoria generale dei sistemi lineari di equazioni e la soluzione di essi con procedimento abbastanza più rapido del processo di E. SCHMIDT [3] al quale entrambe le

[1] s. III, a. IV, n. 3 (settembre 1951) pp. 240-245.

[2] Ibidem, n. 4 (dicembre 1951) pp. 315-317. Nella semplificazione di F. SBRANA il teorema di G. AQUARO può essere sostituito dal fatto che la matrice dei vettori v, giusta quanto qui mostriamo, è a righe e colonne linearmente indipendenti.

[3] *Ueber die Auflösuny linearer Gleîchnngen mit unendlich vielen unbekannten* [« Rend. Pal. », t. XV (1° semestre 1908), pp. 53-77], cap. I, § 5, pag. 61. Il simbolismo vettoriale ivi impiegato non fa riconoscere la laboriosità del processo e rispetto a quello delle matrici riesce, nel calcolo numerico, scarsamente operativo

note mentovate fanno ricorso. Forse è anche opportuno rilevare che, salvo la forma, quest' ultimo si identifica con la diagonalizzazione di una matrice (hermisiana *o*) simmetrica definita, cioè con la riduzione a forma canonica di una forma (hermisiana *o*) quadratica, che si espone e si esegue, pur essa agevolmente, col simbolismo elementare delle matrici [4]. Osserviamo inoltre che il nostro metodo di risoluzione può dirsi essere sostanzialmente quello in uso prima di CRAMER e quello tuttora adoperato dai pratici, nel calcolo numerico.

1. Sia A una matrice complessa di ordine n, b_{-1} una colonna di numeri complessi $b_1, b_2, \dots, b_n$, x_{-1} un' altra colonna di n indeterminate $x_1, x_2, \dots x_n$. La scrittura

$$Ax_{-1} = b_{-1}, \tag{1}$$

rappresenta in matrici il sistema lineare di coefficienti gli elementi di A, termini noti i numeri b_r, incognite le x_r: le righe di A danno i coefficienti delle n equazioni.

Se le colonne di A sono tutte linearmente indipendenti quindi non nulle, si potrà cambiare l'ordine delle equazioni, cioè delle righe di A, in modo che la prima colonna di A presenti al primo posto un elemento $a_1 \neq 0$ e perciò si possa scrivere

$$A = \left(\begin{array}{c|c} a_1 & b \\ \hline c_{-1} & d \end{array}\right)$$

con d matrice di ordine $n-1$, b una riga e c_{-1} una colonna di $n-1$ numeri. Prendendo poi

$$H_1 = \left(\begin{array}{c|c} 1 & 0 \\ \hline -\lambda_{-1} & I_{n-1} \end{array}\right)$$

dove I_{n-1} è la matrice identica di ordine $n-1$, $\lambda_{-1} = a_1^{-1} \cdot c_{-1}$, si ha:

$$A_1 = H_1 A = \left(\begin{array}{c|c} a_1 & b \\ \hline 0 & d^* \end{array}\right), \quad d^* = d - \lambda_{-1} b.$$

Dal sistema (1) si passa così all' altro

$$A_1 x_{-1} = b'_{-1}, \qquad b'_{-1} = H_1 b_{-1} \tag{2}$$

([4]) Vedi le mie *Lezioni di Geometria Analitica....* [Roma, Dante Alighieri (1940)], p. II, cap. VI, § 1, n. 150, p. 161. Trattandosi di matrici definite il procedimento è alquanto più semplice. Il metodo di risoluzione esposto qui appresso richiede un minor numero di operazioni, però analoghe a quelle del processo di diagonalizzazione predetto.

che equivale ad (1) perchè da esso si ritorna al primo moltiplicando ambo i membri, a sinistra, per l'inversa di H_1^{-1} cioè per

$$H_1^{-1} = \left(\begin{array}{c|c} 1 & 0 \\ \hline \lambda_{-1} & I_{n-1} \end{array}\right).$$

Anche A_1 è a colonne linearmente indipendenti, perchè se fosse, per $y_{-1} \neq 0$, $A_1 y_{-1} = 0$, si avrebbe $H_1 z_{-1} = 0$ con $z_{-1} = Ay_{-1}$: e poichè dalla invertibilità di H_1 si ha necessariamente $z_{-1} = 0$, le colonne di A non sarebbero linearmente indipendenti. Segue che le colonne di d^* sono pure indipendenti. Infatti, se la riga di $n-1$ elementi $x' \neq 0$ desse $d^* x'_{-1} = 0$, per $y = (\alpha \mid x')$, con $\alpha = - a_1^{-1} b x'_{-1}$, si avrebbe:

$$A_1 y_{-1} = 0$$

il che si è visto impossibile essendo gli elementi di y non tutti zero.

Dunque, su d^* si può operare come su A, ottenendo, per una K invertibile opportuna [1]:

$$Kd^* = \left(\begin{array}{c|c} a_2 & b^* \\ \hline 0 & d^{**} \end{array}\right),$$

ove $a_2 \neq 0$ e d^{**} è a colonne linearmente indipendenti. Dopo di che si moltiplicherà a sinistra A_1 per la matrice

$$H_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & K \end{array}\right)$$

che anch' essa è invertibile, perchè lo è K. Così seguitando si ottiene la successione di matrici:

$$A_1 = H_1 A, \quad A_2 = H_2 A_1, \ldots, \quad A_n = H_n A_{n-1} = T \tag{3}$$

di cui l'ultima è triangolare alta invertibile, cioè ha tutti gli elementi al disotto della diagonale principale eguali a zero, quelli diagonali essendo diversi da zero. Il sistema proposto equivale, successivamente, a quelli ottenuti sostituendo A con le matrici della successione (3) e b_{-1} con i suoi prodotti, a sinistra, per $H_1, H_2, \ldots, H_n$, successivamente. L'ultimo sistema così ottenuto

$$Tx_{-1} = c_{-1} \tag{4}$$

(1) In K è compendiato anche l'eventuale cambiamento di ordine delle righe di d^*, cioè K è prodotto di due matrici invertibili, ma come la H_1, l'altra ottenuta da I_{n-1} con lo stesso scambio di righe che si vuole effettuare su d^*; la seconda è simmetrica ed involutoria.

ci dà immediatamente x_n e, l' uno dopo l' altra, per sostituzione, le incognite $x_{n-1}, \dots, x_1$, univocamente.

Ecco così risoluto il sistema proposto e dimostrata l' univocità della soluzione. Si trova subito che A ha linearmente indipendenti anche le righe. Infatti, il prodotto $H_n H_{n-1} \dots H_1 = H$ è invertibile, perciò si ha

$$HA = T, \qquad A = H^{-1}T, \qquad c_{-1} = Hb_{-1}$$

e poichè, come si è detto (e come subito si vede) la matrice triangolare T è invertibile, è tale anche A (ha per inversa $T^{-1}H$), quindi è a (colonne e) righe linearmente indipendenti.

I risultati ora ottenuti valgono anche pel sistema trasposto di (1), cioè quando A è a righe linearmente indipendenti e si considera l' equazione

$$(1)' \qquad xA = b,$$

nel quale A risulterà a colonne anch' esse linearmente indipendenti.

2. Il procedimento sopra indicato può effettuarsi anche se A è rettangolare, ad es. ad m righe ed n colonne. Se $n > m$, solo al più m colonne di A, ad es. le prime m, sono indipendenti perchè, avutone m, ogni altra colonna di A, risolvendo un sistema come (1), riesce combinazione lineare di queste m. Analogamente, se $m > n$, il numero delle righe indipendenti di A non supera il numero n delle sue colonne. E se le colonne (le righe) indipendenti di A sono quante le righe (le colonne) queste sono tutte linearmente indipendenti, essendo tali le parti di esse che sono contenute nelle colonne (nelle righe) supposte indipendenti.

Supponiamo che solo m', $0 < m' < m$ righe di A siano linearmente indipendenti, ad es. le prime m'. Allora potrà scriversi

$$(5) \qquad A = \left(\frac{B}{\mu B}\right)$$

dove B è ad m' righe tutte indipendenti e le righe della matrice μ danno coi loro termini le combinazioni lineari da eseguire sulle righe di B per avere le rimanenti righe di A.

Indicando con b'_{-1} la colonna dei primi m' elementi di b_{-1}, con b''_{-1} quella dei rimanenti, la (1) ci dà [1]

$$(6) \qquad Bx_{-1} = b'_{-1} \qquad\qquad (7) \quad \mu b'_{-1} = b''_{-1}$$

[1] Gli elementi di b'_{-1} corrispondono alle righe indipendenti di A.

Quest' nltima è la condizione necessaria e sufficiente per la risolubilità del dato sistema. Se essa è soddisfatta basta risolvere la (6).

Per quel che si è detto poco fa, le colonne di B sono almeno m' ed esattamente m' di esse sono linearmente indipendenti. Se $n = m'$, il sistema (6) ammette una e una sola soluzione, che si otterrà come al n. 1.

Se $n > m'$, posto che le m' colonne indipendenti siano ad es., le prime e indicando con C la matrice da esse formata, si potrà scrivere

$$B = (C \mid C\nu) \tag{8}$$

dove ν è la matrice le cui colonne danno le combinazioni lineari da eseguire sulle colonne di C per avere le rimanenti di B.

Dopo di ciò, separando nella colonna x i primi m' elementi dai rimanenti con lo scrivere $x = (x' \mid x'')$, si ha [1]:

$$Cx'_{-1} + C\nu x''_{-1} = C(x''_{-1} + \nu x''_{-1}) = b'_{-1}. \tag{9}$$

Basta quindi risolvere l' equazione

$$Cz_{-1} = b'_{-1} \tag{10}$$

e poi prendere

$$x'_{-1} = z_{-1} - \nu x''_{-1} \tag{11}$$

lasciando arbitrari gli $n - m'$ elementi di x''.

La (10) è come la (1), quindi è risolubile univocamente; perciò il sistema dato ha $\infty^{n-m'}$ soluzioni, purchè sia soddisfatta la (7).

La teoria generale dei sistemi lineari è così completamente stabilita.

3. La invertibilità di A segue dalla risolubilità della (1) con b_{-1} arbitraria e dalla indipendenza lineare delle sue righe. Infatti, b_{-1} può coincidere con le colonne della matrice identica di ordine n ed x darà, ordinatamente, quelle di una matrice X tale che

$$AX = I. \tag{12}$$

Da questa si ha $AXA = A$, quindi

$$A(XA - I) = 0. \tag{13}$$

Poichè le colonne di A sono indipendenti, questa ci dà $XA = I$.

[1] Gli elementi di x' corrispondono alle colonne indipendenti di B.

Dunque X è l'inversa di A da ambo i lati e può indicarsi con A^{-1}.

Segue, sempre senza calcolo, il teorema sul quale sono basate le due note di AQUARO e SBRANA. Se A è ortogonale a sinistra [2], ossia se $\bar{A}_{-1}A = D$, con D matrice diagonale ad elementi diagonali tutti > 0, indicando con $D^{\frac{1}{2}}$ la matrice diagonale delle radici quadrate (ad es. tutte positive) degli elementi diagonali di D, si ha

$$D^{-\frac{1}{2}}\bar{A}_{-1} \cdot AD^{-\frac{1}{2}} = I. \tag{14}$$

Per quel che ora si è visto, in questa (14) possono invertirsi i due fattori e poi trasporre il prodotto ottenendo:

$$(D^{-\frac{1}{2}}\bar{A}_{-1})_{-1} \cdot (D^{-\frac{1}{2}}A_{-1}) = I; \tag{15}$$

cioè $D^{-\frac{1}{2}}A_{-1}$ è ortonormale a sinistra. Quindi: « Se un sistema di n vettori (le colonne di A) è ortogonale, il sistema trasposto (quello delle colonne di A_{-1}) ottenuto alterando ciascuno dei primitivi vettori di un opportuno fattore scalare reale (gli elementi diagonali di $D^{-\frac{1}{2}}$) risulta ortonormale » [3].