Linear algebra for videogames

MATHEMATICS IMPLEMENTATION IN C++

Vito Domenico Tagliente

Linear Algebra for videogames

Vito Domenico Tagliente

Prefazione

Questo libro presenta una spiegazione degli elementi fondamentali dell'algebra lineare tramite un approccio orientato all'applicazione e l'implementazione dei concetti discussi lungo i vari capitoli che lo compongono. I diversi argomenti saranno spiegati attraverso il formalismo matematico, arricchito di implementazione in codice.

Scopo di questo libro è quello di fornire una infarinatura sugli aspetti matematici su cui si fonda la computer grafica.

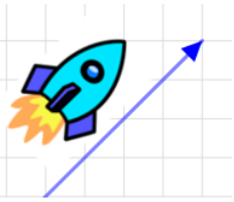
Note sull'autore

Il mio nome è Vito Domenico Tagliente. Mi sono laureato con lode alla facoltà di Ingegneria Informatica del politecnico di Bari, Italia. TODO.

Sommario

1. Vettori

In natura esistono grandezze determinate dal numero che le misura rispetto a una prefissata unità, come per esempio la lunghezza, l'area, il volume, il tempo. Queste grandezze sono dette scalari. Altre grandezze, come per esempio lo spostamento e la velocità, sono rappresentate da un numero, una direzione e un verso. Tali grandezze vengono chiamate grandezze vettoriali e vengono descritte mediante vettori.



Ad esempio, se vogliamo descrivere il movimento di una navicella spaziale, dire che questa si muove ad una certa velocità non è sufficiente. Affinché la descrizione sia completa, occorre specificare anche dove questa si sta muovendo, in particolare in quale direzione e con quale verso questa si sta spostando.

Nell'ambito dei videogiochi, i vettori sono ampiamente utilizzati per esprimere, nella considerazione più semplice possibile, gli spostamenti dei diversi oggetti che compongono la scena. Non solo, i vettori possono essere utilizzati per esprimere una direzione di osservazione applicata alle camere di gioco o semplicemente agli attori della scena che devono spostarsi e quindi ruotarsi e osservare lungo la direzione di spostamento. Gli scenari di applicazioni sono numerosi, l'importante è capire che i vettori devono essere utilizzati lì dove sono necessarie delle informazioni in più rispetto alla semplice grandezza numerica.

1.1 Notazione vettoriale

La notazione matriciale venne introdotta principalmente per esprimere le relazioni dell'algebra lineare in forma compatta, allo scopo di incrementarne la leggibilità. A prova di ciò, consideriamo un insieme di relazioni lineari definite tra un insieme $X = \{x_1, x_2, \dots, x_n\}$ di n elementi e un insieme $Y = \{y_1, y_2, \dots, y_m\}$ di m elementi.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2$$

$$\dots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = y_m$$

Il contenuto informativo della relazione può essere formalmente rappresentato dalla seguente notazione:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

Se consideriamo di associare un nome ai diversi elementi che appaiono nell'equazione $\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \boldsymbol{A}$, $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \boldsymbol{x}$ e $\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \boldsymbol{y}$, due oggetti matematici possono essere individuati:

- La matrice **A**
- I vettori **x** e **y**.

Applicando la nuova notazione, si ricava che la relazione può essere espressa in forma matriciale $\boldsymbol{Ax} = \boldsymbol{y}$.

Definiamo vettore, un insieme di n elementi che può essere espresso come:

$$\begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

Questo oggetto è chiamato vettore colonna. In seguito, sarà chiaro il perché un vettore possa essere considerato come un caso speciale di matrice. Una matrice avente n righe ed una sola colonna.

Un vettore può anche essere espresso secondo la notazione $\boldsymbol{x} = [x_1\ x_2 \dots x_n]$, tale rappresentazione prende il nome di vettore riga. In generale si parla semplicemente di vettori. Se non viene specificato alcun qualificatore, si sottintende si tratti, per convenzione, di un vettore colonna.

I vettori sono rappresentati nei testi di algebra con la notazione $\vec{v}$, in realtà per questioni di comodità relative alla videoscrittura, i vettori vengono anche spesso identificati con una lettera minuscola in grassetto **x**.

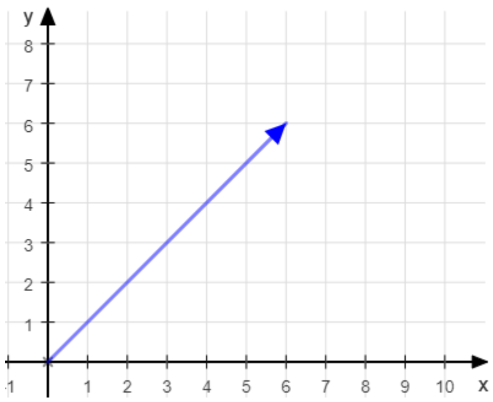
Gli elementi che costituiscono un vettore sono chiamati componenti, dove il simbolo n viene adoperato per indicare la quantità di componenti, l'<u>ordine</u> del vettore. Per esempio, definito il vettore $\boldsymbol{v} = [2 \quad 0]$, la prima componente è 2, la seconda è 0. L'ordine di **v** è 2, in quanto $n = 2$.

Molto importante è il concetto di modulo di un vettore, ovvero la lunghezza del vettore in relazione allo spazio euclideo in cui questo viene rappresentato. Formalmente, il modulo di un vettore, detto anche norma euclidea, si esprime con $||v|| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$. Nel caso in due dimensioni, si ricava $||v|| = \sqrt{x^2 + y^2}$.

1.2 Definizione geometrica di un vettore

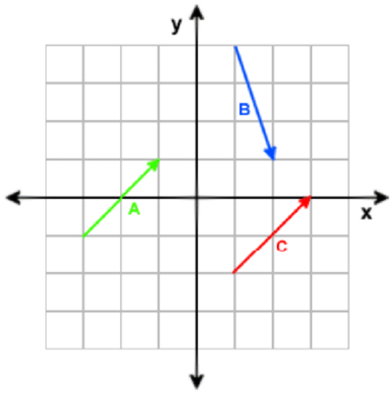
Abbiamo discusso riguardo alla rappresentazione analitica di un vettore, esaminiamo la rappresentazione geometrica. Geometricamente, un vettore può essere rappresentato con un <u>segmento orientato</u>. Un segmento AB può essere percorso in due modi: da A verso B, oppure da B verso A. Nel primo caso, il segmento orientato verrà indicato con la notazione $\overrightarrow{AB}$; nel secondo caso con $\overrightarrow{BA}$. Tale segmento è caratterizzato da una lunghezza, da una direzione e da un verso. Ciò significa che i segmenti $\overrightarrow{AB}$ e $\overrightarrow{BA}$, che sono lo stesso insieme di punti, devono essere considerati diversi come segmenti orientati, perché differisce l'ordine con cui si considerano gli estremi.

In figura è possibile notare la rappresentazione grafica di un vettore in uno spazio a due dimensioni, pertanto poggiato su un sistema cartesiano.



I vettori sono utilizzati per descrivere grandezze in base alla loro direzione e verso. La posizione non è specificata nella definizione del vettore stesso, piuttosto è relativa al punto in cui questo è applicato.

A livello di sistema cartesiano, ogni vettore è descritto dalle due coordinate del sistema, x e y. È da tenere a mente che tali coordinate sono necessarie al fine di identificare la pendenza della retta su cui giace il vettore, ma il punto di applicazione è irrilevante. Vettori aventi componenti identiche, possono essere rappresentati (paralleli tra loro) posti in posizioni differenti.



In figura, infatti, i vettori **a** e **c**, presentano stessa direzione, il che vuol dire che sono descritti dalle medesime componenti. La posizione è differente, questo perché tale informazione prescinde dalla notazione di vettore.

Pertanto, è bene distinguere il concetto di punto, inteso come concetto in grado di descrivere la posizione di un oggetto in un sistema di riferimento, rispetto al concetto di vettore, adatto a descrivere uno spostamento.

I vettori descrivono uno spostamento, sono quindi riferiti a posizioni relative. I punti, invece, descrivono la posizione assoluta di un oggetto. In realtà i due elementi, anche se concettualmente distinti, sono molto simili dal punto di vista matematico. Infatti, se consideriamo di applicare un vettore all'origine di un piano cartesiano. Dire che il vettore $\boldsymbol{v} = [x, y]$ si muove dall'origine alla posizione definita dal punto (x, y), non fa altro che mettere in relazione i due concetti espressi da un formalismo matematico simile.

1.3 Operazioni tra vettori

Di seguito vengono presentate le operazioni tra vettori di comune applicazione nella computer grafica.

1.3.1 Somma di vettori

La somma di due o più vettori si può ottenere, geometricamente, attraverso due regole equivalenti tra loro.

- Regola del parallelogramma: la somma di due vettori applicati ad uno stesso punto corrisponde alla diagonale del parallelogramma che essi definiscono con le rispettive proiezioni.
- Metodo punta coda: la somma di due o più vettori applicati in sequenza corrisponde al vettore che congiunge il punto di applicazione del primo all'estremità dell'ultimo.

Formalmente si ricava, definiti i due vettori $\boldsymbol{a} = [a_1\ a_2 \dots a_n]$ e $\boldsymbol{b} = [b_1\ b_2 \dots b_n]$, che la loro somma corrisponde

$$\boldsymbol{a} + \boldsymbol{b} = [a_1 + b_1 \quad a_2 + b_2 \quad \dots \quad a_n + b_n]$$

1.3.2 Prodotto scalare
Il prodotto scalare, in algebra lineare, si esprime quanto segue

$$\boldsymbol{a} \cdot \boldsymbol{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

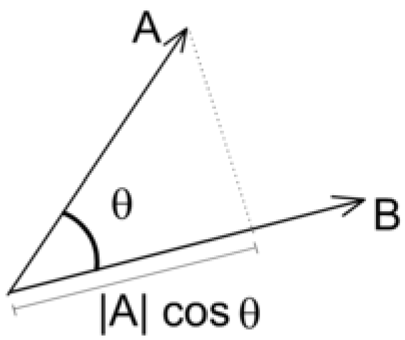
È possibile esplicitare l'operazione appena definita per i casi specifici nelle 2 e 3 dimensioni

- $\boldsymbol{a} \cdot \boldsymbol{b} = a_x b_x + a_y b_y$, in due dimensioni
- $\boldsymbol{a} \cdot \boldsymbol{b} = a_x b_x + a_y b_y + a_z b_z$, in tre dimensioni

Il prodotto scalare è commutativo, ciò vuol dire che vale la relazione $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot$

Esaminiamo l'interpretazione geometrica del prodotto scalare. Considerati due vettori ***a*** e ***b***, applicati nello stesso punto. Il prodotto scalare si definisce come

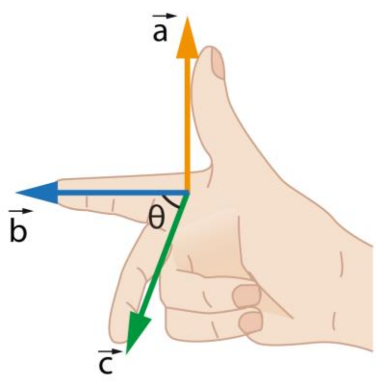
$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$$



A livello geometrico, dato che il prodotto cartesiano dipende dal coseno di theta (l'angolo comrpeso tra i due vettori), si ricava che se il prodotto scalare di due vettori è pari a zero, allora tali vettori sono ortogonali. Se theta è acuto, il prodotto scalare sarà una quantità positiva. Se theta è ottuso, il prodotto scalare risulterà negativo.

È da notare che il risultato di un prodotto scalare è una quantità, non un vettore.

1.3.3 Prodotto vettoriale
Per prodotto vettoriale si intende un'operazione tra due vettori che avviene nello spazio tridimensionale, il cui risultato è un vettore.



Dati due vettori ***a*** e ***b***, il loro prodotto vettoriale, indicato con $\boldsymbol{a} \times \boldsymbol{b}$, è un vettore che ha:

- Direzione perpendicolare al piano che contiene i due vettori ***a*** e ***b***
- Verso dato dalla regola della mano destra, secondo cui se si pone il pollice nel verso del vettore ***a*** e l'indice nel verso di ***b***, il vettore $\boldsymbol{a} \times \boldsymbol{b}$ è uscente dal palmo della mano.
- Modulo pari all'area del parallelogramma generato dai due vettori ***a*** e ***b***

Formalmente il prodotto vettoriale si descrive come

$$\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{n}|\boldsymbol{a}||\boldsymbol{b}| \sin\theta$$

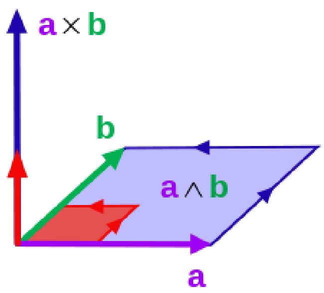
Dove $0 < \theta < \pi$ è l'angolo compreso fra ***a*** e ***b***, ***n*** è un versore normale al piano formato dai due vettori, che fornisce la direzione del prodotto vettoriale. Si nota che $|\boldsymbol{a}||\boldsymbol{b}| \sin\theta$ rappresenta l'area del parallelogramma individuati dai due vettori.

Esplicitamente, definiti ***i***,***j*** e ***k*** i versori di una base ortonormale di R^3, il prodotto di $\boldsymbol{a} = (a_x, a_y, a_z)$ e $\boldsymbol{b} = (b_x, b_y, b_z)$ può essere scritto come il determinante di una matrice (le matrici verranno esaminate nel dettaglio nel capitolo 2):

$$\boldsymbol{a} \times \boldsymbol{b} = det \begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{bmatrix} = (a_y b_z - a_z b_y)\boldsymbol{i} - (a_x b_z - a_z b_x)\boldsymbol{j} + (a_x b_y - a_y b_x)\boldsymbol{k} = \begin{bmatrix} a_y b_z - a_z b_y \\ a_x b_z - a_z b_x \\ a_x b_y - a_y b_x \end{bmatrix}$$

A livello implementativo, poiché in ambito delle applicazioni multimediali quali i videogiochi le prestazioni sono una caratteristica chiave da tenere sempre ben in considerazione, è opportuno esplicitare la formula risultante, permettendo al calcolatore di trascurare calcoli e operazioni non necessarie.

Dal punto di vista geometrico, ricaviamo che il prodotto vettoriale permette di conoscere la natura dei vettori partecipanti nel prodotto. Infatti, se i due sono paralleli tra loro, il prodotto vettoriale sarà nullo, in quanto dipendente dal $\sin \theta$.



Un'ultima considerazione. Il prodotto vettoriale è definito solo nello spazio tridimensionale, trattasi di un caso particolare di semplificazione del prodotto esterno che vale per vettori in spazi ad n dimensioni.

Il prodotto esterno di due vettori è un bivettore, cioè un elemento di piano orientato. Dati due vettori ***a*** e **b**, il bivettore ***a***∧***b*** è il parallelogramma orientato formato dai due vettori.

Con questo piccolo accenno al prodotto esterno, sorvoliamo l'argomento in quanto non è indispensabile per la trattazione matematica di un contesto di applicazione ai videogiochi, strettamente legato al mondo tridimensionale. Per approfondire l'argomento, si consiglia di consultare un testo di algebra lineare adatto.

2. Matrici

Al fine di presentare l'argomento sui vettori, sono già stati introdotti alcune notazioni ed informazioni sulle matrici, di seguito mi occuperò di completarne la spiegazione.

In algebra lineare, una matrice è una tabella ordinata di elementi. Sono largamente utilizzate per la loro capacità di rappresentare in maniera utile e coincisa diversi oggetti matematici, basti considerare che le matrici possono essere adoperati per rappresentare sistemi lineari, come mostrato nel paragrafo [1.2].

Una matrice è caratterizzata da un numero di righe n e di colonne m. Quando $n = m$, la matrice si dice quadrata. Di seguito è presentata una matrice 2x3.

$$\boldsymbol{M} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

Come per i vettori, anche le matrici vengono rappresentate con lettere in grassetto, prediligendo lettere in maiuscolo.

Gli elementi che compongono la matrice, vengono identificati utilizzando la notazione *ij*, dove i identifica l'indice di riga e j, rispettivamente, di colonna. In tal caso, una matrice di dimensione $n \times m$, si può descrivere formalmente come:

$$A = \begin{bmatrix} a_{0,0} & a_{0,1} & \cdots & a_{0,m} \\ a_{1,0} & a_{1,1} & \cdots & a_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,0} & a_{n,1} & \dots & a_{n,m} \end{bmatrix}$$

Nel caso in cui si abbia una matrice quadrata, gli elementi di indice $i = j$, vengono detti elementi della diagonale principale della matrice. Una matrice avente elementi pari a zero, eccetto per quelli costituenti la diagonale principale, viene detta <u>matrice diagonale</u>.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Sulla diagonale principale, in tal caso, possono presentarsi anche uno o più zeri, l'importante è che resti almeno un elemento diverso da zero.

3. Punti