Unit 5: Orthogonality and least square problems.

María Barbero Liñán

Universidad Carlos III de Madrid
Bachelor in Statistics and Business

Mathematical methods II

2011-2012

Let $u, v \in \mathbb{R}^n$. The inner product of u and v is defined by

$$u \cdot v = u^T v = \begin{pmatrix} u_1, u_2, \dots, u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Properties: **For $u, v, w \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$,**

- $u \cdot v = v \cdot u$.
- $(u + v) \cdot w = u \cdot w + v \cdot w$.
- $(\lambda u) \cdot v = \lambda(u \cdot v) = u \cdot (\lambda v)$.
- $u \cdot u \geq 0$, $u \cdot u = 0$ if and only if $u = 0_n$.

The norm or length of $v \in \mathbb{R}^n$ is a nonnegative real number given by

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

A vector v is unit if $\|v\| = 1$.

If v is not unit, it can be normalized by computing $\dfrac{v}{\|v\|}$.

Properties:

- $u \cdot v = \|u\|\|v\| \cos \theta$.
- $\|\lambda v\| = |\lambda|\|v\|$.

The distance between two vectors u and v is...

the length of the vector $\|u - v\| = \mathrm{dist}(u, v)$.

Two vectors are orthogonal, $u \perp v$, if and only if...

$$u \cdot v = 0.$$

A set of vectors $\{v_1, \ldots, v_p\}$ in $\mathbb{R}^n$ is ORTHOGONAL

if each pair of different vectors in the set is orthogonal, that is,

$$v_i \cdot v_j = 0 \text{ for all } i \neq j.$$

The orthogonal subspace of a subspace H in $\mathbb{R}^n$.

$$H^\perp = \{w \in \mathbb{R}^n \mid w \cdot v = 0 \text{ for all } v \in H\}.$$

- $H^\perp$ is a subspace in $\mathbb{R}^n$.
- w is orthogonal to H if and only if w is orthogonal to each vector of the spanning set of H.

A set of vectors $\{v_1, \ldots, v_p\}$ in $\mathbb{R}^n$ is ORTHONORMAL

if the set of vectors is orthogonal and all the vectors are unit, that is,

$$v_i \cdot v_j = 0 \text{ for all } i \neq j, \quad v_i \cdot v_i = 1 \text{ for all } i = 1, \ldots, p.$$

- An orthogonal basis is a basis whose vectors are an orthogonal set of vectors.
- An orthonormal basis is a basis whose vectors are an orthonormal set of vectors.

If a SQUARE matrix $A \in M_{n \times n}$ is ORTHOGONAL,

the rows of A are an orthonormal basis for $\mathbb{R}^n$. $AA^T = A^T A = \mathrm{I}_n$.

The orthogonal projection of v onto a subspace W spanned by a set of linearly independent vectors $a_1, \ldots, a_p$,

$$\mathrm{proj}_W v = A(A^T A)^{-1} A^T v, \text{ where } A = (a_1 \ldots a_p).$$

If $\{a_1, \ldots, a_p\}$ is an orthogonal basis for W, then

$$\mathrm{proj}_W v = \frac{v \cdot a_1}{a_1 \cdot a_1} a_1 + \cdots + \frac{v \cdot a_p}{a_p \cdot a_p} a_p.$$

The orthogonal projection does not depend on the basis for W taken.

Orthogonal decomposition theorem.

Let W be a subspace of $\mathbb{R}^n$. Any vector v in $\mathbb{R}^n$ is uniquely written as the sum of two vectors

$$v = p_W + e$$

such that $p_W = \mathrm{proj}_W v \in W$ and $e \in W^\perp$.

The best approximation theorem.

The vector $\mathrm{proj}_W v$ is the closest point in W to v, in the sense that,

$$\|v - p_W\| < \|v - \widehat{v}\|$$

for all $\widehat{v} \in W$ distinct from p_W.

$\|v - p_W\| = \mathrm{dist}(v, W)\|$

is the distance from a point to a subspace.

If $v \in W$, then $\mathrm{dist}(v, W) = 0$.

Find the distance from $v = (-1, -5, 10)^T$ to the subspace W spanned by $u_1 = (5, -2, 1)^T$ y $u_2 = (1, 2, -1)^T$.

1. Note that the inner product $u_1 \cdot u_2 = 5 - 4 - 1 = 0$.
2. Then
$$p_W = \frac{v \cdot u_1}{u_1 \cdot u_1} u_1 + \frac{v \cdot u_2}{u_2 \cdot u_2} u_2 = \frac{15}{30} u_1 + \frac{-21}{6} u_2 = \begin{pmatrix} -1 \\ -8 \\ 4 \end{pmatrix}.$$
3. $\mathrm{dist}(v, W) = \|v - p_W\| = \left\| \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix} \right\| = \sqrt{0^2 + 3^2 + 6^2} = 3\sqrt{5}.$

Least-square problem.

1. Consider a system $Ax = b$, in general, inconsistent, where $A \in M_{m\times n}$, $b \in \mathbb{R}^m$.
2. A **least-square solution of** $Ax = b$ is $\widehat{x} \in \mathbb{R}^n$ such that
$$\|b - A\widehat{x}\| \leq \|b - Ax\| \text{ for all } x \in \mathbb{R}^n.$$
3. The least-squares solutions are the **solutions of the normal equations:** $A^T Ax = A^T b$.
4. **Least-square error of the approximation is** $\mathrm{dist}(b, A\widehat{x})$.

There exists a unique least-square solution for each $b \in \mathbb{R}^m$

- if and only if $A^T A$ is invertible,
- if and only if the columns of A are linearly independent.

Gram-Schmidt process.

Let $\{a_1, \ldots, a_k\}$ be a basis for the subspace H in $\mathbb{R}^n$.

$$
\begin{aligned}
\text{STEP 1:} \quad q_1 &= \quad a_1. \\
\text{STEP 2:} \quad q_2 &= \quad a_2 - \mathrm{proj}_W a_2 = a_2 - \frac{a_2 \cdot q_1}{q_1 \cdot q_1}\, q_1, \\
& \qquad \text{where } W = \mathrm{span}\{q_1\}. \\
\text{STEP 3:} \quad q_3 &= \quad a_3 - \mathrm{proj}_W a_3 = a_3 - \frac{a_3 \cdot q_1}{q_1 \cdot q_1}\, q_1 - \frac{a_3 \cdot q_2}{q_2 \cdot q_2}\, q_2, \\
& \qquad \text{where } W = \mathrm{span}\{q_1, q_2\}. \\
& \vdots \\
\text{STEP k:} \quad q_k &= \quad a_k - \mathrm{proj}_W a_k = a_k - \frac{a_k \cdot q_1}{q_1 \cdot q_1}\, q_1 - \cdots - \frac{a_k \cdot q_k}{q_k \cdot q_k}\, q_k, \\
& \qquad \text{where } W = \mathrm{span}\{q_1, \ldots, q_k\}.
\end{aligned}
$$

Then:

- $\{q_1, \ldots, q_k\}$ is an **orthogonal basis for** H.
- $\{w_1, \ldots, w_k\} = \left\{ \dfrac{q_1}{\|q_1\|}, \ldots, \dfrac{q_k}{\|q_k\|} \right\}$ is an **orthonormal basis for** H.