PIEMONTE
FONDI STRUTTURALI E DI INVESTIMENTO EUROPEI 2014/2020

fondo sociale europeo

Elementi di Programmazione con Python e Analisi dei Dati
Lezione 2: perché Numpy, alcuni richiami di algebra

Stefano Andreozzi, PhD

Formazione continua individuale – Id Attività: 2530775 Codice Corso: B341-1-2019-0

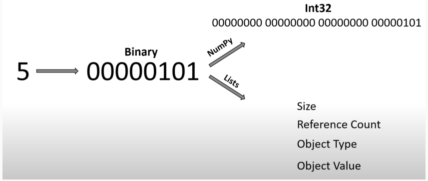
25 novembre 2020

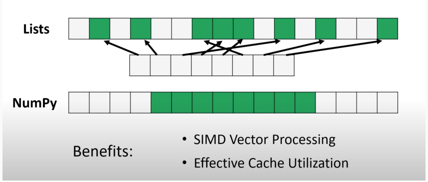
Accademia FormaWeb

UNIONE EUROPEA
REGIONE PIEMONTE
per una crescita intelligente, sostenibile ed inclusiva
www.regione.piemonte.it/europa2020
INIZIATIVA CO-FINANZIATA CON FSE

Performance delle liste

- oggetto lista consiste in due elementi interni: header, vettore di puntatori a indirizzi di memoria all'oggetto (eventualmente riallocabile)
- la dimensione della lista in memoria dipende dal numero di oggetti nella lista, non dalla dimensione degli oggetti puntati
- il tempo per ricavare o settare un singolo elemento è costante (O(1))
- "tempo ammortizzato costante", cioè ad ogni operazione di append viene già allocata più memoria del necessario, per evitare riallocazioni ad ogni chiamata successiva (ok se l'allocatore in memoria è veloce, se no c'è prevalenza di allocazione su scrittura, O(n*n))
- inserimento/rimozione di un elemento nella lista dipende dalla lunghezza e posizione (quanti elementi precedenti ci sono, O(n)); inserimento/rimozione in testa lento, in coda veloce
- inversione proporzionale alla dimensione (O(n))
- sorting dipende dall'algoritmo (caso peggiore O(n log n), ma casi tipici sono di solito migliori





list

a=[1,3,5]

b=[1,2,3]

a*b=ERROR

np.array

a=np.array([1,3,5])

b=np.array([1,2,3])

a*b=np.array([1,6,15])

Risorse su numpy

Leonardo Primavera

- Moduli numpy e matplotlib: introduzione alla libreria di calcolo numerico numpy e alla libreria grafica matplotlib.
- Statistica in python: introduzione all'analisi statistica in python, costruzione di un istogramma e di una PDF, momenti di una PDF, best-fits.

Nicolas P. Rouger

- The Game of Life

Christopher Pease

- An Overview of Monte Carlo Methods

Kelvin Salton do Prado (*solo approfondimento*)

- Understanding time complexity with Python examples

$\underline{v} = (v_1, v_2, \dots, v_n)$	Vettore
$\|\|\underline{v}\|\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$	Modulo o norma
$\begin{bmatrix} v_1 \\ v_2 \\ \dots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_n \end{bmatrix} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \dots \\ v_n + w_n \end{bmatrix}$	Somma tra vettori
$k\underline{v} = (kv_1, kv_2, \dots, kv_n)$	Prodotto scalare-vettore
$\underline{v} \cdot \underline{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n$	Prodotto scalare
$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} \times \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} y_1 z_2 - z_1 y_2 \\ z_1 x_2 - x_1 z_2 \\ x_1 y_2 - y_1 x_2 \end{bmatrix}$	Prodotto vettoriale
$\|\|\underline{v} + \underline{w}\|\|^2 = \|\|\underline{v}\|\|^2 + \|\|\underline{w}\|\|^2$	Se $v \perp w$
$\cos\theta = \frac{\underline{v} \cdot \underline{w}}{\|\|\underline{v}\|\| \cdot \|\|\underline{w}\|\|}$	Angolo tra vettori
$\underline{v} \perp \underline{w} \Leftrightarrow \underline{v} \cdot \underline{w} = 0$	Vettori perpendicolari
$\underline{v} \parallel \underline{w} \Leftrightarrow \underline{v} \times \underline{w} = \underline{0}$	Vettori paralleli

Combinazione lineare:

$\underline{w}$ è combinazione lineare di $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ sse $\exists a_1, a_2, \ldots, a_n$ tali che $\underline{w} = a_1\underline{v}_1 + a_2\underline{v}_2 + \ldots + a_n\underline{v}_n$

$\underline{0}$ è combinazione lineare di un qualunque insieme di vettori.

Dipendenza lineare:

$\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ si dicono *linearmente dipendenti* sse $\exists a_1, a_2, \ldots, a_n$ non tutti nulli t.c. $a_1\underline{v}_1 + a_2\underline{v}_2 + \ldots + a_n\underline{v}_n = \underline{0}$

$\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ si dicono *linearmente indipendenti* se non sono linearmente dipendenti.

Se un insieme di vettori è linearmente indipendente $a_1 = a_2 = \ldots = a_n = 0$.

$$v_1 = (1, -3, 7) \quad v_2 = (2, -1, -1) \quad v_3 = (-4, 2, 2)$$

$$\begin{cases} x + 2y - 4z = 0 \\ -3x - y + 2z = 0 \\ 7x - y + 2z = 0 \end{cases} \rightarrow \begin{cases} x = 0 \\ y = 2z \\ 0 = 0 \end{cases} \rightarrow t \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \forall t \in \mathbb{R} \text{ sono l.d.}$$

$$0v_1 + 2tv_2 + tv_3 = 0$$

Se l'unica soluzione fosse stata $(0, 0, 0)$ allora sarebbero l.i.

Proprietà del prodotto scalare-vettore e del prodotto scalare

Prodotto scalare-vettore:

$$(t_1 + t_2)\underline{v} = t_1\underline{v} + t_2\underline{v} \qquad (t_1 t_2)\underline{v} = t_1(t_2\underline{v})$$

$$t(\underline{v}_1\underline{v}_2) = t\underline{v}_1 + t\underline{v}_2 \qquad 1\underline{v} = \underline{v}$$

Prodotto scalare:

$$\langle \underline{v}, \underline{w} \rangle = \langle \underline{w}, \underline{v} \rangle$$

$$\langle t\underline{v}, t\underline{w} \rangle = t\langle \underline{v}, \underline{w} \rangle$$

$$\langle \underline{v}_1 + \underline{v}_2, \underline{w} \rangle = \langle \underline{v}_1, \underline{w} \rangle + \langle \underline{v}_2, \underline{w} \rangle$$

$$\langle t_1\underline{v}_1 + t_2\underline{v}_2, \underline{w} \rangle = t_1\langle \underline{v}_1, \underline{w} \rangle + t_2\langle \underline{v}_2, \underline{w} \rangle$$

$$\langle \underline{v}, \underline{v} \rangle \geq 0$$

$\langle \underline{u}, \underline{v} \times \underline{w} \rangle = 0$ sse sono l.d.

Un insieme di oggetti in cui è definita l'operazione di somma e l'operazione prodotto scalare-vettore che rispetta le proprietà indicate precedentemente viene definito *spazio vettoriale*.

Esempi di spazi vettoriali: vettori liberi, matrici, polinomi, funzioni lineari.

$\mathcal{L}(\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n)$ è lo spazio vettoriale generato dalla combinazione lineare di $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$.

Un insieme di vettori è un generatore sse ogni vettore dello spazio è combinazione lineare del generatore.

Una base è un generatore formato da vettori l.i.

$\dim V$ = numero di vettori di una base qualunque. $\dim\{\underline{0}\} = 0$.

Esempio

$V = \mathcal{L}(\begin{bmatrix} 1 & 1 \end{bmatrix}) \quad W = \mathcal{L}(\begin{bmatrix} 1 & 0 \end{bmatrix})$

$V + W = \mathcal{L}([1, 1], [1, 0])$ sono l.i. quindi $|V + W| = 2$

Per Grassmann $|V \cap W| = 0$

$U \subset V$ U è sottospazio di V se è uno spazio vettoriale rispetto le operazioni di V.

$\underline{u}_1 + \underline{u}_2 \in U \quad k\underline{u} \in U \quad \underline{0} \in U \quad -\underline{u} \in U$

Operazioni

U, W sottospazi di $V \Rightarrow U \cap W$ è sottospazio di V.

$U + W = \{\underline{v} \in V | \underline{v} = \underline{u} + \underline{w}, \underline{u} \in U, \underline{w} \in W\}$ sottospazio di V.

$|U + W| = |U| + |W| - |U \cap W|$ Formula di Grassmann

$U \oplus W : \forall \underline{v} \in V\ \exists \underline{u} \in U, \underline{w} \in W$ t.c. $\underline{v} = \underline{u} + \underline{w}$ in modo unico

$\begin{cases} U \cap W = \{0\} \\ U + W = V \end{cases}$ W è generato dai vettori che aggiunti alla base di U formano una base di V.

Trovare una base

Se $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n$ sono generatori, la base è un sottoinsieme l.i. di quei vettori. Rimuovo i vettori nulli, da sinistra tolgo i vettori linearmente dipendenti dai precedenti presi. Se non sono generatori accosto a destra i vettori di una base qualunque e ottengo un insieme di generatori.

Verifica di sottospazio

$$S = \{\underline{v} = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$$

$$\underline{v} = (x_1, x_2, x_3) \qquad \underline{w} = (y_1, y_2, y_3)$$

$$\underline{0} \in S \quad 0 + 0 + 0 = 0$$

$$\underline{v} + \underline{w} = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$$

$$(x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) =$$

$$= (x_1 + x_2 + x_3) + (y_1 + y_2 + y_3) = 0$$

$$\lambda\underline{v} = (\lambda x_1, \lambda x_2, \lambda x_3)$$

$$\lambda x_1 + \lambda x_2 + \lambda x_3 = \lambda(x_1 + x_2 + x_3) = 0$$

$A_{m,n} \cdot B_{n,p} = C_{m,p}$

$C_{i,j} = \sum_{k=1}^{n} a_{ik} b_{kj}$ (i-esima riga di A $\cdot$ j-esima colonna di B)

Non commutativa	$AB \neq BA$
Associativa	$A(BC) = (AB)C$
Elemento neutro	$AI_n = I_nA = A$
Distributiva	$A(B+C) = AB + AC$
Trasposta	$A^T = [\alpha_{ij}] \quad \alpha_{ij} = a_{ji}$
Inversa	$AA^{-1} = A^{-1}A = I_n \quad (\exists \text{ sse } \det A \neq 0)$
	$(AB)^{-1} = B^{-1}A^{-1}$

Prodotto tra matrici

$$\begin{bmatrix} 1 & 4 \\ 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 & 4 \\ 3 & 5 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 21 & 4 & 16 \\ 13 & 17 & 3 & 17 \\ 8 & 11 & 2 & 10 \end{bmatrix}$$

Inversa della matrice

$$A|I_2 = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 5 & 2 & -1 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 2/5 & -1/5 \end{array}\right] \to \left[\begin{array}{cc|cc} 1 & 0 & 1/5 & 2/5 \\ 0 & 1 & 2/5 & -1/5 \end{array}\right] \to A^{-1} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$$

Una matrice è a scala se $\underline{R}_i = 0 \Rightarrow \underline{R}_{i+1} = 0$ e le posizioni dei primi elementi non nulli (*pivot*) sono strettamente crescenti.

$\mathrm{rk} A$ = numero di *pivot* di A ridotta a scala.

Mosse di Gauss:

- Moltiplicare una riga per $k \neq 0$
- Sostituire una riga con la c.l. di due righe
- Scambiare due righe

$[A|I_n] \to [I_n|A^{-1}]$ riducendo a scala

$A^{-1} = \left(1/\det A\right) \, [b_{ik}] \quad b_{ik} = C_{ki}$ (compl. algebrico trasposto)

$$A = \left[\begin{array}{c|c|c|c} \underline{C}_1 & \underline{C}_2 & \cdots & \underline{C}_n \end{array}\right] = \left[\begin{array}{c} \underline{R}_1 \\ \hline \vdots \\ \hline \underline{R}_m \end{array}\right]$$

$\mathrm{Row} A = \mathcal{L}(\underline{R}_1, \underline{R}_2, \dots, \underline{R}_m) \quad \mathrm{Col} A = \mathcal{L}(\underline{C}_1, \underline{C}_2, \dots, \underline{C}_n)$

$\dim \mathrm{Row} A = \dim \mathrm{Col} A = \mathrm{rk} A$

$\det : M_{n,n}(K) \to K$

$n = 1 \quad A = [a] \quad \det A = a$

$n > 1$ Ricorsivamente

A_{ik} ottenuta da A togliendo riga i e colonna k

$M_{ik} = \det A_{ik}$ (detto minore complementare)

$C_{ik} = (-1)^{i+k} M_{ik}$ (detto complemento algebrico)

$\det A = \sum_{i=1}^{n} a_{1i} C_{1i}$

I th. di Laplace: si può usare una riga o una colonna qualsiasi.

$\det A = \det A^T$

Se una riga è di zeri: $\det A = 0$

Se si scambiano 2 righe: $\det A' = -\det A$

Se due righe parallele sono proporzionali: $\det A = 0$

Moltiplicando una riga per t: $\det A' = t \det A$

$\det tA = t^n \det A$

In una matrice triangolare: $\det A = \prod_{i=1}^{n} a_{ii}$

T. di Binet: $\det AB = \det A \cdot \det B$

Regola di Kronecker: se esiste una sottomatrice quadrata $A'_{n,n}$ con $\det A' \neq 0$ allora $\mathrm{rk} A \geq n$. Se tutte le matrici ottenute per orlatura da A' hanno $\det = 0$ allora $\mathrm{rk} A = n$.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

$$\det A = ad - bc \qquad \det B = aei + bfg + cdh - ceg - afh - bdi$$

$$\begin{vmatrix} 2 & 0 & -2 & 1 \\ 3 & 1 & 3 & 4 \\ -1 & 2 & 0 & 2 \\ 1 & 0 & 1 & -1 \end{vmatrix} = 1(-)^4 \begin{vmatrix} 2 & -2 & 1 \\ -1 & 0 & 2 \\ 1 & 1 & -1 \end{vmatrix} + 2(-)^5 \begin{vmatrix} 2 & -2 & 1 \\ 3 & 3 & 4 \\ 1 & 1 & -1 \end{vmatrix} = (0-1-4-0+2-4)-2(-6-8+3-3-8-6) = 49$$

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases} \qquad A|\underline{b} = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right]$$

Se $b_1 = b_2 = \ldots = b_m = 0$ il sistema si dice omogeneo ed ammette sempre almeno una soluzione ($\underline{0}$). Le soluzioni non banali vengono chiamate *autosoluzioni*. Riducendo a scala $A|\underline{b}$ si ottiene un sistema equivalente.

Se $m = n$ e $\det A \neq 0$ per Cramer $x_i = \det A_i / \det A$ con A_i ottenuta da A sostituendo la i-esima colonna con $\underline{b}$.

Dato un sistema lineare e le matrici associate A e $A|\underline{b}$:

$\mathrm{rk}A \neq \mathrm{rk}A|\underline{b}$ Sistema impossibile

$\mathrm{rk}A = \mathrm{rk}A|\underline{b}$ Sistema possibile

In un sistema possibile in n incognite con $\mathrm{rk}A = r$:

$r = n$ Sis. det. 1! soluzione

$r < n$ Sis. ind. ∞^{n-r} soluzioni ($n - r$ variabili libere)

Un sistema lineare omogeneo ha autosoluzioni sse $\mathrm{rk}A < n$.

$$\begin{cases} 2x+4y+4z=4 \\ x-z=1 \\ -x+3y+4z=2 \end{cases} \qquad A|\underline{b} = \left[\begin{array}{ccc|c} 2 & 4 & 4 & 4 \\ 1 & 0 & -1 & 1 \\ -1 & 3 & 4 & 2 \end{array}\right]$$

Riducendo $A|\underline{b}$ a scala

$$\left[\begin{array}{ccc|c} 2 & 4 & 4 & 4 \\ 1 & 0 & -1 & 1 \\ -1 & 3 & 4 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 1 & 2 & 2 & 2 \\ -1 & 3 & 4 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 2 & 3 & 1 \\ 0 & 3 & 3 & 3 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 2 & 3 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right] \to$$

$$\begin{cases} x-z=1 \\ 2y+3z=1 \\ z=-1 \end{cases}$$