Eliminazione di Gauss come cambiamento di base: fattorizzazione LU

Le operazioni che definiscono l'algoritmo della *eliminazione di Gauss* possono essere viste come le operazioni corrispondenti ad un particolare cambiamento di base.

Si consideri di nuovo l'isomorfismo

$$A : \mathcal{U} \to \mathcal{V}$$

e la sua matrice $\mathbf{A}$ in corrispondenza di una base $\{b_1 \dots b_n\}$ di $\mathcal{U}$ e una base $\{d_1 \dots d_n\}$ di $\mathcal{V}$. Si modifichi la base di $\mathcal{V}$ sostituendo il vettore d_1 con il vettore, parallelo a Ab_1,

$$\acute{d}_1 := d_1 + c_{21} d_2 + \ldots + c_{n1} d_n$$

essendo

$$\begin{aligned} c_{21} &:= a_{21}/a_{11} \\ c_{31} &:= a_{31}/a_{11} \\ &\dots \\ c_{n1} &:= a_{n1}/a_{11} \end{aligned}$$

e avendo supposto che Ab_1 non appartenga al sottospazio generato dai vettori $\{d_2 \dots d_n\}$, ovvero che sia $a_{11} \neq 0$ (altrimenti si otterrebbe, invece di una nuova base, una famiglia di vettori dipendenti).

Si osservi che per un generico vettore $v \in \mathcal{V}$ in corrispondenza delle due basi si ha

$$\begin{aligned} v &= v_1 d_1 + v_2 d_2 + \ldots + v_n d_n \\ v &= \acute{v}_1 \acute{d}_1 + \acute{v}_2 d_2 + \ldots + \acute{v}_n d_n \\ v &= \acute{v}_1 d_1 + (\acute{v}_2 + c_{21}\acute{v}_1) d_2 + \ldots + (\acute{v}_n + c_{n1}\acute{v}_1) d_n \end{aligned}$$

Le componenti nelle due basi risultano dunque tali che

$$\begin{aligned} \acute{v}_1 &= v_1 \\ \acute{v}_2 &= v_2 - c_{21} v_1 \\ &\dots \\ \acute{v}_n &= v_n - c_{n1} v_1 \end{aligned}$$

Poiché le colonne della matrice $\mathbf{A}$ sono costituite dalle componenti dei vettori Ab_i, la matrice corrispondente alla base appena definita si ottiene sostituendo i termini di ciascuna riga, dalla seconda in poi, con le espressioni seguenti

$$\begin{array}{llll} a^{(1)}_{21} := 0 & a^{(1)}_{22} := a_{22} - c_{21} a_{12} & \dots & a^{(1)}_{2n} := a_{2n} - c_{21} a_{1n} \\ a^{(1)}_{31} := 0 & a^{(1)}_{32} := a_{32} - c_{31} a_{12} & \dots & a^{(1)}_{3n} := a_{3n} - c_{31} a_{1n} \\ \dots & & & \\ a^{(1)}_{n1} := 0 & a^{(1)}_{n2} := a_{n2} - c_{n1} a_{12} & \dots & a^{(1)}_{nn} := a_{nn} - c_{n1} a_{1n} \end{array}$$

Le operazioni che conducono alla nuova matrice di A appena descritte sono proprio quelle definite dalla *eliminazione di Gauss* nel primo dei cicli più esterni.

Le successive operazioni dello stesso algoritmo consistono nella costruzione delle matrici di A corrispondenti a nuove basi ottenute nel modo seguente.

Si modifichi la base di $\mathcal{V}$ sostituendo d_2 con il vettore

$$\acute{d}_2 := d_2 + c_{32}d_3 + \ldots + c_{n2}d_n$$

essendo

$$\begin{aligned} c_{32} &:= a_{32}^{(1)}/a_{22}^{(1)} \\ c_{42} &:= a_{42}^{(1)}/a_{22}^{(1)} \\ &\ldots \\ c_{n2} &:= a_{n2}^{(1)}/a_{22}^{(1)} \end{aligned}$$

e avendo supposto che sia $a_{22}^{(1)} \neq 0$. La famiglia di vettori così ottenuta è ancora una base.

Proseguendo con la sostituzione dei successivi vettori d_i con i vettori

$$\acute{d}_i := d_i + c_{i+1,i}d_{i+1} + \ldots + c_{ni}d_n$$

si ottiene infine una base di $\mathcal{V}$ tale che la matrice di A risulta essere triangolare superiore.

La matrice ottenuta è dunque la matrice $\mathbf{U}$ prodotta dall'algoritmo della *eliminazione di Gauss*, poiché le operazioni eseguite sugli elementi della matrice originaria sono proprio quelle definite da quell'algoritmo.

La matrice le cui colonne sono costituite dalle componenti dei vettori $\acute{d}_i$ della nuova base rispetto ai vettori d_i della base originaria risulta

$$\begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ c_{21} & 1 & 0 & \ldots & 0 \\ c_{31} & c_{32} & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & c_{n3} & \ldots & 1 \end{pmatrix}$$

Indicando con $\mathbf{L}$ tale matrice, la relazione tra le matrici di A nelle due basi si scrive

$$\mathbf{A} = \mathbf{LU}$$

Si noti che la matrice $\mathbf{L}$ risulta essere in ogni caso *triangolare inferiore* con gli elementi della diagonale uguali ad 1.

Riferendosi al primo dei successivi cambiamenti di base, si noti che se $a_{11} = 0$ non risulta più definita la espressione data per $\acute{d}_1$.

Si può in tal caso modificare l'ordine dei vettori della base di $\mathcal{V}$ in modo che a_{11}, componente di Ab_1 sul primo di tali vettori, risulti non nullo.

Si noti che ad un riordinamento dei vettori della base di $\mathcal{V}$ corrisponde un riordinamento delle componenti dei vettori in $\mathcal{V}$. Ne consegue un riordinamento dei termini di ciascuna colonna di $\mathbf{A}$ e delle colonne di $\mathbf{L}$. Tale riordinamento è esattamente quello realizzato nella *eliminazione di Gauss* con lo scambio di righe.

In corrispondenza di un cambiamento di base consistente in un semplice riordinamento dei vettori, la matrice le cui colonne sono costituite dalle componenti dei vettori della nuova base rispetto alla base originaria si ottiene dalla matrice identità assegnando alle colonne l'ordine corrispondente al nuovo ordinamento dei vettori.

Indicando con $\mathbf{B}$ la matrice che risulta definita da tutti gli scambi effettuati durante l'esecuzione dell'algoritmo della *eliminazione di Gauss*, si può vedere la matrice di A nella nuova base come ottenuta attraverso il solo riordinamento dei vettori definito da $\mathbf{B}$, seguito dalla sostituzione di tali vettori, dal primo in poi secondo il nuovo ordine, con le combinazioni lineari

$$\acute{d}_i := d_i + c_{i+1,i} d_{i+1} + \ldots + c_{ni} d_n$$

La relazione tra le matrici di A nelle basi iniziale e finale può pertanto essere scritta

$$\mathbf{A} = \mathbf{BLU}$$

oppure, ponendo $\mathbf{P} := \mathbf{B}^{-1}$,

$$\mathbf{PA} = \mathbf{LU}$$

L'algoritmo che restituisce, oltre alla matrice $\mathbf{U}$, anche la matrice $\mathbf{L}$ si può ottenere direttamente dalla *eliminazione di Gauss* assegnando i valori dei coefficienti a_{ij}/a_{jj} agli elementi dell'*array* `a` con indici `i>j`, con la seguente modifica del ciclo più interno

```
for i:=j+1 to n do
begin
  c := a[i,j]/a[j,j];
  a[i,j] := c;
  for k:=j+1 to n do  a[i,k] := a[i,k] - c * a[j,k];
  b[i] := b[i] - c * b[j];
end;
```

Non è invece conveniente rappresentare la matrice $\mathbf{P}$ con un *array* bidimensionale. La informazione riguardante gli scambi di righe eseguiti può più semplicemente essere rappresentata attraverso un *array* `p` di interi il cui valore iniziale è definito da

```
for i:=1 to n do p[i]:=i;
```

e che in corrispondenza di uno scambio della riga m con la riga j ($j < m$) viene modificato così

```
p[j]:=m;
```

L'algoritmo così definito si chiama *fattorizzazione LU*, dalla espressione $\mathbf{PA} = \mathbf{LU}$.

La soluzione di un sistema di equazioni lineari

$$\mathbf{Ax} = \mathbf{b}$$

si può ottenere calcolando per prima cosa il vettore che corrisponde a $\mathbf{b}$ nella nuova base, valutando successivamente

$$\mathbf{z} := \mathbf{Pb}$$
$$\mathbf{y} := \mathbf{L}^{-1}\mathbf{z}$$

La valutazione di $\mathbf{z}$ consiste nell'applicare alla *n–pla* $\mathbf{b}$ le permutazioni definite dall'*array* `p` nel seguente modo

```
for j:=1 to n-1 do if p[j]>j then
begin
   m:=p[j]; s:=b[j]; b[j]:=b[m]; b[m]:=s;
end;
```

La valutazione di $\mathbf{y}$ consiste nell'eseguire la *sostituzione in avanti*

```
for i:=1 to n do
begin
     s:=0; for k:=1 to i-1 do  s := s + a[i,k] * b[k];
     b[i] := b[i] - s;
end;
```

conservando $\mathbf{z}$ nello stesso *array* `b`.
Essendo infine

$$\mathbf{x} := \mathbf{U}^{-1}\mathbf{y}$$

la *n–pla* $\mathbf{x}$ può essere valutata attraverso la *sostituzione all'indietro* già definita.