Numero di condizionamento rispetto alla inversione

Siano $\mathcal{U}$ e $\mathcal{V}$ spazi vettoriali di dimensione n e dotati di norma. Assegnata una trasformazione lineare

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

si consideri l'insieme $\mathcal{P}$ dei vettori di $\mathcal{U}$ aventi norma uguale ad 1 e l'insieme $\mathcal{Q}$ dei vettori di $\mathcal{V}$ ottenuti applicando A ai vettori di $\mathcal{P}$.

Sia v_m uno dei vettori di $\mathcal{Q}$ aventi norma minima e v_M uno dei vettori aventi norma massima.

Se $\|v_m\| = 0$ allora la trasformazione A non è iniettiva. Se invece A è un isomorfismo, siano u_m e u_M i vettori in $\mathcal{U}$ corrispondenti ai vettori v_m e v_M.

Supponendo di aver scelto v_m e v_M indipendenti (questo è sempre possibile per $n > 1$), anche i vettori u_m e u_M risultano indipendenti. Si consideri il sottospazio U generato dai vettori u_m e u_M e il vettore di tale sottospazio

$$u := u_m + \epsilon u_M$$

Riguardando tale vettore come ottenuto dal vettore u_m aggiungendo una “perturbazione”, è interessante esaminare come tale perturbazione si riflette sul vettore v_m. Si osservi che

$$\frac{\|u - u_m\|}{\|u_m\|} = |\epsilon| \frac{\|u_M\|}{\|u_m\|} = |\epsilon|$$

$$v := Au = v_m + \epsilon v_M \quad \Rightarrow \quad \frac{\|v - v_m\|}{\|v_m\|} = |\epsilon| \frac{\|v_M\|}{\|v_m\|}$$

La perturbazione è dunque in generale amplificata, essendo $\|v_M\|/\|v_m\| \geq 1$. Una perturbazione del vettore u_M

$$u := u_M + \epsilon u_m$$

è invece in generale smorzata, essendo

$$v := Au = v_M + \epsilon v_m \quad \Rightarrow \quad \frac{\|v - v_M\|}{\|v_M\|} = |\epsilon| \frac{\|v_m\|}{\|v_M\|}$$

Facendo riferimento al primo caso, visto come caso sfavorevole, al rapporto

$$\kappa(A) := \frac{\|v_M\|}{\|v_m\|}$$

si può attribuire il significato di misura della sensibilità dell'isomorfismo A alle perturbazioni dei vettori a cui è applicato.

Utilizzando la definizione di *norma di una trasformazione lineare* si osservi che

$$\|A\| := \max \frac{\|Au\|}{\|u\|} = \|v_M\|$$

$$\|A^{-1}\| := \max \frac{\|A^{-1}v\|}{\|v\|} = \max \frac{\|u\|}{\|Au\|} = 1/\min \frac{\|Au\|}{\|u\|} = 1/\|v_m\|$$

$$\kappa(A) := \|v_M\|/\|v_m\| = \|A\|\|A^{-1}\|$$

Si noti che $\kappa(A) = \kappa(A^{-1})$. Poiché il problema più interessante è quello di valutare la immagine inversa di un vettore di $\mathcal{V}$ e non quello di valutare l'immagine di un vettore di $\mathcal{U}$, a $\kappa(A)$ si dà il nome di *numero di condizionamento di A rispetto alla inversione*.

Si dice anche che una trasformazione A è *ben condizionata* o *mal condizionata* a seconda che $\kappa(A)$ sia vicino ad 1 (il suo valore minimo) o molto più grande di 1.

Tutto quanto definito si può riferire in particolare ad una matrice riguardando questa come la matrice di una trasformazione lineare.

Una illuminante interpretazione di κ si ottiene utilizzando la norma indotta da un prodotto interno. Indicando con

$$A = RU$$

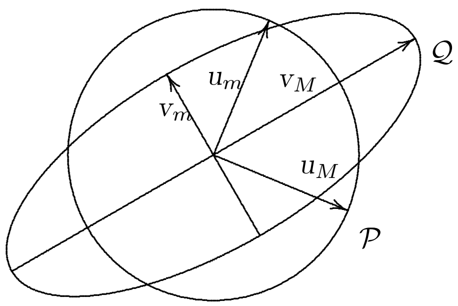
la decomposizione polare di A, si osservi che

$$\max \frac{\|Au\|}{\|u\|} = \max \frac{\|RUu\|}{\|u\|} = \max \frac{\|Uu\|}{\|u\|} = \max \sqrt{\mu_i}$$
$$\min \frac{\|Au\|}{\|u\|} = \min \frac{\|RUu\|}{\|u\|} = \min \frac{\|Uu\|}{\|u\|} = \min \sqrt{\mu_i}$$

avendo indicato con μ_i gli autovalori di A^*A.
Risulta pertanto

$$\kappa(A) = \frac{\max \sqrt{\mu_i}}{\min \sqrt{\mu_i}}$$

Una efficace rappresentazione grafica si può dare nel caso in cui $\mathcal{U} = \mathcal{V} = \mathbb{R}^2$, interpretando i vettori come *vettori posizione* di punti di un piano. All'insieme $\mathcal{P}$ corrisponde un cerchio di raggio unitario. Applicando U ai vettori di $\mathcal{P}$ si ottiene un'ellisse con assi corrispondenti ai due autospazi (ortogonali) di U. Applicando R si ottiene poi la figura corrispondente all'insieme $\mathcal{Q}$. Le lunghezze dei semiassi dell'ellisse sono gli autovalori $\sqrt{\mu}_1$, $\sqrt{\mu}_2$ corrispondenti ai due autospazi di U. Pertanto $\kappa(A)$ può essere interpretato come il rapporto tra le lunghezze dei semiassi. Tale valore è tanto più grande quanto più l'ellisse risulta *schiacciata*.



Questa interpretazione si può estendere al caso in cui $\mathcal{U} = \mathcal{V} = \mathbb{R}^n$, limitando l'attenzione al sottospazio generato da una coppia di vettori, uno appartenente all'autospazio di U corrispondente a $\max \sqrt{\mu}_i$ e l'altro appartenente all'autospazio di U corrispondente a $\min \sqrt{\mu}_i$.