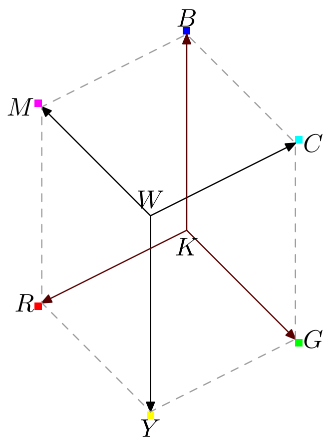
Spazio affine e colori

In questo foglio vogliamo dare qualche rapido cenno di alcune codifiche dei colori e di come queste si descrivano facilmente tramite la nozione di spazio affine. Iniziamo dalla cosiddetta codifica *rgb*, per cui i colori corrispondono ai punti del cubo unitario

$$[0,1]^3 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{A}(\mathbb{R}^3) \mid x,y,z \in [0,1] \right\}.$$



Nell'origine è posto il nero, $K = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, mentre i tre colori corrispondenti alle direzioni fondamentali sono il *rosso* $R = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, il *verde* $G = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, e il *blu* $B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Ogni colore è quindi identificato da tre coordinate $\begin{pmatrix} r \\ g \\ b \end{pmatrix}$, nel sistema di riferimento che ha l'origine in K e come direzioni fondamentali i vettori $R-K$, $G-K$, $B-K$; con la limitazione che le coordinate appartengono tutte all'intervallo unitario $[0,1] \subset \mathbb{R}$. Dunque il colore X, che si ottiene *sovrapponendo r parti di rosso, g parti di verde e b parti di blu* corrisponde al punto di coordinate $\begin{pmatrix} r \\ g \\ b \end{pmatrix}$ nel cubo, ovvero $X = K + r(R-K) + g(G-K) + b(B-K)$. Ad esempio, possiamo considerare i colori *ciano* $C = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, ovvero $C = K + (G-K) + (B-K)$, *magenta* $M = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, *giallo* $Y = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, e il *bianco* $W = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Si veda ad esempio il disegno a fianco, dove sono stati posti dei piccoli quadratini colorati vicino ai punti corrispondenti.

La scelta, definita da un comitato internazionale nel 1931, sta alla base della rappresentazione dei colori su uno schermo di computer, dove a ogni pixel si associa un colore secondo la codifica *rgb*, e i valori delle tre coordinate non variano arbitrariamente in $[0,1]$, ma usualmente sono numeri binari da 8 cifre, ovvero interi compresi tra 0 e $255 = 2^8 - 1$ e devono perciò essere interpretati come il numeratore di una frazione con denominatore 255. Si parla in questo caso di colori a 24bit, e si può ottenere in tal modo una gamma di oltre 16 milioni di colori (precisamente $2^{24} = 16777216$), ciascuno rappresentato da 24 cifre binarie, ovvero da un numero di 6 cifre in base 16 [(ad esempio, il bianco, W, è `#FFFFFF`, ovvero $(255, 255, 255)$]. Quindi per ogni punto (pixel) dello schermo ci sono tre sorgenti dei colori fondamentali e, a seconda delle diverse intensità luminose delle tre sorgenti, si ottengono i vari colori: quando sono tutte e tre spente il pixel è nero, mentre diventa bianco quando sono tutte alla massima intensità luminosa.

Se però dobbiamo passare dallo schermo a una stampante la situazione cambia, perché siamo in grado di ottenere i colori su un foglio di carta, partendo dal foglio bianco e aggiungendo pigmenti colorati, che, mi è sembrato di capire, agiscono impedendo selettivamente per alcune lunghezze d'onda, la riflessione completa della luce che si ha col colore bianco. Nelle stampanti sono presenti, di solito, tre pigmenti fondamentali, ciano, magenta e giallo (proprio i colori indicati sopra). Dobbiamo quindi cambiare il riferimento affine nel cubo, prendendo come origine W e come direzioni fondamentali $C - W = -(R-K)$, $M - W = -(G-K)$, $Y - W = -(B-K)$; le coordinate in questo riferimento si indicano di solito con le lettere *cmy*. I due riferimenti sono stati rappresentati nel disegno sopra. Il cambio di coordinate si esprime quindi osservando che

$$\begin{aligned} X &= K + r(R-K) + g(G-K) + b(B-K) \\ &= W + (K-W) - r(C-W) - g(M-W) - b(Y-W) \\ &= W + (1-r)(C-W) + (1-g)(M-W) + (1-b)(Y-W) \\ &= W + c(C-W) + m(M-W) + y(Y-W), \end{aligned}$$

essendo $K - W = -(W-K) = -[(R-K) + (G-K) + (B-K)] = (C-W) + (M-W) + (Y-W)$. È immediato dedurre da questo il cambiamento inverso.

Bisogna però osservare che ben poche stampanti utilizzano questo modo di rappresentare i colori; molto più di frequente viene utilizzato il cosiddetto sistema *cmyk* che usa 4 coordinate per rappresentare i colori, aggiungendo il colore nero ai tre già indicati(*). Per ottenere queste nuove coordinate per il colore X, basta partire dalla scrittura

$$\begin{aligned} X &= W + (K - W) - r(C - W) - g(M - W) - b(Y - W) \\ &= W + (1 - h)(K - W) + (h - r)(C - W) + (h - g)(M - W) + (h - b)(Y - W) \end{aligned}$$

ove si è sommato il vettore nullo, ovvero $0 = -h(K - W) + h(C - W) + h(M - W) + h(Y - W)$, per un opportuno h affiché i coefficienti $1 - h$, $h - r$, $h - g$ e $h - b$ siano tutti maggiori o uguali a zero e uno almeno tra gli ultimi tre sia nullo. Per ottenere questo, basta (ed è necessario) prendere $h = \max(r, g, b)$.

Sembrerebbe ragionevole prendere i coefficienti $h - r$, $h - g$, $h - b$ e $1 - h$ come valori delle nuove coordinate, ma non è questa la consuetudine: si considerano solo i colori $X \neq K$ e si pone

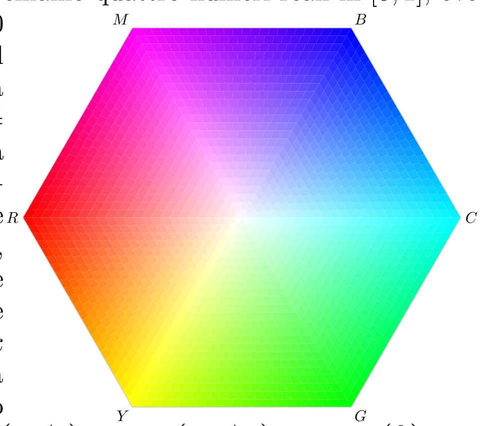
$$c = \frac{h - r}{h}, \qquad m = \frac{h - g}{h}, \qquad y = \frac{h - b}{h}, \qquad k = 1 - h$$

(si confronti ad esempio con `http://www.rapidtables.com/convert/color/rgb-to-cmyk.htm`). In questo modo ai punti del cubo dei colori si associano le quattro coordinate nel sistema *cmyk*, aggiungendo la quaterna $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ per il punto K, ovvero per il colore nero. La corrispondenza è biunivoca e le formule inverse sono

$$r = (1 - k)(1 - c), \qquad g = (1 - k)(1 - m), \qquad b = (1 - k)(1 - y);$$

ove $k \in [0, 1)$ e la terna $\begin{pmatrix} c \\ m \\ y \end{pmatrix}$ ha le coordinate in $[0, 1]$ e una delle tre è nulla. La corrispondenza si estende al colore nero.

Abbiamo quindi due modi di attribuire coordinate ai colori del cubo unitario. Col sistema *rgb* abbiamo una terna di numeri reali compresi tra 0 e 1; col sistema *cmyk* otteniamo quattro numeri reali in $[0, 1]$, ove almeno uno tra i primi tre è sempre nullo ($c = 0$ oppure $m = 0$ oppure $y = 0$) e la coordinata k è in $[0, 1)$ e vale 1 solo per il nero. Possiamo descrivere in modo un po' più geometrico questa corrispondenza: un punto $X = K + r(R - K) + g(G - K) + b(B - K) \neq K$ del cubo dei colori, definisce una retta $K \vee X$ che interseca la superficie del cubo, oltre che nel punto K, nel punto $X_0 = K + \frac{1}{h}(X - K)$, ove $0 < h = \max(r, g, b) \leq 1$. Il punto X_0 appartiene quindi a una delle facce $r = 1$ o $g = 1$ o $b = 1$ (ovvero $c = 0$, $m = 0$ o $y = 0$) e si può scrivere $X = hX_0 + (1 - h)K$ (coordinate baricentriche). Le coordinate *cmy* del punto X_0 e il coefficiente $k = 1 - h$ di K nella scrittura precedente danno le coordinate *cmyk* di X. Possiamo dare una rappresentazione tridimensionale (diversa dal cubo) dei punti di coordinate *cmyk*. Consideriamo lo spazio affine $\mathbb{A}(\mathbb{R}^3)$ con le coordinate $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ e i vettori $v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} -1/2 \\ \sqrt{3}/2 \\ 0 \end{pmatrix}$, $v_3 = \begin{pmatrix} -1/2 \\ -\sqrt{3}/2 \\ 0 \end{pmatrix}$ e $v_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Al colore X di coordinate $\begin{pmatrix} c \\ m \\ y \\ k \end{pmatrix}$ (con $c = 0$ oppure $m = 0$ oppure $y = 0$), possiamo associare il punto



$$X' = (1 - k)(O + cv_1 + mv_2 + yv_3) + k(O + v_4).$$

(*) Come abbiamo già osservato, il vettore $K - W$ dipende linearmente da $C - W$, $M - W$ e $Y - W$, quindi non si ha un sistema di coordinate associato a una base dello spazio vettoriale. È necessario fare alcune restrizioni sulle quaterne che si ottengono per poter stabilire una corrispondenza biunivoca tra i punti del cubo e queste nuove coordinate. Uno dei motivi per cui si giustifica questa scelta è che ogni colore viene rappresentato utilizzando al più due pigmenti e un'eventuale quantità di nero e il costo di quest'ultimo è inferiore rispetto a quello dei pigmenti colorati.

I punti con $k = 0$, ovvero quelli che corrispondono a colori che giacciono sulla faccia esterna del cubo, formano l'esagono del piano $x_3 = 0$ rappresentato nella figura qui sopra (in cui abbiamo ridotto la densità dei colori che sarebbero circa 200000). I vertici dell'esagono corrispondono agli omonimi colori e nel centro dell'esagono (l'origine) c'è il colore bianco. A partire da ogni punto X_0 dell'esagono, spostandosi lungo il segmento che lo congiunge al punto $K = O + v_4$ (che corrisponde al colore nero), si possono trovare le "sfumature" più scure della stessa tinta, ovvero il punto $(1 - t)X_0 + tK$ corrisponde al colore X_0 con l'aggiunta di una quantità $t \in [0, 1]$ di nero. Si ottengono così tutti i punti di una una piramide con base esagonale e vertice in $K^{(\dagger)}$.

Il segmento che congiunge tale punto con l'origine (l'altezza della piramide) corrisponde alla diagonale del cubo dei colori ovvero al segmento che congiunge il bianco con il nero: la cosiddetta *scala dei grigi*. La proiezione dei colori del cubo sul segmento WK, permette di trasformare un'immagine in una corrispondente immagine in scala di grigi (che ha bisogno quindi di solo 8bit per pixel, in confronto con i 24 necessari per l'immagine a colori). La scelta più naturale per la proiezione del colore X consiste nel fare la proiezione ortogonale su $\langle W - K\rangle$ del vettore $X - K$, nelle coordinate rgb. In tal caso il valore di grigio corrispondente al punto X, di coordinate $\begin{pmatrix} r \\ g \\ b \end{pmatrix}$, è

$$\mu = \frac{(X - K) \cdot (W - K)}{(W - K) \cdot (W - K)} = \frac{r + g + b}{3},$$

ovvero la media dei valori dei tre colori. Per motivi estetici, non si usa quasi mai questa proiezione, ma si preferiscono delle scelte empiriche differenti. In diverse pagine online, si può trovare l'indicazione della forma lineare $\zeta = (0.2126, 0.7152, 0.0722)$ per calcolare la quantità di grigio o il corrispondente valore $\lambda = \zeta \circ (X - K) = 0.2126r + 0.7152g + 0.0722b$; ma l'indicazione non è univoca (ad esempio, in altre pagine si può trovare $\xi = (0.2979, 0.5879, 0.1142)$ o altri valori, nella maggior parte dei casi, giustificati da ragioni estetiche).

$(\dagger)$ Se si fosse usata la corrispondenza $\begin{pmatrix} c \\ m \\ y \\ k \end{pmatrix} \mapsto O + cv_1 + mv_2 + yv_3 + kv_4$, si avrebbe avuto, in luogo della piramide, un prisma privato della faccia superiore dove vi sarebbe stato il solo punto corrispondente al nero. Anche questo compare in alcune descrizioni.