Práctica 6
Ortogonalidad y proyecciones ortogonales

Índice

1. Producto escalar, norma y distancia

Recordaremos en este apartado diversos conceptos presentados en la teoría. Sean $\vec{u}$ y $\vec{v}$ dos vectores de $\mathbb{R}^n$.

- El *producto escalar* (o producto interior) de $\vec{u}$ y $\vec{v}$ es el número real

$$\vec{u}\cdot\vec{v} = \vec{u}^t\vec{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n.$$

- Si $\vec{u}\cdot\vec{v} = 0$, se dice que los vectores $\vec{u}$ y $\vec{v}$ son *ortogonales*.

- Si W es un subespacio vectorial de $\mathbb{R}^n$, se denomina *complemento ortogonal* de W al subespacio
$$W^\perp = \{\vec{z}\in\mathbb{R}^n : \vec{z}\cdot\vec{w} = 0 \quad \text{para todo } \vec{w}\in W\}.$$

- La *norma* (o longitud) de $\vec{v}$ es

$$\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

  i $\|\vec{v}\|^2 = \vec{v}\cdot\vec{v}$.

- Un vector se dice que es *unitario* si su norma es 1. El vector $(1/\|\vec{v}\|)\vec{v}$ siempre es unitario.

- La *distancia* entre $\vec{u}$ y $\vec{v}$ es $\|\vec{u}-\vec{v}\|$.

- Dada una matriz real A con m filas y n columnas, se denomina *subespacio columna* de A (y se denota por Col A) al subespacio vectorial de $\mathbb{R}^m$ generado por los vectores columna de A.
- Dada una matriz real A con m files y n columnas, se denomina *subespacio fila* de A (y se denota por Fil A) al subespacio vectorial de $\mathbb{R}^n$ generado por los traspuestos de los vectores fila de A.

Teorema 1. Sea A una matriz real con m filas y n columnas. Se satisface lo siguiente:

$$(\operatorname{Fil} \mathsf{A})^\perp = \operatorname{Nul} A \quad \text{i} \quad (\operatorname{Col} \mathsf{A})^\perp = \operatorname{Nul} A^t.$$

Ejemplo 1. Sean

$$\vec{u} = \begin{bmatrix} 5 \\ -4 \\ 0 \\ -3 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} -4 \\ 1 \\ -3 \\ 8 \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} 3 \\ 3 \\ 5 \\ -1 \end{bmatrix},$$

y consideremos $W = \langle \vec{u}, \vec{v}, \vec{w} \rangle$ y

$$\mathsf{A} = \begin{bmatrix} 5 & -4 & 3 \\ -4 & 1 & 3 \\ 0 & -3 & 5 \\ 3 & 8 & -1 \end{bmatrix}.$$

- Los vectores $\vec{u}$ y $\vec{v}$ son ortogonales porque $\vec{u} \cdot \vec{v} = 5 \cdot (-4) + (-4) \cdot 1 + 0 \cdot (-3) + 3 \cdot 8 = 0$.
- $\|\vec{w}\| = \sqrt{3^2 + 3^2 + 5^2 + (-1)^2} = 2\sqrt{11}$.
- El vector $\frac{1}{\|\vec{w}\|}\vec{w} = \begin{bmatrix} 3/2\sqrt{11} \\ 3/2\sqrt{11} \\ 5/2\sqrt{11} \\ -1/2\sqrt{11} \end{bmatrix}$ es unitario.
- $\|\vec{u} - \vec{v}\| = \sqrt{(5-4)^2 + (-4-1)^2 + (0-(-3))^2 + (3-8)^2} = 2\sqrt{15}$.
- A es equivalente por filas a

$$\begin{bmatrix} 5 & -4 & 3 \\ 0 & 7 & 9 \\ 0 & 0 & 62 \\ 0 & 0 & 0 \end{bmatrix},$$

  por tanto

$$(\operatorname{Fil} A)^\perp = \left(\left\langle \begin{bmatrix} 5 \\ -4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 7 \\ 9 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 62 \end{bmatrix} \right\rangle \right)^\perp = \operatorname{Nul} \mathsf{A} \quad \text{i} \quad (\operatorname{Col} A)^\perp = W^\perp = \operatorname{Nul} \mathsf{A}^t$$

Con Scilab:

- Introducimos los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$:

```
-->u=[5;-4;0;3], v=[-4;1;-3;8], w=[3;3;5;-1]
 u  =

    5.
  - 4.
    0.
    3.
 v  =

  - 4.
    1.
  - 3.
    8.
 w  =

    3.
    3.
    5.
  - 1.
```

- Para comprobar que los vectores $\vec{u}$ y $\vec{v}$ son ortogonales, calculamos su producto escalar:

```
-->uv=u'*v
 uv  =

    0.
```

- Para determinar la norma del vector $\vec{w}$, usaremos el comando **norm()**:

```
-->norm(w)
 ans  =

    6.6332496
```

- Para determinar un vector unitario asociado a $\vec{w}$ calculamos

```
-->t=w/norm(w)
 t  =

    0.4522670
    0.4522670
```

```
    0.7537784
  - 0.1507557
```

- Comprobamos que $\vec{t}$ es unitario:

```
-->norm(t)
 ans  =

    1.
```

- Para determinar la distancia entre los vectores $\vec{u}$ y $\vec{v}$, hacemos lo siguiente:

```
-->norm(u-v)
 ans  =

    11.83216
```

- Construimos la matriz A que tiene como columnas a los vectores $\vec{u}$, $\vec{v}$ y $\vec{w}$:

```
-->A=[u v w]
 A  =

    5.  - 4.    3.
  - 4.    1.    3.
    0.  - 3.    5.
    3.    8.  - 1.
```

- Para calcular una base del subespacio $(\mathrm{Fil}\,\mathrm{A})^{\perp}$, calcularemos el núcleo de la matriz A:

```
-->NF=kernel(A)
 NF  =

     []
```

Es decir: el vector nulo es el único vector ortogonal al subespacio fila $\mathrm{Fil}\,\mathrm{A}$.

- De forma análoga, para determinar una base del subespacio $(\mathrm{Col}\,\mathrm{A})^{\perp}$ tendremos que calcular una base del núcleo de A^t:

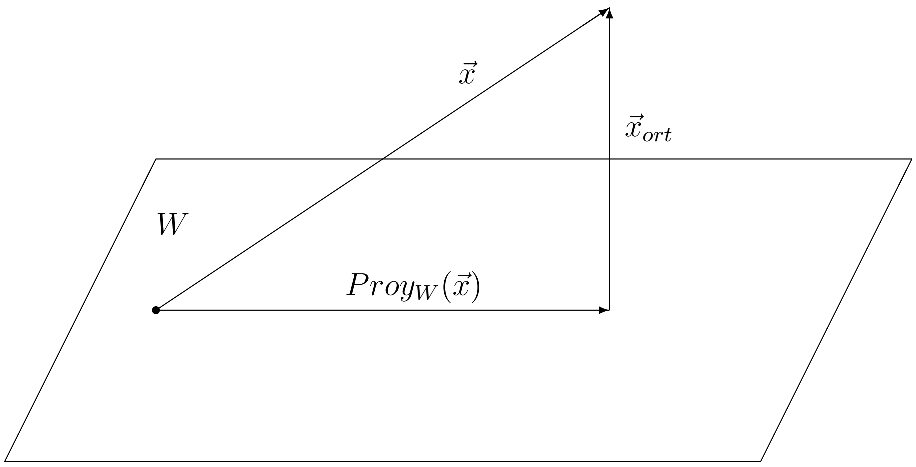
```
-->NC=kernel(A')
 NC  =

    0.4958960
    0.5689754
  - 0.6524947
  - 0.0678594
```

2. Proyecciones ortogonales

2.1. Definición

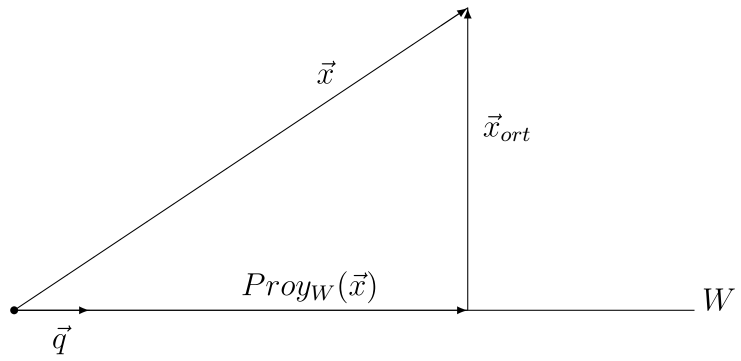
Sea W un subespacio vectorial de $\mathbb{R}^n$. Para cada vector $\vec{x} \in \mathbb{R}^n$ existen un *único* vector $\vec{w} \in W$ y un único vector $\vec{x}_{ort}$ en el complemento ortogonal $W^\perp$ tales que $\vec{x} = \vec{w} + \vec{x}_{ort}$. El vector $\vec{w}$ se denomina *proyección ortogonal* de $\vec{x}$ sobre el subespacio W, y lo denotaremos por $Proy_W(\vec{x})$. Esta definición formaliza y generaliza la idea greométrica de "proyección perpendicular":



El vector $\vec{x}$ se puede descomponer, de forma única, como una suma de dos *componentes*: una de ellas ($Proy_W(\vec{x})$) sobre W y la otra ($\vec{x}_{ort}$) es ortogonal a todos los vectores de W (es decir, pertenece a $W^\perp$). Cuando hablemos de proyección ortogonal de $\vec{x}$ sobre W estaremos hablando de la componente $Proy_W(\vec{x})$ sobre W en esta descomposición.

2.2. Caso fácil: proyección ortogonal sobre una recta

Consideremos el caso en el cual W es una recta de $\mathbb{R}^n$, es decir, está generado por un único vector no nulo. Podemos dividir este generador entre su norma y transformalo en un vector unitario (que continuará generando la recta). Tomamos $S = \{\vec{q}\}$, donde $\vec{q}$ es un vector <u>unitario</u> que genera la recta.



Consideremos un vector no nulo cualquiera $\vec{x} \in \mathbb{R}^n$. Queremos calcular la proyección ortogonal $Proy_W(\vec{x})$ de $\vec{x}$ sobre W. Como $Proy_W(\vec{x})$ pertenece a la recta W, existe un escalar λ tal que $Proy_W(\vec{x}) = \lambda\vec{q}$. Pero el vector $\vec{x} - \lambda\vec{q} = \vec{x}_{ort}$ es ortogonal a W y, por tanto, $(\vec{x} - \lambda\vec{q}) \cdot \vec{q} = 0$. Es decir, $\vec{x} \cdot \vec{q} - \lambda\vec{q} \cdot \vec{q} = 0$; como $\vec{q}$ es unitario tenemos que $\lambda = \vec{q} \cdot \vec{x}$ o, utilizando la notación de producto fila-columna en lugar de la notación del producto escalar, $\lambda = \vec{q}^{\,t}\vec{x}$. Concluimos, por tanto, la siguiente propiedad:

La proyección ortogonal de un vector $\vec{x}$ sobre una recta W es el vector

$$Proy_W(\vec{x}) = (\vec{q}^{\,t}\vec{x})\vec{q}$$

donde $\vec{q}$ es un generador <u>unitario</u> de la recta.

Ejemplo 2. Consideremos la recta $W = \langle(1, -2, 5)\rangle \subseteq \mathbb{R}^3$. Calcularemos a continuación la proyección ortogonal del vector $\vec{x} = (0, 1, 1)$ sobre W. Primero calcularemos un generador unitario de la recta, dividiendo este vector entre su norma:

```
-->u=[1; -2; 5];

-->q=u/norm(u)
 q  =

    0.1825742
  - 0.3651484
    0.9128709
```

Ahora calcularemos la proyección usando la fórmula de antes:

```
-->x=[0; 1; 1];

-->(q'*x)*q
 ans  =
```

```
  0.1
- 0.2
  0.5
```

Por tanto, la proyección ortogonal es $(0{,}1, -0{,}2, 0{,}5)$.

2.3. Caso general

Queremos ahora calcular la proyección ortogonal de un vector sobre un subespacio vectorial cualquiera W. Suponemos conocido un sistema generador $S = \{\vec{u}_1, \ldots, \vec{u}_r\}$ de W. Entonces W puede verse como el subespacio columna de la matriz $\mathsf{M}(S)$ (la matriz que tiene, como columnas, los vectores $\vec{u}_i$). Por tanto

$$W^{\perp} = \mathrm{Nul}(\mathsf{M}(S)^t).$$

Como $\vec{x} = Proy_W(\vec{x}) + \vec{x}_{ort}$ se tiene que $\vec{x} - Proy_W(\vec{x}) = \vec{x}_{ort}$ y, por tanto,

$$\boxed{\vec{x} - Proy_W(\vec{x}) \text{ ha de ser ortogonal a } W,}$$

es decir,

$$\vec{x} - Proy_W(\vec{x}) \in W^{\perp} = \mathrm{Nul}(\mathsf{M}(S)^t).$$

Así, tenemos que

$$\mathsf{M}(S)^t(\vec{x} - Proy_W(\vec{x})) = \vec{0},$$

es decir,

$$\mathsf{M}(S)^t Proy_W(\vec{x}) = \mathsf{M}(S)^t\vec{x}. \tag{1}$$

Por otro lado

$$\boxed{\text{el vector } Proy_W(\vec{x}) \text{ pertenece a } W}$$

y esto quiere decir que puede escribirse como combinación lineal de los vectores de S (que es un sistema generador de W). Denotamos por y_i a los coeficientes de esta combinación lineal:

$$Proy_W(\vec{x}) = y_1\vec{u}_1 + \cdots + y_r\vec{u}_r \tag{2}$$

y definimos el vector

$$\vec{y} := \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix}.$$

(Observa que calcular $Proy_W(\vec{x})$ equivale a calcular $\vec{y}$).

La igualdad (2) significa que

$$Proy_W(\vec{x}) = \mathsf{M}(S)\vec{y}.$$

Sustituyendo esta nueva expresión de $Proy_W(\vec{x})$ en la igualdad (1) se tiene que

$$M(S)^t\mathsf{M}(S)\vec{y} = \mathsf{M}(S)^t\vec{x}.$$

Observa que $M(S)^t\mathsf{M}(S)$ es una matriz cuadrada de orden r y que la igualdad anterior es la expresión matricial de un sistema de r ecuaciones lineales con incógnitas $y_1, \ldots, y_r$ y con vector de términos independientes $M(S)^t\vec{x}$ (recuerda que $\vec{x}$ es el vector que estamos proyectando). Puede razonarse fácilmente que todas las soluciones $(y_1, \ldots, y_r)$ de este sistema dan lugar a la misma combinación lineal $y_1\vec{u}_1 + \cdots + y_r\vec{u}_r$, que es la proyección ortogonal deseada.

Hemos deducido, por tanto, el siguiente resultado:

> La proyección ortogonal de un vector $\vec{x}$ sobre un subespacio vectorial W generado por un conjunto de vectores $S = \{\vec{u}_1, \ldots, \vec{u}_r\}$ es el vector $Proy_W(\vec{x}) = y_1\vec{u}_1 + \cdots + y_r\vec{u}_r$, donde $\vec{y} = (y_1, \ldots, y_r)$ es una solución del sistema
>
> $$M(S)^t\mathsf{M}(S)\vec{y} = \mathsf{M}(S)^t\vec{x}. \tag{3}$$

Supongamos ahora que S és linealmente independiente, es decir, que es una base de W.

En este caso, $\mathrm{rang}(\mathsf{M}(S)^t\mathsf{M}(S)) = r$ y esto implica que el sistema (3) **tiene solución única** que nos da los coeficientes $\vec{y}_i$ (respecto de la base S) de la proyección ortogonal de $\vec{x}$. Como la matriz $\mathsf{M}(S)^t\mathsf{M}(S)$ es invertible se tiene que la igualdad (3) equivale a

$$\vec{y} = (\mathsf{M}(S)^t\mathsf{M}(S))^{-1}\mathsf{M}(S)^t\vec{x}.$$

Como las componentes de $\vec{y}$ son las coordenadas respecto de la base S de la proyección de $\vec{x}$, el producto $\mathsf{M}(S)\vec{y}$ será igual a $Proy_W(\vec{x})$:

$$Proy_W(\vec{x}) = \mathsf{M}(S)(\mathsf{M}(S)^t\mathsf{M}(S))^{-1}\mathsf{M}(S)^t\vec{x}.$$

Así, cuando el sistema de generadores S es una base de W, hemos obtenido una formula elegante para proyectar sobre W cualquier vector:

> La proyección ortogonal de un vector $\vec{x}$ sobre un subespacio vectorial W generado por <u>una base</u> S de W es
>
> $$Proy_W(\vec{x}) = \mathsf{P}_W\vec{x}, \tag{4}$$
>
> donde
>
> $$\mathsf{P}_W = \mathsf{M}(S)(\mathsf{M}(S)^t\mathsf{M}(S))^{-1}\mathsf{M}(S)^t$$
>
> se denomina *matriz de proyección* sobre W.

(La matriz de proyección P_W es independiente de la base S que consideremos).

Ejemplo 3. Sea W el subespacio vectorial de $\mathbb{R}^3$ con base $S = \{(1,2,3), (-3,5,1)\}$ (observa que S es, efectivamente, una base!) y consideremos el vector $\vec{x} = (2,3,4) \in \mathbb{R}^3$. Calcularemos, con la ayuda de Scilab, la proyección ortogonal de $\vec{x}$ sobre W.

Primero definiremos la matriz $\mathsf{M}(S)$:

```
-->u1=[1; 2; 3;]; u2=[-3; -5; 1]; MS=[u1 u2]
 MS  =

    1.  - 3.
    2.  - 5.
    3.    1.
```

Calculemos la matriz de proyección P_W:

```
-->x=[2; 3; 4];
-->PW=MS*inv(MS'*MS)*MS'
 PW  =

    0.2589744    0.4358974  - 0.0435897
    0.4358974    0.7435897    0.0256410
  - 0.0435897    0.0256410    0.9974359
```

Multiplicando esta matriz por el vector $\vec{x}$ obtendremos la proyección ortogonal:

```
-->PW*x
 ans  =

    1.6512821
    3.2051282
    3.9794872
```

Por tanto $Proy_W(\vec{x}) = (1{,}6512821, 3{,}2051282, 3{,}9794872)$.