AYUDANTÍA S10

MÍNIMOS CUADRADOS

Ayudante: Francisco Manríquez Novoa

Jueves 15 de mayo del 2025

Computación Científica

Combinación lineal [1]

Dados dos vectores $\mathbf{v}_1 = (1, 1, 0)$ y $\mathbf{v}_2 = (2, 0, -1)$, se puede pensar en todas sus posibles **combinaciones lineales**: vectores de la forma

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$$

Combinación lineal [2]

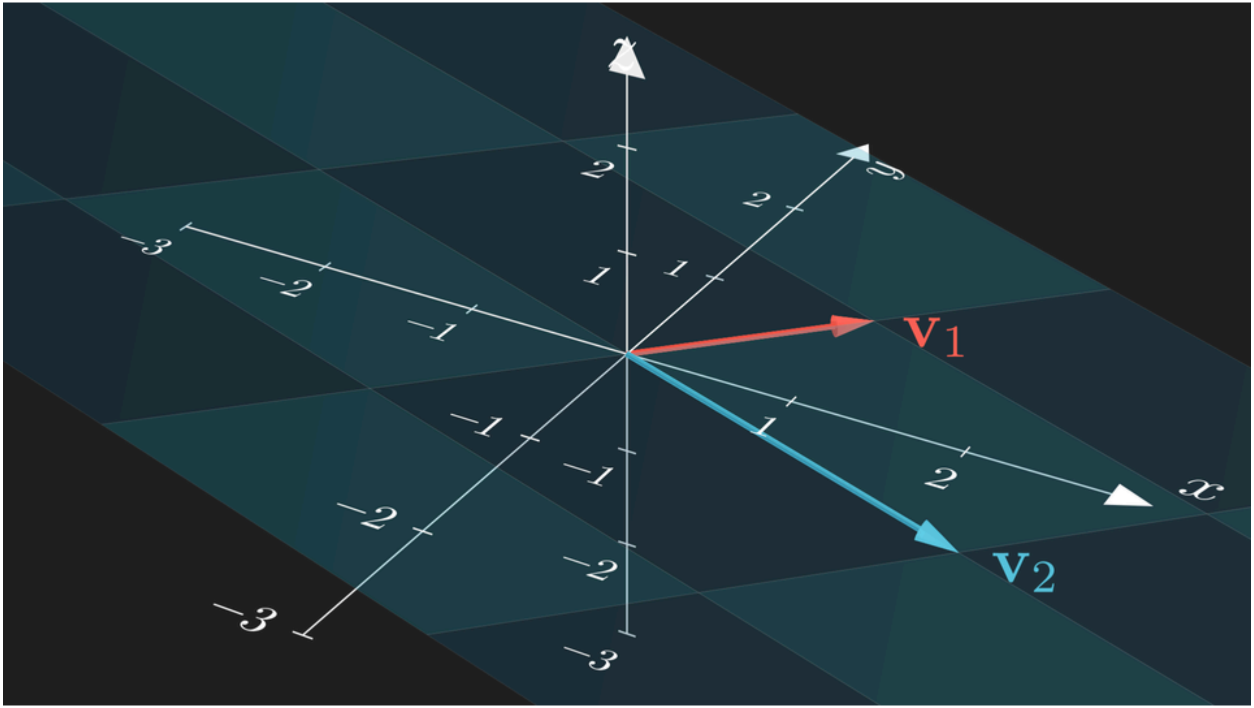
Dados dos vectores $v_1 = (1, 1, 0)$ y $v_2 = (2, 0, -1)$, se puede pensar en todas sus posibles **combinaciones lineales**: vectores de la forma

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$$

Todas estas combinaciones forman un PLANO en R^3.

Jerga de MAT022: $\{v_1, v_2\}$ GENERA el plano. El ESPACIO GENERADO de $\{v_1, v_2\}$ es el plano.

Combinación lineal [3]



Combinación lineal [4]

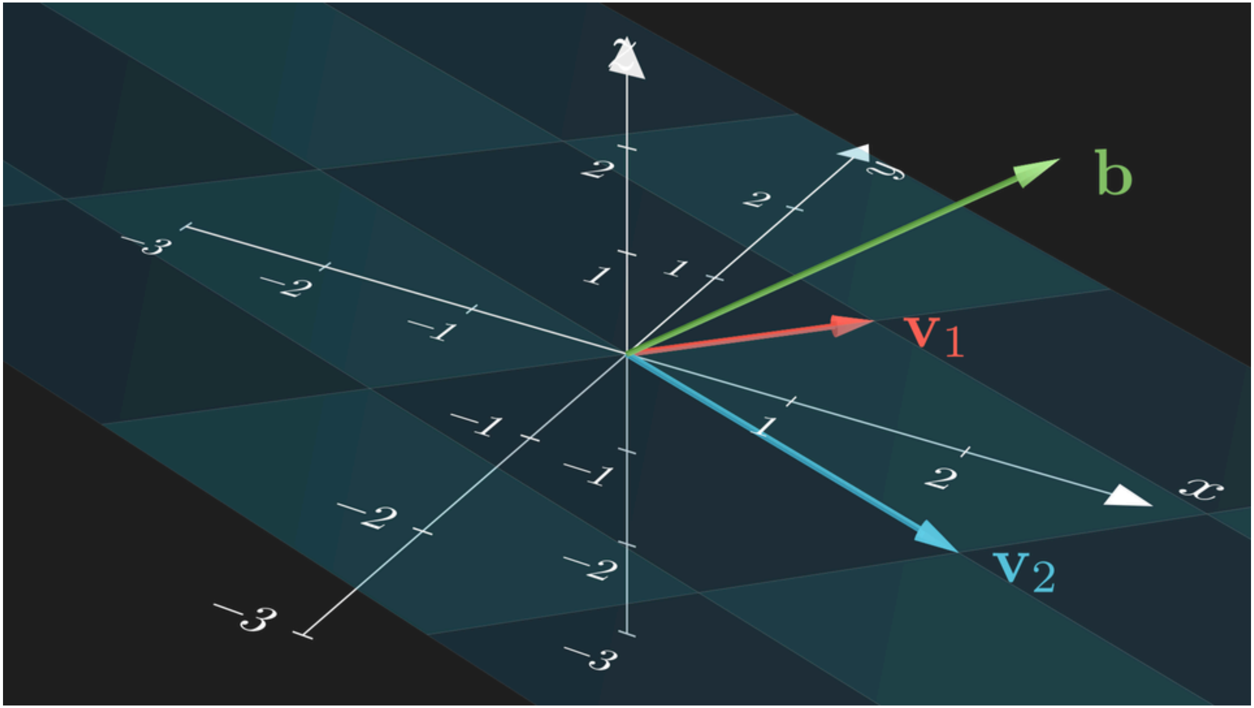
Dados dos vectores $\mathbf{v}_1 = (1, 1, 0)$ y $\mathbf{v}_2 = (2, 0, -1)$, se puede pensar en todas sus posibles **combinaciones lineales**: vectores de la forma

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$$

Todos los vectores en el plano se pueden expresar como combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$.

Sin embargo, ¿qué pasa con un vector que NO está en el plano? $\mathbf{b} = (2, 1, 2)$

Combinación lineal [5]



La mejor aproximación

Es imposible expresar **b** como combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$.

Sin embargo, es posible APROXIMAR **b** con la combinación lineal que más se acerque.

La mejor aproximación

Es imposible expresar **b** como combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$.

Sin embargo, es posible APROXIMAR **b** con la combinación lineal que más se acerque.

Queremos MINIMIZAR la distancia entre **b** y nuestra aproximación $\mathbf{b}_{\text{aprox}} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$.

$$\min \|\mathbf{b} - \mathbf{b}_{\mathrm{aprox}}\|_2$$

norma-2

Conceptos [1]

La combinación lineal $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$ se puede condensar como el producto matriz-vector $V\boldsymbol{\alpha}$:

$$\alpha_1 \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{v}_1} + \alpha_2 \underbrace{\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}}_{\mathbf{v}_2} = \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 0 & -1 \end{pmatrix}}_{V} \underbrace{\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}}_{\boldsymbol{\alpha}}$$

V es la matriz cuyas columnas son $\mathbf{v}_1$ y $\mathbf{v}_2$.
$\boldsymbol{\alpha}$ es el vector de coeficientes (α_1, α_2).

Conceptos [2]

Entonces, nuestra aproximación de **b**, es decir, $\mathbf{b}_{\text{aprox}} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2$, se puede expresar como $V\boldsymbol{\alpha}$.

Debemos minimizar la distancia entre **b** y $V\boldsymbol{\alpha}$:

$$\min \|\mathbf{b} - V\boldsymbol{\alpha}\|_2$$

Nota: en el apunte, se usa $A\mathbf{x}$ en vez de $V\boldsymbol{\alpha}$. Se dice que $A\mathbf{x} \approx \mathbf{b}$. En estas slides, para evitar ambigüedad, diremos que $V\boldsymbol{\alpha} \approx \mathbf{b}$, porque la matriz V viene de los vectores $\mathbf{v}_i$ y el vector $\boldsymbol{\alpha}$ viene de los coeficientes α_i.

Conceptos [3]

A la diferencia **b** - **b**$_{aprox}$ la llamamos "RESIDUO" y la denotamos como **r**.

$$\mathbf{r} = \mathbf{b} - V\boldsymbol{\alpha}$$

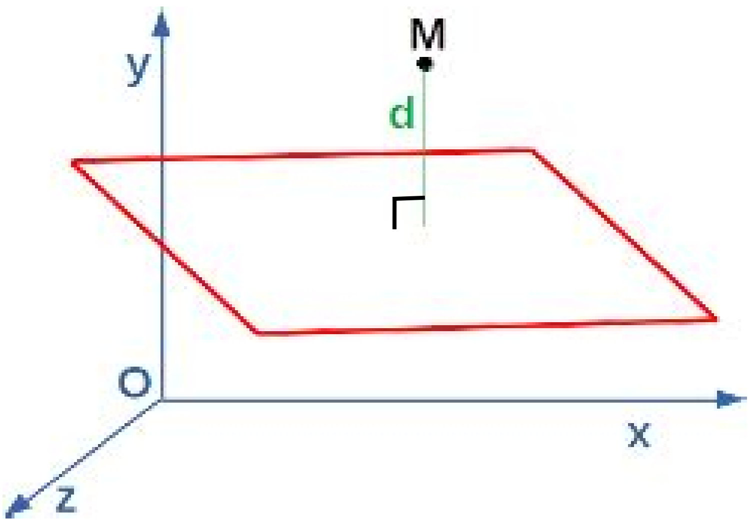
o, en el apunte,

$$\mathbf{r} = \mathbf{b} - A\mathbf{x}$$

Sea como sea, buscamos MINIMIZAR su norma.

Residuo mínimo [1]

La distancia mínima entre un punto y plano es perpendicular a este plano.



Residuo mínimo [2]

Entonces, el vector residuo mínimo $\mathbf{r}_{min}$ es el perpendicular al plano generado por $\mathbf{v}_1$ y $\mathbf{v}_2$.

Residuo mínimo [3]

Para que el residuo $\mathbf{r}_{\text{min}}$ sea perpendicular al plano generado por $\mathbf{v}_1$ y $\mathbf{v}_2$, basta con que sea perpendicular a $\mathbf{v}_1$ y perpendicular a $\mathbf{v}_2$. Dos vectores son perpendiculares si su producto punto es 0:

$$
\begin{aligned}
\mathbf{v}_1^T \mathbf{r}_{\text{min}} &= 0 \\
\mathbf{v}_2^T \mathbf{r}_{\text{min}} &= 0
\end{aligned}
$$

Esto basta para que $\mathbf{r}$ sea automáticamente perpendicular a todos los $\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2$ del plano.

Residuo mínimo [4]

Las ecuaciones anteriores se reescriben como:

$$\begin{aligned} \mathbf{v}_1^T \mathbf{r}_{\min} &= 0 \\ \mathbf{v}_2^T \mathbf{r}_{\min} &= 0 \end{aligned}$$

$$\begin{aligned} \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \mathbf{r}_{\min} &= 0 \\ \begin{pmatrix} 2 & 0 & -1 \end{pmatrix} \mathbf{r}_{\min} &= 0 \end{aligned}$$

$$\underbrace{\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}}_{V^T} \mathbf{r}_{\min} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$V^T \mathbf{r}_{\min} = \mathbf{0}$$

Residuo mínimo [5]

Finalmente, si $r_{min} = b - V\bar{\alpha}$:

$$V^T \mathbf{r}_{\min} = \mathbf{0}$$

$$V^T(\mathbf{b} - V\bar{\boldsymbol{\alpha}}) = \mathbf{0}$$

$$V^T\mathbf{b} - V^T V\bar{\boldsymbol{\alpha}} = \mathbf{0}$$

$$V^T\mathbf{b} = V^T V\bar{\boldsymbol{\alpha}}$$

$$V^T V\bar{\boldsymbol{\alpha}} = V^T\mathbf{b}$$

Residuo mínimo [6]

$$V^T V \bar{\mathbf{\alpha}} = V^T \mathbf{b}$$

Estas son las ecuaciones normales (en el apunte, $A^T A \bar{\mathbf{x}} = A^T \mathbf{b}$). Nótese la barra sobre $\bar{\mathbf{\alpha}}$ y $\bar{\mathbf{x}}$.

Resolverlas nos dará los coeficientes $\bar{\mathbf{\alpha}} = (\bar{\alpha}_1, \bar{\alpha}_2)$ que minimizan el residuo.

Con ellos, obtenemos también la mejor aproximación a $\mathbf{b}$: $V\bar{\mathbf{\alpha}} = \bar{\alpha}_1 \mathbf{v}_1 + \bar{\alpha}_2 \mathbf{v}_2$.

Residuo mínimo [7]

Resolvamos. Recuerda que $\mathbf{v}_1 = (1, 1, 0)$, $\mathbf{v}_2 = (2, 0, -1)$ y $\mathbf{b} = (2, 1, 2)$:

$$V^T V \bar{\boldsymbol{\alpha}} = V^T \mathbf{b}$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \bar{\alpha}_1 \\ \bar{\alpha}_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} \bar{\alpha}_1 \\ \bar{\alpha}_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

Resultado: $(\bar{\alpha}_1, \bar{\alpha}_2) = (2.5, -1)$.

Residuo mínimo [8]

Con esto, la mejor aproximación a **b** = (2, 1, 2) es

$$\underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 0 & -1 \end{pmatrix}}_{V} \underbrace{\begin{pmatrix} 2.5 \\ -1 \end{pmatrix}}_{\bar{\alpha}} = 2.5 \underbrace{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{v}_1} -1 \underbrace{\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}}_{\mathbf{v}_2} = \begin{pmatrix} 0.5 \\ 2.5 \\ 1 \end{pmatrix}$$

y la distancia entre **b** y la aproximación es

$$\left\| \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 0.5 \\ 2.5 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 1.5 \\ -1.5 \\ 1 \end{pmatrix} \right\| = \sqrt{1.5^2 + (-1.5)^2 + 1^2} = \sqrt{5.5} \approx 2.345$$

Residuo mínimo [9]

¡Resolvimos un problema de mínimos cuadrados!

Resumen: Dados vectores $\mathbf{v}_1$, $\mathbf{v}_2$, ..., ¿cuál es la mejor aproximación a un vector $\mathbf{b}$ con la combinación lineal $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots$?

1. Define $\mathbf{b}_{\text{aprox}} = V\boldsymbol{\alpha} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots$
2. Define el residuo $\mathbf{r} = \mathbf{b} - \mathbf{b}_{\text{aprox}} = \mathbf{b} - V\boldsymbol{\alpha}$.
3. El residuo debe ser perpendicular a $\mathbf{v}_1$, $\mathbf{v}_2$, ... Es decir, $\mathbf{v}_1{}^T\mathbf{r} = 0$, $\mathbf{v}_2{}^T\mathbf{r} = 0$, ... => $V^T\mathbf{r} = \mathbf{0}$.
4. Surgen las ecuaciones normales $V^TV\bar{\boldsymbol{\alpha}} = V^T\mathbf{b}$. Resuelve esas ecuaciones para encontrar $\bar{\alpha}_i$.

Un problema

El problema con las ecuaciones $V^T V \boldsymbol{\alpha} = V^T \mathbf{b}$ (o, en general, $A^T A \mathbf{x} = A^T \mathbf{b}$) es que la matriz $V^T V$ suele ser mal condicionada.

Por eso, en vez de resolver las ecuaciones normales directamente, intentamos resolver alternativas que nos entreguen la misma solución, pero con una matriz mejor condicionada.

Vectores ortonormales

Si los vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ fueran ortonormales, es decir, ortogonales entre sí ($\mathbf{v}_1{}^T\mathbf{v}_2 = 0$) y de norma 1 ($|\mathbf{v}_i|^2 = \mathbf{v}_i{}^T\mathbf{v}_i = 1$), **¡entonces la matriz $V^TV = I$!**

¿Por qué? Si recuerdas que el producto matricial *AB* involucra hacer productos punto entre las filas de *A* y las columnas de *B*:

$$V^TV = \begin{pmatrix} — & \mathbf{v}_1^T & — \\ — & \mathbf{v}_2^T & — \end{pmatrix} \begin{pmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{pmatrix} = \begin{pmatrix} \mathbf{v}_1^T\mathbf{v}_1 & \mathbf{v}_1^T\mathbf{v}_2 \\ \mathbf{v}_2^T\mathbf{v}_1 & \mathbf{v}_2^T\mathbf{v}_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

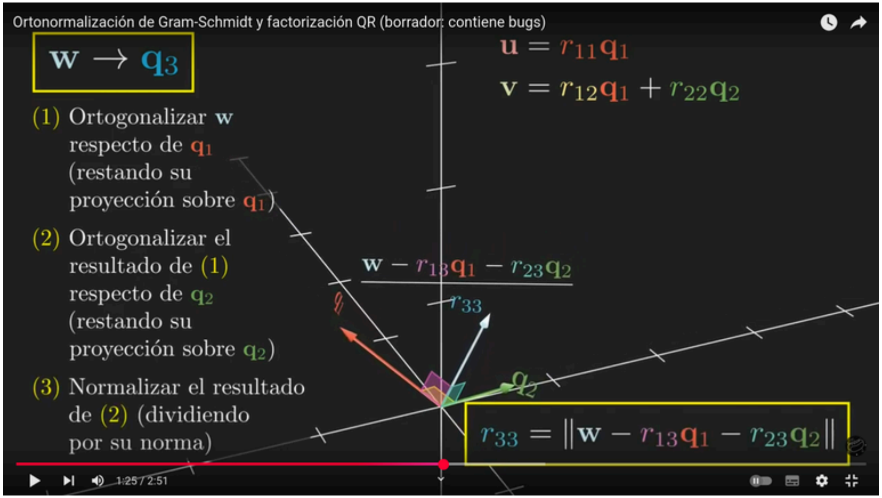
Gram-Schmidt [1]

Muy rara vez los vectores v_1 y v_2 parten siendo ortonormales, pero es posible construir un 2° conjunto de vectores q_1 y q_2 que sí lo sea y genere el mismo plano que v_1 y v_2.

Para construirlos, se puede usar el proceso de ortonormalización de Gram-Schmidt.

Gram-Schmidt [2]

Para entender el proceso, mira este video:
https://www.youtube.com/watch?v=uT1ZCDx0P6Y



Factorización QR [1]

El proceso genera la factorización QR:

$$\underbrace{\begin{pmatrix} | & | \\ \mathbf{v}_1 & \mathbf{v}_2 \\ | & | \end{pmatrix}}_{V} = \underbrace{\begin{pmatrix} | & | \\ \mathbf{q}_1 & \mathbf{q}_2 \\ | & | \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix}}_{R}$$

Factorización QR [2]

Reemplazando en las ecuaciones normales:

$$V^T V \bar{\alpha} = V^T \mathbf{b}$$

$$R^T \cancel{Q^T Q} R \bar{\alpha} = R^T Q^T \mathbf{b}$$

$$\cancel{R^T} R \bar{\alpha} = \cancel{R^T} Q^T \mathbf{b}$$

$$R \bar{\alpha} = Q^T \mathbf{b}$$

Factorización QR [3]

Entonces, si se posee la factorización V = QR, se puede resolver este sistema simplificado:

$$R\bar{\alpha} = Q^T\mathbf{b}$$

- *R* es una matriz triangular superior, así que se resuelve con backward substitution en $O(n^2)$.
- Este sistema está mucho mejor condicionado que el de las ecuaciones normales y, aún así, tiene la misma solución.

Factorización QR [4]

Ejercicio: encontrar la factorización $V = QR$ y resolver el sistema anterior.

$$V = \begin{pmatrix} 1 & 2 \\ 1 & 0 \\ 0 & -1 \end{pmatrix}$$

¿Dudas?