EL ESPACIO EUCLÍDEO

INTRODUCCIÓN.

Trataremos en este tema de llevar a los espacios vectoriales nociones geométricas como ortogonalidad, ángulo, longitud, distancias, áreas...

Veremos que todo ello se puede obtener al introducir la operación de **producto escalar**.

La geometría euclídea se desarrolla en los siglos XIX y XX, tras la aparición del concepto de espacio vectorial. Recibe su nombre en honor a Euclides, matemático griego (~300 a.C.) quien estudió los conceptos básicos de la Geometría plana, aunque por supuesto no en un contexto vectorial.

Para generalizar esos conceptos geométricos, observamos el comportamiento de los vectores del plano. En $\mathbb{R}^2$ tenemos definido el producto escalar usual

$$(a_1, a_2) \cdot (b_1, b_2) = a_1 b_1 + a_2 b_2$$

Es una operación entre dos vectores, cuyo resultado es un número, es decir, un escalar (de ahí el nombre "producto escalar").

El producto escalar permite reconocer a los vectores ortogonales ("ángulo recto"): dos vectores son ortogonales si su producto escalar es cero: por ejemplo, (1,3) y (–6,2), etc.

Observemos las propiedades de esta operación:

Propiedades del producto escalar usual.

1. Conmutativa. $u \cdot v = v \cdot u$
2. Distributiva. $u \cdot (v + w) = u \cdot v + u \cdot w$
3. Reubicación del escalar. $\alpha (u \cdot v) = (\alpha u) \cdot v = u \cdot (\alpha v)$
4. Definida positiva: $v \cdot v \geq 0$, y se da la igualdad $v \cdot v = 0$ solamente para el vector $v = \vec{0}$.

Definición: Producto escalar en cualquier espacio. Espacio euclídeo.

Cualquier operación en un espacio vectorial que cumpla las anteriores propiedades, diremos que es un producto escalar (aunque no se trate del producto escalar usual).

Por eso se habla de "un" producto escalar, y no "el" producto escalar, ya que hay muchas posibilidades para definirlo.

Llamaremos espacio euclídeo a un espacio vectorial dotado de un producto escalar.

El producto escalar se denotará por $u \cdot v$. También se puede utilizar la notación <u,v>.

Ejemplos de producto escalar.

1. El producto escalar usual en $\mathbb{R}^n$:

$(a_1,a_2,...,a_n) \cdot (b_1,b_2,...,b_n) = a_1 b_1 + a_2 b_2 + ... + a_n b_n$

Puede verse como el producto de una matriz fila por una matriz columna:

$$(a_1,a_2,...,a_n) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1 b_1 + a_2 b_2 + ... + a_n b_n$$

2. En $\mathbb{R}^3$ podemos "inventar" otra operación que cumpla también las propiedades anteriores, y por tanto podremos llamarla un producto escalar. Por ejemplo,

$(a,b,c) \cdot (a',b',c') = aa' + 2bb' + 3\,cc'$

Compruébese que cumple las propiedades.

3. En el espacio $\mathbf{M_2}$ de matrices 2x2 con términos reales, podemos definir el siguiente producto escalar:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} = aa'+bb'+cc'+dd'$$

Este producto escalar se obtiene de identificar $\mathbf{M_2}$ con el espacio $\mathbb{R}^4$, identificando la matriz $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con el vector (a,b,c,d), y aplicando el producto escalar usual.

También podemos definir este otro (comprobar que cumple las propiedades):

$A \cdot B =$ traza $(A\,A^t)$ (Nota: la traza de una matriz es la suma de sus elementos diagonales).

4. En el espacio $\mathbf{M_2}$, el producto ordinario de matrices no es un producto escalar. (Su resultado no es un escalar; además no es conmutativo, etc.)

5. En el espacio vectorial ***C***[a,b] de las funciones continuas en el intervalo [a,b], definimos el producto escalar:

$$f \cdot g = \int_b^a f(x)g(x)\,dx$$

Compruébese que cumple todas las propiedades de un producto escalar.

6. En el espacio $\mathbf{P_2}$**={** $ax^2 + bx + c : a, b, c \in \mathbb{R}$ **}** de los polinomios de grado ≤ 2, podemos definir el producto escalar

$(ax^2 + bx + c) \cdot (a' x^2 + b' x + c') = a\,a' + b\,b' + c\,c'$

Otra posibilidad es considerar a los polinomios como funciones continuas en un intervalo, y utilizar el producto escalar del ejemplo anterior.

Notar que el producto ordinario de polinomios no es un producto escalar (su resultado no es un escalar, etc.)

Conceptos geométricos obtenidos del producto escalar.

Por analogía con lo que ocurre en el plano o el espacio con el producto escalar usual, podemos definir los siguientes conceptos, siempre referidos a un cierto producto escalar.

Nos situamos en V, un espacio euclídeo.

1. Vectores ortogonales.

Dos vectores **u, v** son ortogonales si su producto escalar es cero: **u · v = 0.** Se denota $\mathbf{u} \perp \mathbf{v}$.

Diremos que un conjunto de vectores es un conjunto ortogonal si cada uno de ellos es ortogonal a todos los demás. (Exigimos además que ninguno de los vectores sea el $\vec{0}$).

- Notar que si dos vectores **u, v** son ortogonales entonces también lo son sus múltiplos $\alpha\,\mathbf{u}$ y $\beta\,\mathbf{v}$ (α, β escalares).

Si un conjunto ortogonal es además base de un espacio o subespacio, diremos que es una base ortogonal del mismo.

2. Norma o módulo de un vector.

La norma o módulo de un vector es $|\,v\,| = \sqrt{v \cdot v}$ La noción corresponde, intuitivamente, a la "longitud" de la flecha que representa el vector. También se puede denotar $\|\,v\,\|$.

Si un vector tiene norma 1, se dice que es unitario.

- Con cualquier producto escalar, el único vector de módulo cero es el $\vec{0}$.
- Notar también que el módulo de un vector es el mismo que el de su opuesto, y que el módulo de $\alpha\mathbf{v}$ es $|\,\alpha\mathbf{v}\,| = |\,\alpha|\;|\,\mathbf{v}\,|$ (es decir, el módulo queda multiplicado por el valor absoluto del escalar).
- Además se cumple para cualesquiera **u, v** la desigualdad triangular: $|\,\mathbf{u}+\mathbf{v}\,| \leq |\,\mathbf{u}\,| + |\,\mathbf{v}\,|$

  El nombre de esta desigualdad se debe a la forma en que se suman vectores gráficamente. Si $|\,\mathbf{u}\,|$ y $|\,\mathbf{v}\,|$ son los lados de un triángulo, $|\,\mathbf{u}+\mathbf{v}\,|$ es el tercer lado y por tanto ha de ser menor que la suma de los otros dos.

3. Distancia entre dos vectores.

La distancia entre **u** y **v** es la norma del vector diferencia entre ambos.

$$\mathbf{dist\,(u, v) = |\,u - v\,|}$$

4. Ángulo entre dos vectores.

Es sabido que para el producto escalar usual de $\mathbb{R}^2$ se tiene que **u · v = | u | | v |** cos α, donde α es el ángulo que forman ambos vectores. Por tanto, para generalizar la noción de ángulo a cualquier espacio euclídeo, definimos

$$\textbf{ángulo(u, v) = arc cos } \frac{u \cdot v}{|u|\,|v|}$$

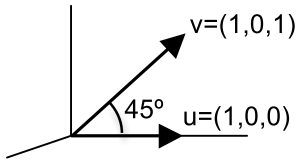
Notar que si alguno de los dos vectores es nulo, no podemos dividir por su módulo y por tanto el ángulo no está definido. En efecto, geométricamente el vector nulo no forma ángulo ninguno.

Ejemplos.

1. En $\mathbb{R}^3$ con el producto escalar usual:

Sea u=(1,0,0), v=(1,0,1).

- Sus módulos son: $|u| = \sqrt{(1,0,0)\begin{pmatrix}1\\0\\0\end{pmatrix}} = \sqrt{1} = 1$, $|v| = \sqrt{(1,0,1)\begin{pmatrix}1\\0\\1\end{pmatrix}} = \sqrt{2}$
- Su distancia es | v – u | = | (0,0,1) | = 1
- El ángulo que forman es arc cos $\dfrac{(1,0,0)\begin{pmatrix}1\\0\\1\end{pmatrix}}{1\cdot\sqrt{2}}$ = arc cos $\dfrac{1}{\sqrt{2}}$ = 45º.



2. En $\mathbb{R}^2$ con un producto escalar diferente:

Dados dos vectores u=(a,b), v=(a', b') en $\mathbb{R}^2$, definimos su producto escalar como:

u · v = 2aa' – ab' – ba' + bb'

Calculemos el producto escalar de los vectores u=(2,1) y v=(1,3):

u · v = 2·2·1 – 2·3 – 1·1 + 1·3 = 0

Entonces, con este producto escalar, los vectores (2,1) y (1,3) son ortogonales, a pesar de que no lo son para el producto escalar usual.

Por tanto, este producto escalar define *una geometría diferente* en el plano, con todas sus nociones de ángulos, distancias, etc. que no son como los usuales pero son igualmente coherentes.

3. **En el espacio *C*[0,1] de funciones continuas en el intervalo [0, 1]**

Sean los vectores (funciones) $f(x)=x^2$, $g(x)=x+1$.

Respecto al producto escalar $f \cdot g = \int_b^a f(x)g(x)\,dx$ tenemos:

- Sus módulos son: $|f| = \sqrt{\int_0^1 x^2 \cdot x^2 \, dx} = \frac{1}{\sqrt{5}}$, $|g| = \sqrt{\int_0^1 (x+1)\,(x+1)\, dx} = \sqrt{\frac{7}{3}}$
- Su distancia es: $\sqrt{\int_0^1 \left[x^2 - (x+1) \right]^2 dx} = \sqrt{\frac{41}{30}}$
- Ángulo que forman: arc cos $\frac{f \cdot g}{|f||g|}$ =arc cos $\frac{\int_0^1 x^2(x+1)\,dx}{\frac{1}{\sqrt{5}} \cdot \sqrt{\frac{7}{3}}}$ =arc cos $\frac{\frac{7}{12}}{\frac{1}{\sqrt{5}} \cdot \sqrt{\frac{7}{3}}}$ = 0,547 rad

Como se ve en los ejemplos, si las anteriores nociones se aplican a $\mathbb{R}^2$ o $\mathbb{R}^3$ con el producto escalar usual, producen los conceptos geométricos usuales. Si cambiamos el producto escalar o trabajamos en otro espacio vectorial las nociones de longitud, ángulo, etc. serán diferentes, formando otra geometría.

Teorema (Desigualdad de Cauchy-Schwarz)

Para cualesquiera u, v en un espacio euclídeo, se tiene que su producto escalar, en valor absoluto, es menor o igual que el producto de sus módulos.

| u · v | $\leq$ **| u | · | v |**

y además la igualdad **| u · v | = | u | · | v |** sólo se da cuando **u** es múltiplo de **v**.

Teorema.

Todo conjunto ortogonal es linealmente independiente.

Demostración (sólo la hacemos para dos vectores)

Sabemos que el conjunto {u, v} es ortogonal, es decir, u · v =0. Debemos ver que es linealmente independiente. Para ello, bastará ver que un vector no es múltiplo del otro.

En efecto, si así fuera, tendríamos u=αv. Entonces al ser ortogonales,

u · v =0 → (αv) · v =0 → α(v · v) = 0 → o bien α= 0, o bien v · v =0 (y por tanto v=0).

Como ni αni v pueden ser nulos, concluimos que no es posible que sea u=αv. Por tanto {u, v} son linealmente independientes.

Normalización. Conjunto ortonormal.

- **Normalizar** un vector es reducirlo a otro vector (múltiplo suyo) de norma 1. Es decir, es hacerlo unitario conservando su dirección y sentido.

Ello se consigue dividiendo el vector por su módulo, es decir, multiplicándolo por el número $\frac{1}{|\,v\,|}$

Ejemplo: El vector (3,4) tiene norma 5 ("mide" 5 unidades de longitud). Por tanto, al dividirlo por 5 se hará unitario.

$\frac{(3,4)}{|\,(3,4)\,|} = = \frac{(3,4)}{5} = \left(\frac{3}{5},\frac{4}{5}\right)$ es el vector normalizado.

Comprobamos que es múltiplo del anterior y que es unitario: $\left(\frac{3}{5}\right)^2 + \left(\frac{4}{5}\right)^2 = 1.$

- Se llama **conjunto ortonormal** a un conjunto ortogonal de vectores unitarios.

Se puede obtener normalizando un conjunto ortogonal.

Si un conjunto ortonormal es además base de un espacio o subespacio, diremos que es una base ortonormal del mismo.

Propiedad de los conjuntos ortonormales:

Si tenemos un conjunto ortonormal y hacemos el producto escalar de los vectores del conjunto dos a dos, obtendremos 0 (si los vectores son distintos) o bien 1 (si los dos vectores son el mismo). Es decir:

u · v = 0 para u ≠ v.

Dos vectores distintos en el conjunto presentan producto escalar 0, ya que son ortogonales;

u · u = 1 .

un vector del conjunto multiplicado escalarmente por sí mismo resulta 1, ya que es unitario.

Ejemplos:

1) La base canónica (1,0,0), (0,1,0), (0,0,1) de $\mathbb{R}^3$ es un conjunto ortonormal de tres vectores. Es además una base ortonormal de $\mathbb{R}^3$.

2) El conjunto (3,0,0), (0,0,2) es un conjunto ortogonal de dos vectores en $\mathbb{R}^3$. (No es una base). Si lo normalizamos, obtenemos el conjunto (1,0,0), (0,0,1) que es ortonormal.

Definición (Matriz ortogonal).

Una matriz cuadrada nxn se dice que es una matriz ortogonal si sus columnas son **n** vectores ortonormales de $\mathbb{R}^n$ (con el producto escalar usual), es decir, si sus columnas son una base ortonormal de $\mathbb{R}^n$.

(Nota. No nos dejemos confundir por la nomenclatura, pues se llama matriz ortogonal, a pesar de que sus columnas son ortonormales y no sólo ortogonales.)

Ejemplos. $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ en $\mathbb{R}^3$; $\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$ en $\mathbb{R}^2$.

Propiedades de las matrices ortogonales.

1. Como sus columnas forman base, toda matriz ortogonal es regular o inversible (su determinante es no nulo).

2. Al multiplicar $A^t A$ se obtiene la matriz identidad. En efecto, podemos considerar que A^t tiene los vectores ortonormales en las filas, mientras que A los tiene en columnas. Al hacer el producto de matrices, vamos multiplicando cada fila por cada columna. Entonces, por la anterior propiedad de los conjuntos ortonormales, cada vector multiplicado por sí mismo resulta 1, y cada vector multiplicado por otro distinto da 0. Así se obtiene la matriz identidad.

3. Como consecuencia de 3), una matriz A es ortogonal si y solo si su inversa coincide con su traspuesta. $A^{-1} = A^t$.

En lo sucesivo veremos la utilidad de las bases ortonormales (p.ej. la base canónica de $\mathbb{R}^n$), pues es más sencillo trabajar con ellas que con bases cualesquiera.

SUBESPACIOS ORTOGONALES.

Definición (vector ortogonal a un subespacio)

Se dice que un vector **v** es ortogonal a un subespacio **S** (y se denota **v** $\perp$ **S**) si **v** es ortogonal a todos los vectores de **S**. (Basta con que **v** sea ortogonal los vectores de una base de **S**).

Ejemplo. En $\mathbb{R}^3$, el vector (0,0,1) es ortogonal al plano XY.

Definición: Subespacios ortogonales.

Diremos que **un subespacio S es ortogonal a otro subespacio T** (se denota $S \perp T$) si todo vector de S es ortogonal a todo vector de T, es decir:

$u \cdot v = 0$ para todo $u \in S$, $v \in T$.

Basta con que los vectores de una base de S sean ortogonales a los vectores de una base de T.

Propiedad: Si dos subespacios son ortogonales entonces su intersección ha de ser $\{\vec{0}\}$.

En efecto, si tuviéramos un $v \neq \vec{0}$ en la intersección, tendríamos $\mathbf{v} \cdot \mathbf{v} \neq \mathbf{0}$ con $\mathbf{v} \in \mathbf{S}$, $\mathbf{v} \in \mathbf{T}$. Ello impide que los subespacios sean ortogonales.

Ejemplos:

1) En $\mathbb{R}^3$, el eje X es ortogonal al plano YZ, ya que cualquier vector de la forma $(\alpha,0,0)$ es ortogonal a cualquier vector de la forma $(0, \beta,\gamma)$. Esto responde a la noción geométrica de dos subespacios que se cortan perpendicularmente, y que se cortan solamente en el $\{\vec{0}\}$.

2) El plano XY no es ortogonal al plano YZ, ya que no es cierto que todo vector de uno de los planos sea ortogonal a todo vector del otro. Por ejemplo, (1,1,0) en el plano XY no es ortogonal a (0,1,1) en el plano YZ.

Otra manera de ver que estos subespacios no son ortogonales, es ver que su intersección es el eje Y, y por tanto no es $\{\vec{0}\}$.

Por ello, la ortogonalidad de subespacios no es exactamente lo mismo que la noción geométrica de “estar colocado en ángulo recto”. Se requiere además que la intersección de ambos subespacios sea $\{\vec{0}\}$.

Definición: Complemento ortogonal.

• Recordemos que en un espacio vectorial, dado un subespacio S podemos definir su suplementario (o complementario) como T tal que $S \oplus T$ sea el espacio total.

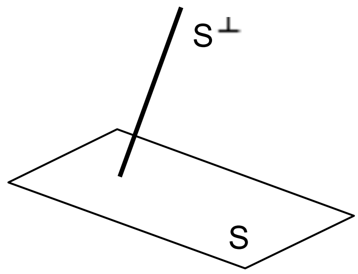
Todo subespacio S (salvo el $\{\vec{0}\}$ y el total) tienen infinitos suplementarios, pero solo uno de ellos es ortogonal a S.

Dado un subespacio S, su complemento ortogonal es el único subespacio (denotado por $S^{\perp}$) que cumple:

- **dim S + dim $S^{\perp}$ = n** (donde n es la dimensión del espacio total),
- **$S^{\perp}$ es ortogonal a S**.

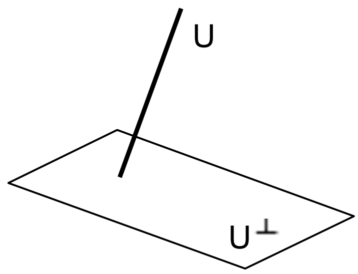
Es decir, el complemento ortogonal responde a la idea de "lo que le falta" al subespacio S para completar el total, colocado de forma ortogonal a él.

Ejemplo.



En $\mathbb{R}^3$, el subespacio S=planoXY tiene infinitos suplementarios (toda recta que no esté contenida en S). Pero de ellos solo uno es su complemento ortogonal, que es el eje Z.

Propiedades:



- $S^{\perp}$ está formado por todos los vectores del espacio que son ortogonales a S.
- $S^{\perp}$ es el mayor posible de todos los subespacios ortogonales a S.
- El complemento ortogonal es una relación recíproca entre subespacios. Es decir, que si S es el complemento ortogonal de T, entonces T lo es de S.

Construcción del complemento ortogonal.

Nota. En lo sucesivo, la palabra "complemento" significará "complemento ortogonal".

Se trata de encontrar todos los vectores ortogonales a S. Basta con que sean ortogonales a su base.

Por tanto planteamos un sistema de ecuaciones como en el ejemplo siguiente. Es un sistema compatible indeterminado, cuyo espacio solución será (en forma paramétrica) $S^{\perp}$.

Observemos que la dimensión de $S^{\perp}$ (número de parámetros) deberá ser n – dim S.

Ejemplo.

Calculemos el complemento del siguiente subespacio de $\mathbb{R}^4$: U= { $(\alpha, 0, 2\alpha, \alpha+\beta) : \alpha,\beta\in \mathbb{R}$ }

Primero necesitamos una base de U: esta es (1,0,2,1), (0,0,0,1).

Buscamos los vectores que sean ortogonales a ambos, es decir, los (x,y,z,t) tales que:

$$(1, 0, 2, 1)\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0, \quad (0, 0, 0, 1)\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0 \quad \rightarrow \quad \text{por tanto,} \quad \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

sistema compatible indeterminado cuya solución es $\{ (2\lambda,\ \mu, -\lambda, 0) : \lambda, \mu \in \mathbb{R} \}$

Entonces una base de $U^{\perp}$ será (2, 0, –1, 0), (0, 1, 0, 0).

Notar que efectivamente dim $U^{\perp}$ + dim U = 2+2 = 4, la dimensión total del espacio.

Comprobar también que cada uno de los vectores de la base de $U^{\perp}$ es ortogonal a cada uno de los vectores de la base de U.

Observación.

No debe confundirse esto con el hecho de que la base de U sea o no una base ortogonal, y lo mismo la base de su complemento $U^{\perp}$.

En el ejemplo anterior, tenemos:

Base de U: (1,0,2,1), (0,0,0,1). El producto escalar de estos dos vectores entre sí no es 0, lo cual indica que esta no es una <u>base ortogonal</u> dentro del subespacio U. Para este ejercicio no es necesario que lo sea.

Base de $U^{\perp}$: (2,0,–1,0), (0,1,0,0). El producto escalar de estos dos vectores entre sí es 0, lo cual indica que se trata de una <u>base ortogonal</u> dentro del subespacio $U^{\perp}$, aunque no era necesario.

Ahora bien, el producto de cada vector de la base de U por cada vector de la base de $U^{\perp}$ (es decir, multiplicando los vectores de <u>un</u> subespacio por los del <u>otro</u>) sí es cero:

(1,0,2,1) · (2,0,–1,0) = 0

(1,0,2,1) · (0,1,0,0) = 0

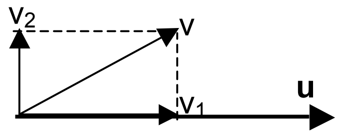
(0,0,0,1) · (2,0,–1,0) = 0

(0,0,0,1) · (0,1,0,0) = 0

y esto indica que ambos subespacios, U y $U^{\perp}$, son <u>ortogonales entre sí</u>, como pedía el ejercicio.

PROYECCIONES ORTOGONALES.

1. Proyección ortogonal de un vector sobre otro.



Proyección de un vector **v** sobre otro vector **u**:

v se puede descomponer de manera única como **v = $\mathbf{v_1}$ + $\mathbf{v_2}$** con **$\mathbf{v_1}$** en la dirección de **u**, y **$\mathbf{v_2}$** ortogonal a **u.**

La componente **$\mathbf{v_1}$** se llama proyección ortogonal de **v** sobre **u**, y se denota **proy $_{\mathbf{u}}$ (v).**

Se calcula : $$\boxed{\mathbf{proy}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathrm{u} \cdot \mathbf{v}}{\mathrm{u} \cdot \mathrm{u}}\, \mathbf{u}}$$

Notar que esta expresión es un escalar multiplicado por **u.**

Así pues, **proy $_{\mathbf{u}}$ (v)** es un múltiplo de **u**. Dicho de otro modo, al proyectar sobre un vector **u**, obtenemos siempre vectores en la dirección de **u**.

Una vez hallada la componente **$\mathbf{v_1}$** , se puede calcular la otra componente **$\mathbf{v_2}$** como

$$\mathbf{v_2} = \mathbf{v} - \mathbf{v_1}$$

Ejemplo:

En $\mathbb{R}^2$, proyectamos **v**=(1,2) sobre **u**=(3,1).

$$\mathbf{proy}_{\mathbf{u}}(\mathbf{v}) = \frac{\mathrm{u} \cdot \mathrm{v}}{\mathrm{u} \cdot \mathrm{u}}\mathrm{u} = \frac{(3,1)\begin{pmatrix} 1 \\ 2 \end{pmatrix}}{(3,1)\begin{pmatrix} 3 \\ 1 \end{pmatrix}}(3,1) = \frac{5}{10}(3,1) = \left(\frac{3}{2},\frac{1}{2}\right)$$

Así **$\mathbf{v_1}$ =** $\left(\frac{3}{2},\frac{1}{2}\right)$, y la otra componente es **$\mathbf{v_2}$** = **v – $\mathbf{v_1}$** = (1,2) – $\left(\frac{3}{2},\frac{1}{2}\right) = \left(\frac{-1}{2},\frac{3}{2}\right)$

Entonces el vector **v** queda expresado como (1,2) = $\left(\frac{3}{2},\frac{1}{2}\right) + \left(\frac{-1}{2},\frac{3}{2}\right)$, dos componentes de las cuales la primera va en la dirección de **u** y la segunda es ortogonal a **u** (compruébese).

En particular, las dos componentes son ortogonales entre sí:

$\left(\frac{3}{2},\frac{1}{2}\right) \cdot \left(\frac{-1}{2},\frac{3}{2}\right)$ =0. Este es un modo útil de comprobar que los cálculos son correctos.

2. Proyección ortogonal de un vector sobre un subespacio.

Lo que hemos hecho hasta ahora es proyectar **sobre un vector**, que se puede ver como una proyección **sobre una recta** (la recta definida por el vector).

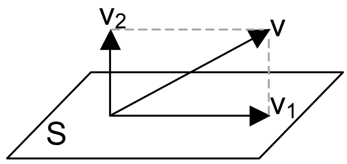
También se puede proyectar sobre otros subespacios que no sean una recta (un plano, o subespacios de mayor dimensión en $\mathbb{R}^n$). Veámoslo.

Dado un vector **v** y un subespacio S, **v** se puede descomponer de manera única como

$\mathbf{v} = \mathbf{v_1} + \mathbf{v_2}$, con $\mathbf{v_1} \in S$, y $\mathbf{v_2}$ ortogonal a S**.**

La componente $\mathbf{v_1}$ se llama proyección ortogonal de **v** sobre S, y se denota **proy $_S$ (v).**

Se calcula : $\boxed{\textbf{proy}_{\,s}\,\textbf{(v)} = \text{proy}_{u_1}(v) + \ldots + \text{proy}_{u_n}(v)}$



es decir: $\boxed{\textbf{proy}_{\,s}\,\textbf{(v)} = \dfrac{u_1 \cdot \mathbf{v}}{u_1 \cdot u_1} u_1 + \ldots + \dfrac{u_n \cdot \mathbf{v}}{u_n \cdot u_n} u_n}$

(Fórmula de la Proyección)

donde $\{u_1, \ldots, u_n\}$ son una base ortogonal de S.

(Si además la base es ortonormal, los denominadores pueden suprimirse, pues valen 1).

Igualmente, una vez hallada la componente $\mathbf{v_1}$, se puede calcular $\mathbf{v_2}$ como

$\mathbf{v_2} = \mathbf{v} - \mathbf{v_1}$

Observación.

Notar que la anterior fórmula de la proyección **solo es válida si la base de S es ortogonal.**

Si no disponemos de una base ortogonal de S, deberemos calcular una, o bien utilizar otros métodos que se verán (**descomposición por coordenadas** o **matriz de proyección**).

Ejemplo:

En $\mathbb{R}^3$, proyectamos v=(3,2,2) sobre el subespacio S generado por: u_1= (2,0,1), u_2= (0,3,0).

Lo primero, observamos que u_1 , u_2 forman base de S, y además base ortogonal. Por tanto puede utilizarse la fórmula de la proyección:

$$\text{proy}_s(v) = \frac{u_1 \cdot v}{u_1 \cdot u_1} u_1 + \frac{u_2 \cdot v}{u_2 \cdot u_2} u_2 = \frac{(2,0,1)\begin{pmatrix}3\\2\\2\end{pmatrix}}{(2,0,1)\begin{pmatrix}2\\0\\1\end{pmatrix}} u_1 + \frac{(0,3,0)\begin{pmatrix}3\\2\\2\end{pmatrix}}{(0,3,0)\begin{pmatrix}0\\3\\0\end{pmatrix}} u_2 = \frac{8}{5} u_1 + \frac{6}{9} u_2 = \frac{8}{5}(2,0,1) + \frac{6}{9}(0,3,0) =$$

$= \left(\frac{16}{5},2,\frac{8}{5}\right)$

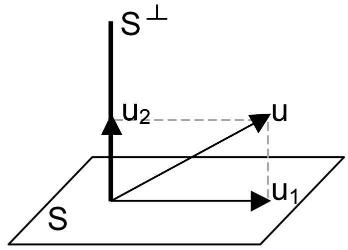
Así $\mathbf{v_1} = \left(\frac{16}{5},2,\frac{8}{5}\right)$, y por tanto $\mathbf{v_2} = v - v_1 = (3,2,2) - \left(\frac{16}{5},2,\frac{8}{5}\right) = \left(\frac{-1}{5},0,\frac{2}{5}\right)$

y el vector **v** queda expresado como $(3,2,2) = \left(\frac{16}{5},2,\frac{8}{5}\right) + \left(\frac{-1}{5},0,\frac{2}{5}\right)$, dos componentes de las cuales la primera, $\mathbf{v_1}$, pertenece a S y la segunda, $\mathbf{v_2}$, es ortogonal a S.

Se puede comprobar, para ver si es correcto, que las dos componentes son ortogonales entre sí: $\left(\frac{16}{5},2,\frac{8}{5}\right) \cdot \left(\frac{-1}{5},0,\frac{2}{5}\right) = 0$

(Este mismo ejemplo lo haremos después por otros métodos).

Observación.



$u_1 = \text{proy}_S(u)$

$u_2 = \text{proy}_{S^\perp}(u)$

Dado cualquier subespacio S, se tiene que $S \oplus S^\perp = V$, donde V es el espacio total.

Por tanto, todo vector $u \in V$ puede descomponerse como:

$$\mathbf{u = u_1 + u_2}, \qquad \text{con } u_1 \in S,\ u_2 \in S^\perp.$$

De hecho, estas dos componentes no son otras que las proyecciones ortogonales de u sobre S y sobre $S^\perp$, es decir,

$$\boxed{\mathbf{u = \text{proy}_S(u) + \text{proy}_{S^\perp}(u)}}$$

Así pues, puede calcularse primero la componente $u_1 = \text{proy}_S(u)$ (con la fórmula de la proyección o la matriz de proyección) y obtener después la componente u_2 como $u_2 = u - u_1$.

Otra posibilidad es a la inversa: empezar calculando $u_2 = \text{proy}_{S^\perp}(u)$ para luego obtener $u_1 = u - u_2$.

Ejemplo.

Anteriormente hemos proyectado v = (3,2,2) sobre el subespacio S generado por (2,0,1), (0,3,0), obteniendo

$$v_1 = \text{proy}_S(v) = \left(\frac{16}{5},2,\frac{8}{5}\right), \qquad v_2 = v - v_1 = (3,2,2) - \left(\frac{16}{5},2,\frac{8}{5}\right) = \left(\frac{-1}{5},0,\frac{2}{5}\right)$$

Otro modo de resolverlo es hallar primero v_2 como $\text{proy}_{S^\perp}(u)$.

Para ello hallamos $S^\perp$, que es la recta generada por w = (1,0,–2) (compruébese). Entonces,

v_2 = proy $_{S^\perp}$ (u) = proy $_w$ (u) = $\dfrac{w \cdot u}{w \cdot w} w = \dfrac{-1}{5}(1,0,-2) = \left(\dfrac{-1}{5}, 0, \dfrac{2}{5}\right)$

y de ahí $v_1 = v - v_2 = (3,2,2) - \left(\dfrac{-1}{5}, 0, \dfrac{2}{5}\right) = \left(\dfrac{16}{5}, 2, \dfrac{8}{5}\right)$

Observaciones.

1. Notar que si S es un subespacio de dimensión 1, cuya base es un solo vector **u**, entonces proyectar sobre S es lo mismo que proyectar sobre **u**. Esta proyección es fácil de calcular con la fórmula.

2. Si un vector **w** ya está en el subespacio S, al proyectarlo sobre S no varía. La descomposición $\mathbf{w = w_1 + w_2}$ sería en realidad $\mathbf{w = w + 0}$, donde $\mathbf{w} \in S$, $\mathbf{0} \in S^\perp$.

3. Si un vector **u** está en el subespacio $S^\perp$, entonces ocurre lo contrario: la descomposición $\mathbf{u = u_1 + u_2}$ sería en realidad $\mathbf{u = 0 + u}$, donde $\mathbf{0} \in S$, $\mathbf{u} \in S^\perp$.

Ejemplo: Subespacio S anterior, generado por (2,0,1), (0,3,0).

Descomponemos el vector w = (4,3,2). Haciendo el cálculo se tiene $\text{proy}_s(w)$ = (4,3,2).

Ello ocurre porque $w \in S$ (notar que el determinante que forma w con la base de S es nulo, lo que significa que es linealmente dependiente de dicha base).

Por tanto se descompone como (4,3,2) = (4,3,2) + (0,0,0), con (4,3,2) $\in S$, (0,0,0)$\in S^\perp$.

Otro método: La descomposición por coordenadas.

Basándonos en que $S \oplus S^\perp = V$, podemos formar una base del espacio total V de modo que sus primeros n vectores estén en S (donde n=dim S) y el resto estén en $S^\perp$. Dicho de otro modo, unimos una base de S y una base de $S^\perp$ para formar una base del espacio total V.

Entonces, para proyectar un vector **u**, solo tenemos que expresarlo en coordenadas respecto a esta base. La parte correspondiente a S será $\mathbf{u_1}$ **,** y la parte correspondiente a $S^\perp$ será $\mathbf{u_2}$.

Ejemplo: Subespacio S anterior, generado por (2,0,1), (0,3,0), proyectamos v = (3,2,2).

Para ello formamos una base del espacio total $\mathbb{R}^3$ de modo que sus dos primeros vectores pertenezcan a S, y el tercer vector a $S^\perp$.

Los dos primeros vectores pueden ser los que ya tenemos de S (no hace falta que sean ortogonales entre sí, aunque en este caso lo son): (2,0,1), (0,3,0).

El tercer vector ha de ser ortogonal a estos dos, por tanto ha de cumplir:

$$\left.\begin{aligned}(2,0,1) \cdot (x,y,z) &= 0\\ (0,3,0) \cdot (x,y,z) &= 0\end{aligned}\right\} \Rightarrow \left.\begin{aligned}2x+z &= 0\\ 3y &= 0\end{aligned}\right\} \Rightarrow (\alpha, 0, -2\alpha)$$

de donde se obtiene una base de $S^\perp$, el vector (1,0,–2)

Así formamos una base del total $\mathbb{R}^3$: $\{ \underbrace{(2,0,1),\ (0,3,0)}_{\in S},\ \underbrace{(1,0,-2)}_{\in S^{\perp}} \}$

A continuación expresamos el vector que queremos proyectar, v = (3,2,2), como combinación lineal de esta base.

$$(3,2,2) = \alpha(2,0,1) + \beta\,(0,3,0) + \gamma\,(1,0,-2) \rightarrow \alpha=\frac{8}{5},\ \beta=\frac{2}{3},\ \gamma=\frac{-1}{5}$$

$$(3,2,2) = \underbrace{\frac{8}{5}(2,0,1) + \frac{2}{3}(0,3,0)}_{\in S} + \underbrace{\frac{-1}{5}(1,0,-2)}_{\in S^{\perp}} = \underbrace{\left(\frac{16}{5},2,\frac{8}{5}\right)}_{\in S} + \underbrace{\left(\frac{-1}{5},0,\frac{2}{5}\right)}_{\in S^{\perp}}$$

que es la misma descomposición que se obtuvo por otro método.

Coordenadas de un vector en una base ortogonal.

Como decíamos anteriormente, si un vector **w** ya está en el subespacio S, al proyectarlo sobre S no varía, por tanto su descomposición $\mathbf{w} = \mathbf{w_1} + \mathbf{w_2}$ sería en realidad $\mathbf{w} = \mathbf{w} + \mathbf{0}$.

Esto nos proporciona una manera de obtener las coordenadas de un vector $w \in S$ respecto a una base ortogonal de S.

En efecto, si $w \in S$, y si tenemos una base ortogonal $\{ u_1, \ldots, u_n \}$ de S, entonces proyectamos w y no varía, con lo cual

$$w = \text{proy}_S(w) = \frac{u_1 \cdot w}{u_1 \cdot u_1} u_1 + \ldots + \frac{u_n \cdot w}{u_n \cdot u_n} u_n$$

y esto significa (por la definición de coordenadas) que los escalares $\frac{u_1 \cdot w}{u_1 \cdot u_1}, \ldots, \frac{u_n \cdot w}{u_n \cdot u_n}$ son las coordenadas de w en la base dada $\{ u_1, \ldots, u_n \}$.

<u>Además si la base es ortonormal</u>, los denominadores pueden suprimirse pues valen 1, y así las coordenadas son simplemente $(\mathbf{u_1 \cdot w}, \ldots, \mathbf{u_n \cdot w})$.

<u>Ejemplo</u>.

Consideremos en $\mathbb{R}^2$ la base (1,1), (1,–1) que es ortogonal. Hallemos las coordenadas de w=(5,4) en esta base, que son:

$$\left(\frac{(1,1)\begin{pmatrix}5\\4\end{pmatrix}}{(1,1)\begin{pmatrix}1\\1\end{pmatrix}},\ \frac{(1,-1)\begin{pmatrix}5\\4\end{pmatrix}}{(1,-1)\begin{pmatrix}1\\-1\end{pmatrix}} \right) = \left(\frac{9}{2},\frac{1}{2}\right)$$

En efecto, comprobamos que $(5,4)=\frac{9}{2}(1,1) + \frac{1}{2}(1,-1)$.

CÁLCULO DE BASES ORTOGONALES

Hemos visto que para aplicar la fórmula de la proyección sobre un subespacio S, se requiere una base ortogonal de S. Veremos cómo obtener una base ortogonal de S, si la que tenemos no lo es.

Proceso de ortogonalización de Gram-Schmidt

Supongamos que tenemos una base { $u_1, \dots, u_n$ } de S.

Vamos a construir otra base de S, { $a_1, \dots, a_n$ } que será ortogonal.

Como primer vector tomamos el propio u_1.

$\mathbf{a_1} = \mathbf{u_1}$

Para construir el segundo vector tomamos u_2 y lo proyectamos sobre el anteriormente construido $\mathbf{a_1}$, quedándonos con la componente ortogonal a $\mathbf{a_1}$.

$\mathbf{a_2} = \mathbf{u_2} - \text{proy}_{a_1}(u_2)$

Así { $\mathbf{a_1}$, $\mathbf{a_2}$ } son dos vectores que están en S (pues todas las operaciones se han hecho entre vectores de S, dentro del subespacio) y además son ortogonales entre sí.

Para construir el tercer vector tomamos u_3 y lo proyectamos sobre los anteriormente construidos $\mathbf{a_1}$, $\mathbf{a_2}$ quedándonos con la componente ortogonal a ellos.

$\mathbf{a_3} = \mathbf{u_3} - \text{proy}_{a_1}(u_3) - \text{proy}_{a_2}(u_3)$

Etc.

Este proceso conduce a una base <u>ortogonal</u> de S. Si queremos obtener una base <u>ortonormal</u>, bastará normalizar finalmente los vectores $\mathbf{a_i}$ obtenidos (o bien normalizarlos en cada paso).

Ejemplo.

Obtener una base ortogonal de $\mathbb{R}^3$ a partir de la base $\mathbf{u_1}=(1,1,0)$, $\mathbf{u_2}=(0,1,1)$, $\mathbf{u_3}=(1,0,1)$.

$\mathbf{a_1} = \mathbf{u_1} = (1,1,0)$

$$\mathbf{a_2} = \mathbf{u_2} - \text{proy}_{a_1}(u_2) = \mathbf{u_2} - \frac{a_1 \cdot u_2}{a_1 \cdot a_1} a_1 = (0,1,1) - \frac{(1,1,0)\begin{pmatrix}0\\1\\1\end{pmatrix}}{(1,1,0)\begin{pmatrix}1\\1\\0\end{pmatrix}}(1,1,0) = (0,1,1) - \frac{1}{2}(1,1,0) = \left(\frac{-1}{2},\frac{1}{2},1\right)$$

$$\mathbf{a_3} = \mathbf{u_3} - \text{proy}_{a_2}(u_3) - \text{proy}_{a_3}(u_3) = \mathbf{u_3} - \frac{a_1 \cdot u_3}{a_1 \cdot a_1} a_1 - \frac{a_2 \cdot u_3}{a_2 \cdot a_2} a_2 = (1,0,1) - \frac{1}{2}(1,1,0) - \frac{1}{3}\left(\frac{-1}{2},\frac{1}{2},1\right) = \left(\frac{2}{3},\frac{-2}{3},\frac{2}{3}\right)$$

Por tanto la base ortogonal es $(1,1,0)$, $\left(\frac{-1}{2},\frac{1}{2},1\right)$, $\left(\frac{2}{3},\frac{-2}{3},\frac{2}{3}\right)$. (Compruébese que es ortogonal).

Para hacerla ortonormal dividimos cada vector por su norma, obteniéndose la base

$$\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2},0\right),\ \left(\frac{-\sqrt{6}}{6},\frac{\sqrt{6}}{6},\frac{\sqrt{6}}{3}\right),\ \left(\frac{\sqrt{3}}{3},\frac{-\sqrt{3}}{3},\frac{\sqrt{3}}{3}\right)$$ (Compruébese que es ortonormal).

LA MATRIZ DE PROYECCIÓN.

Veamos otra manera de hallar la proyección ortogonal de un vector sobre un subespacio S, sin necesidad de hallar una base ortogonal de éste. El siguiente método solamente es válido para subespacios de $\mathbb{R}^n$ y con el producto escalar usual.

Dada una base de S, no necesariamente ortogonal,

1. Formamos la matriz A que contiene la base en sus columnas. Tendrá pues, tantas columnas como indique dim S. (No es necesariamente cuadrada).

2. Formamos la matriz cuadrada

$$\boxed{\mathbf{P_S = A\,(A^t A)^{-1} A^t}}$$

(la existencia de la inversa de $A^t A$ está garantizada, por haber formado A con una base).

La matriz $\mathbf{P_S}$ se llama **matriz de proyección sobre el subespacio S.**

Esta matriz sirve para proyectar sobre S cualquier vector v :

$$\boxed{\mathbf{proy_S(v) = P_S \cdot v}}$$

Por tanto, es especialmente útil si tenemos que proyectar varios vectores sobre un mismo subespacio.

La matriz de proyección de un subespacio S es única, y no depende de la base de la que hayamos partido.

Ejemplo.

En $\mathbb{R}^3$, vamos a proyectar v = (3,2,2) sobre el subespacio S generado por: u_1= (2,0,1), u_2 = (0,3,0). (Ya lo hicimos por otro método, comprobaremos que el resultado es el mismo).

Tomamos la base (2,0,1), (0,3,0), (que es ortogonal pero no sería necesario que lo fuese), y formamos la matriz

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \\ 1 & 0 \end{pmatrix}$$

Con ella calculamos $P_S = A\,(A^t A)^{-1} A^t = \begin{pmatrix} 2 & 0 \\ 0 & 3 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 9 \end{pmatrix}^{-1} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \end{pmatrix} = \begin{pmatrix} \frac{4}{5} & 0 & \frac{2}{5} \\ 0 & 1 & 0 \\ \frac{2}{5} & 0 & \frac{1}{5} \end{pmatrix}$

Ahora proyectamos el vector v = (3,2,2) :

$\text{proy}_S(v) = P_S \cdot v = \begin{pmatrix} \frac{4}{5} & 0 & \frac{2}{5} \\ 0 & 1 & 0 \\ \frac{2}{5} & 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{16}{5} \\ 2 \\ \frac{8}{5} \end{pmatrix}$ que coincide con el resultado ya obtenido.

Propiedades de la matriz de proyección.

Toda matriz P de proyección sobre un subespacio de $\mathbb{R}^n$ es:

1. **Cuadrada nxn**
2. **Simétrica**
3. **Idempotente (P^2=P)**

El hecho de que matriz sea idempotente ($P^2 = P$) significa que efectuar la proyección dos veces da el mismo resultado que una sola vez.

Por otra parte, toda matriz Q que cumpla las tres propiedades anteriores, resulta ser matriz de proyección de un cierto subespacio de $\mathbb{R}^n$. Este subespacio es aquel que generan sus columnas.

Ejemplo. Comprobemos que la matriz $Q=\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$ cumple las anteriores propiedades.

Por tanto, Q es matriz de proyección de un cierto subespacio S. Éste será el generado por sus columnas $(\frac{1}{2},\frac{1}{2},0)$, $(\frac{1}{2},\frac{1}{2},0)$, (0,0,1), es decir, el generado por (1,1,0), (0,0,1).

Observación.

Si tenemos que proyectar un vector sobre un subespacio S, y no disponemos de una base ortogonal de S, tenemos dos opciones:

- Calcular una base ortogonal por Gram-Schmidt para utilizar la fórmula de la proyección,

- O bien hallar la matriz de proyección P_S y utilizarla para proyectar el vector.

OTRAS NOCIONES GEOMÉTRICAS.

1. Vector simétrico.

En un espacio euclídeo podemos generalizar la noción de vector simétrico respecto a un plano, recta o subespacio en general, que actuará como un “espejo”.

Dado un vector v y un subespacio S, descomponemos:

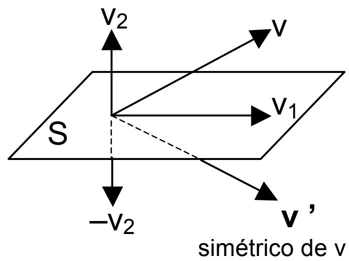
$$v = v_1 + v_2 = \text{proy}_S(v) + \text{proy}_{S^\perp}(v)$$

Entonces se define el **vector simétrico de v respecto a S** como

$$\mathbf{v'} = v_1 - v_2$$

es decir

$$\boxed{\mathbf{v'} = \text{proy}_S(v) - \text{proy}_{S^\perp}(v)}$$



2. Cálculo de áreas y volúmenes.

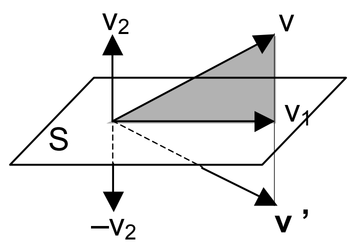
Puesto que en un espacio euclídeo sabemos “medir” distancias y longitudes (con la noción de módulo de un vector), ello nos permite aplicar las fórmulas geométricas usuales para calcular áreas y volúmenes en $\mathbb{R}^2$ y $\mathbb{R}^3$.

Ejemplo.

Dado el subespacio S generado por (2,0,1),(0,3,0) hallar el simétrico **v’** del vector **v** = (3,2,2). Calcular el área del triángulo formado por **v** y **v’**.

- Como ya hemos calculado, $v = v_1 + v_2 = \left(\frac{16}{5}, 2, \frac{8}{5}\right) + \left(\frac{-1}{5}, 0, \frac{2}{5}\right)$.

Entonces, $v' = v_1 - v_2 = \left(\frac{16}{5}, 2, \frac{8}{5}\right) - \left(\frac{-1}{5}, 0, \frac{2}{5}\right) = \left(\frac{17}{5}, 2, \frac{6}{5}\right)$



El área pedida es dos veces el área sombreada, por tanto

$$\text{área} = 2\,\frac{\text{base}\cdot\text{altura}}{2} = 2\,\frac{|v_1|\cdot|v_2|}{2} = |v_1| \cdot |v_2| = \sqrt{\left(\tfrac{16}{5}\right)^2 + 2^2 + \left(\tfrac{8}{5}\right)^2} \cdot \sqrt{\left(\tfrac{-1}{5}\right)^2 + 0 + \left(\tfrac{2}{5}\right)^2} \approx 1{,}833 \text{ unidades}^2$$

APLICACIONES PRÁCTICAS.

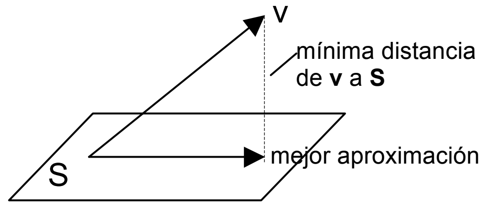
Veremos cómo aplicar las nociones del espacio euclídeo para:

- Aproximar una función continua por medio de un polinomio.
- Encontrar la solución aproximada de un sistema incompatible.
- Ajustar a una gráfica (recta, curva) una nube de puntos.

Para ello introducimos la noción de "mejor aproximación"

Mejor aproximación de un vector en un subespacio.

En un espacio euclídeo, dado un vector **v** y un subespacio **S**, de entre todos los vectores de **S** hay uno que es el más próximo a **v.** Se llama **mejor aproximación de v en S**, y es precisamente la proyección ortogonal **proy$_S$(v).**



Para cualquier otro $w \in S$ la distancia a v es mayor, es decir

$|\ proy_S(v) - v\ | < |\ w - v\ |$ para cualquier $w \in S$, $w \neq proy_S(v)$

Se llama **error cuadrático medio** (o **error cuadrático**, o **desviación cuadrática**) al cuadrado de la distancia que separa v de su aproximación:

$$\boxed{\textbf{Error} = |\ \textbf{proy}_S(\textbf{v}) - \textbf{v}\ |^2}$$

Nota. También es posible usar como medida de error la norma $|\ proy_S(v) - v\ |$ en lugar de la norma al cuadrado.

1. Aproximación de una función continua por un polinomio.

Para facilitar los cálculos con funciones trigonométricas, logarítmicas, exponenciales..., podemos sustituirlas por un polinomio que se aproxime a ellas en un cierto intervalo [a, b].

Para ello consideramos el espacio euclídeo

C[a,b] = { funciones continuas en el intervalo [a,b] }.

con el producto escalar dado por: $\quad f \cdot g = \int_a^b f(x)g(x)\,dx$

y consideraremos el subespacio **P**$_r$ = {polinomios de grado $\leq$ r }, calculando entonces la mejor aproximación del vector (la función dada) en dicho subespacio.

Podemos elegir el grado **r** dependiendo de la precisión que queramos alcanzar. Cuanto mayor sea el grado, menor será el error.

Ejemplo.

Queremos aproximar la función e^x por un polinomio de grado 2, en el intervalo [0, 1].

Para ello consideramos el espacio ***C***[0,1] con el producto escalar $f \cdot g = \int_a^b f(x)g(x)\,dx$

el vector (función) $\mathbf{f} = \mathbf{e^x}$, y el subespacio $\mathbf{P}_2$ = {polinomios de grado $\leq$ 2 },

Calculemos la mejor aproximación de **f** en $\mathbf{P}_2$, es decir, $\text{proy}_{\mathbf{P}_2}(f)$.

Para ello necesitamos una base ortogonal de $\mathbf{P}_2$. Partiendo de la base estándar $\{1, x, x^2\}$ podemos hallar una ortogonal por Gram-Schmidt (utilizando siempre el producto escalar dado por la integral), obteniéndose la base $a_1 = 1$, $a_2 = x - \frac{1}{2}$, $a_3 = x^2 - x + \frac{1}{6}$. Así,

$$\text{proy}_{\mathbf{P}_2}(f) = \frac{a_1 \cdot f}{a_1 \cdot a_1} a_1 + \frac{a_2 \cdot f}{a_2 \cdot a_2} a_2 + \frac{a_3 \cdot f}{a_3 \cdot a_3} a_3$$

lo que se calcula mediante las integrales correspondientes

(p. ej. $a_1 \cdot f = \int_0^1 1 \cdot e^x \, dx = e - 1$, etc.)

resultando el polinomio

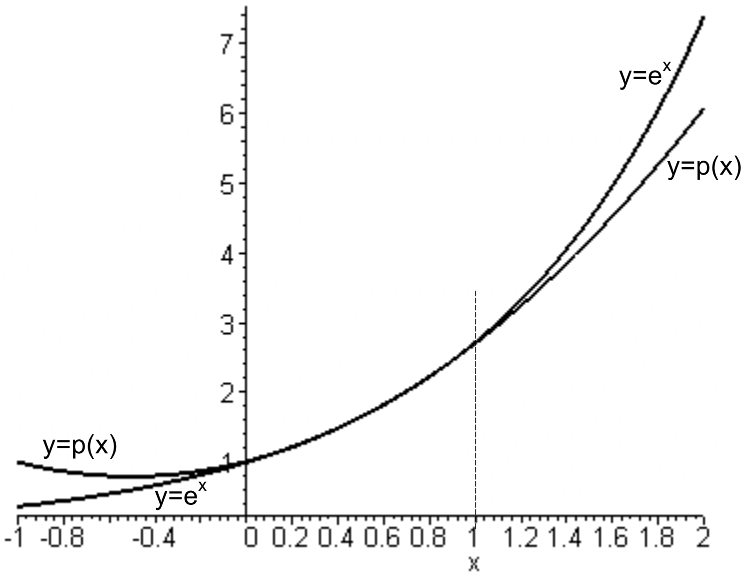
$p(x) \approx 0'84\, x^2 + 0'85\, x + 1'01$

El error cuadrático es

$|\, p - f \,|^2 = \int_0^1 \left[(p(x) - f(x) \right]^2 dx =$

$= 0'0000386$

El error es muy pequeño, lo que significa que la aproximación es buena. En efecto, en la figura se ve que la gráfica del polinomio p(x) y la de e^x están "bastante próximas" en el intervalo [0,1].



(De hecho, el error representa el área comprendida entre las dos curvas. Este método de aproximación hace que dicha área sea mínima).

Si lo aproximáramos por un polinomio de grado 3, 4,... obtendríamos una aproximación aún mejor.

Otras posibilidades de aproximación.

- Dada una función f(x), en lugar de aproximarla por un polinomio, también podemos aproximarla por otro tipo de funciones. Basta hacer la transformación adecuada para reducirlo al problema de aproximación por polinomios. Por ejemplo,

1) Aproximar f(x) = tg x por una función de la forma $\frac{1}{ax+b}$:

Equivale a aproximar $\frac{1}{\text{tg } x}$ por un polinomio $ax + b$.

2) Aproximar f(x) = log x por una función de la forma $\sqrt{ax+b}$:

Equivale a aproximar $(\log x)^2$ por un polinomio $ax + b$.

3) Aproximar f(x) = cos x por una función de la forma e^{ax+b} :
Equivale a aproximar $\log(\cos x)$ por un polinomio $ax + b$.

- En otras ocasiones será necesario un cambio de variable:

4) Aproximar f(x) = cos(x) por una función de la forma $\frac{a}{x} + b$:

Hacemos el cambio $\frac{1}{x} = t$ y así equivale a aproximar $\cos\left(\frac{1}{t}\right)$ por un polinomio $at + b$.

5) Aproximar f(x) = sen(x) por una función de la forma $a \log(x) + b$:
Hacemos el cambio $\log(x) = t$ y por tanto $x = e^t$, y así equivale a aproximar $\text{sen}(e^t)$ por un polinomio $at + b$.

6) Aproximar f(x) = cos(x) por una función de la forma $ae^x + b$:
Hacemos el cambio $e^x = t$ y por tanto $x = \log(t)$, y así equivale a aproximar $\cos(\log(t))$ por un polinomio $at + b$.

- También es posible tomar logaritmos:

7) Aproximar f(x) = cos(x) por una función de la forma ae^{bx} :
Equivale a aproximar $\log(\cos(x))$ por una función de la forma $\log(a e^{bx})$, es decir, $\log(a) + bx$.
y poniendo $c = \log(a)$, equivale a aproximar $\log(\cos(x))$ por una función $c + bx$ que ya es un polinomio.
Una vez hallado c, recuperamos $a = e^c$, pues los coeficientes pedidos eran a y b.

2. Solución aproximada de sistemas incompatibles.

Observemos que resolver un sistema lineal $A\vec{x}=\vec{b}$ consiste en poner la columna $\vec{b}$ de términos independientes como combinación lineal de las columnas de la matriz:

Por ejemplo, $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ es encontrar x, y que cumplan $x\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}+y\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$

Existirán tales x, y si el vector $\vec{b}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ es combinación lineal de $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$, $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$, es decir, si pertenece al subespacio S generado por ellas. En ese caso el sistema será compatible.

Si es incompatible, se debe a que $\vec{b}$ no pertenece al subespacio S.

En ese caso, podemos sustituir $\vec{b}$ por otro vector $\vec{c}$ que sí se encuentre en S, y para cometer el menor error posible, $\vec{c}$ será la mejor aproximación de $\vec{b}$ en S.

Por tanto, en lugar del sistema incompatible $A\vec{x}=\vec{b}$, resolvemos otro sistema:

$A\vec{x}=\vec{c}$, donde $\vec{c}=\text{proy}_S(\vec{b})$, S = subespacio generado por las columnas de A.

Este nuevo sistema ya es compatible. Su solución, si bien no cumple las ecuaciones del sistema original $A\vec{x}=\vec{b}$, las satisface "aproximadamente".

El error cuadrático es $|\vec{c}-\vec{b}|^2$ (trabajando en $\mathbb{R}^n$ con el producto escalar usual), o dicho de otra manera, $|A\vec{x}-\vec{b}|^2$ donde $\vec{x}$ es la solución aproximada que hemos hallado.

Observar que si $\vec{x}$ fuera una solución exacta del sistema, el error $|A\vec{x}-\vec{b}|^2$ sería cero.

Este método se llama **método de mínimos cuadrados** ya que hace mínimo el error cuadrático.

Ejemplo 1.

$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ Este sistema es incompatible, lo cual es debido a que b = $\begin{pmatrix} 3 \\ 5 \end{pmatrix}$ no pertenece al subespacio S generado por $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ en $\mathbb{R}^2$. Entonces sustituimos b por su mejor aproximación en S. Para ello, una base de S es el vector (1,2).

$$c = \text{proy}_S(b) = \text{proy}_{(1,2)}(b) = \frac{(1,2)\begin{pmatrix} 3 \\ 5 \end{pmatrix}}{(1,2)\begin{pmatrix} 1 \\ 2 \end{pmatrix}}\ (1,2) = \frac{13}{5}(1,2) = \left(\frac{13}{5},\frac{26}{5}\right)$$

y ahora resolvemos el sistema compatible (en este caso indeterminado)

$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{13}{5} \\ \frac{26}{5} \end{pmatrix}$ cuya solución es $(\frac{13}{5} - 2\lambda,\ \lambda)$. Obtenemos infinitas soluciones que satisfacen aproximadamente el sistema original.

El error cuadrático es $|\, b - c\,|^2 = \left|(3,5) - \left(\frac{13}{5},\frac{26}{5}\right)\right|^2 = \left|\left(\frac{2}{5},\frac{-1}{5}\right)\right|^2 = \frac{1}{5} = 0{,}2$.

Ejemplo 2.

$\begin{pmatrix} 2 & 3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ Igualmente es incompatible puesto que $\vec{b} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ no pertenece al subespacio S generado por las columnas $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$ en $\mathbb{R}^3$.

Calculamos por tanto la mejor aproximación de $\vec{b}$ en S. Una base de S está formada por ambas columnas, puesto que son linealmente independientes. Con ella podemos calcular la matriz de proyección, que será $P_S = A\,(A^t A)^{-1} A^t$.

Así, $\vec{c} = \text{proy}_S(\vec{b}) = P_S \cdot \vec{b} = \begin{pmatrix} \frac{5}{14} \\ \frac{2}{7} \\ \frac{-1}{14} \end{pmatrix}$

y se resuelve el sistema $\begin{pmatrix} 2 & 3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \frac{5}{14} \\ \frac{2}{7} \\ \frac{-1}{14} \end{pmatrix}$ compatible determinado, cuya solución es

$$\begin{cases} x = \dfrac{2}{7} \\ y = \dfrac{-1}{14} \end{cases}$$

El error cuadrático es $|\ \vec{b} - \vec{c}\ |^2 = \left| (0,1,1) - \left(\frac{5}{14}, \frac{2}{7}, \frac{-1}{14} \right) \right|^2 = \frac{25}{14} \approx 1{,}79$.

Obtención directa de la solución.

En algunos casos podremos obtener la solución directamente, sin siquiera plantear ni resolver el sistema compatible.

- Dado el sistema incompatible $A\vec{x} = \vec{b}$, comprobar si las columnas de A son linealmente independientes (ello ocurrirá cuando el rango de A coincida con el nº de columnas).
- Si es así, la solución aproximada por mínimos cuadrados es única, y es

$$\boxed{\vec{x} = (A^t A)^{-1} A^t \vec{b}}$$

Explicación:

En el ejemplo 2 anterior, las columnas de A son linealmente independientes. Ello nos ha permitido utilizar A para para calcular la matriz de proyección $A(A^tA)^{-1} A^t$, y ésta para calcular $\vec{c} = A (A^t A)^{-1} A^t \vec{b}$.

Pero, observemos la expresión $A \underbrace{(A^t A)^{-1} A^t \vec{b}} = \vec{c}$

Como el sistema que hay que resolver es $A \vec{x} = \vec{c}$, la expresión anterior nos indica que

$(A^t A)^{-1} A^t \vec{b}$ (señalado arriba con la llave), que es un vector columna, es solución del sistema $A \vec{x} = \vec{c}$.

Así pues, esa es la solución (única, pues el sistema $A \vec{x} = \vec{c}$ es compatible determinado).

Caso particular: Ajuste de nubes de puntos.

Dada una nube de puntos, (es decir, un conjunto de puntos en el plano, obtenidos mediante resultados experimentales, estadísticos, etc) buscamos la recta, parábola, etc. que mejor se ajuste a dichos puntos.

Si planteamos que la curva pase por todos los puntos, obtendremos un sistema incompatible que podemos resolver mediante el método anterior.

Ejemplo

Buscamos la recta que mejor se ajuste a los puntos (1,2), (2,3), (3,5).

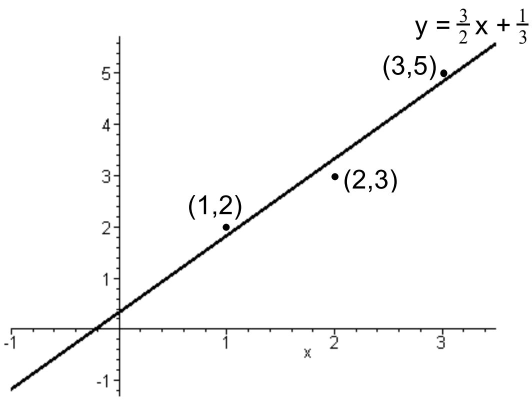
Una recta es de la forma y = m x + b. Si pasara por los tres puntos, debería cumplirse

$$\left.\begin{array}{l} 2 = a\cdot 1 + b \\ 3 = a\cdot 2 + b \\ 5 = a\cdot 3 + b \end{array}\right\} \quad \text{es decir, un sistema en las incógnitas } m, b: \quad \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

Como las columnas A (matriz de coeficientes) son linealmente independientes, la solución se puede hallar directamente como

$$(A^t A)^{-1} A^t \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{4}{3} & \frac{1}{3} & -\frac{2}{3} \end{pmatrix}\begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ \frac{1}{3} \end{pmatrix}$$

es decir $m = \dfrac{3}{2}$, $b = \dfrac{1}{3}$



y así la recta buscada es $y = \dfrac{3}{2}x + \dfrac{1}{3}$. Comprobar en la figura que, si bien esta recta no pasa por ninguno de los puntos dados, se aproxima a todos ellos.

El error cuadrático es $|A\vec{x} - \vec{b}|^2 = \left\| \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}\begin{pmatrix} \frac{3}{2} \\ \frac{1}{3} \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} \right\|^2 = \left\| \begin{pmatrix} \frac{11}{6} \\ \frac{10}{3} \\ \frac{29}{6} \end{pmatrix} - \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix} \right\| = \dfrac{1}{6} = 0{,}167$

El error es pequeño porque los puntos dados estaban “casi” alineados.

- **Otras posibilidades**: ajustar la nube de puntos mediante una parábola $\mathbf{y = ax^2 + bx + c}$, o mediante polinomios de grado mayor; o también mediante una función de cualquier tipo como por ejemplo $\mathbf{y = a \cos(x) + b \log(x)}$, etc.

En todos los casos se debe "hacer pasar" la función por cada uno de los puntos, introduciendo la abscisa del punto en la **x** y la ordenada en la **y**, para obtener una ecuación por cada punto. Finalmente obtendremos un sistema incompatible en las incógnitas **a, b, c...** que resolveremos por mínimos cuadrados.

Observación.

Puede ocurrir que el método de mínimos cuadrados se aplique a un sistema que ya era compatible. Es lo que ocurre si, por ejemplo, tratamos de ajustar por una recta una nube de puntos que ya estaban alineados.

En este caso no hay ningún impedimento para aplicar el método, y la solución "aproximada" es en realidad la solución "exacta", la misma que se obtendría resolviendo el sistema compatible por cualquiera de los métodos conocidos.

En este caso, el error cuadrático sería cero.