Section 6.2

Orthogonal Complements

Orthogonal Complements

Definition

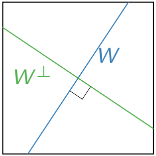
Let W be a subspace of $\mathbf{R}^n$. Its **orthogonal complement** is

$$W^{\perp} = \{ v \text{ in } \mathbf{R}^n \mid v \cdot w = 0 \text{ for all } w \text{ in } W \} \qquad \text{read “}W\text{ perp”}.$$

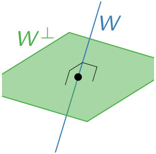
$W^{\perp}$ is orthogonal complement
A^T is transpose

Pictures:

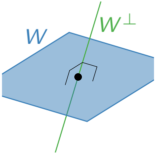
The orthogonal complement of a line in $\mathbf{R}^2$ is the perpendicular line. [interactive]



The orthogonal complement of a line in $\mathbf{R}^3$ is the perpendicular plane. [interactive]



The orthogonal complement of a plane in $\mathbf{R}^3$ is the perpendicular line. [interactive]



Poll

Poll

Let W be a 2-plane in $\mathbf{R}^4$. How would you describe $W^\perp$?

A. The zero space $\{0\}$.

B. A line in $\mathbf{R}^4$.

C. A plane in $\mathbf{R}^4$.

D. A 3-dimensional space in $\mathbf{R}^4$.

E. All of $\mathbf{R}^4$.

For example, if W is the xy-plane, then $W^\perp$ is the zw-plane:

$$\begin{pmatrix} x \\ y \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ z \\ w \end{pmatrix} = 0.$$

Let W be a subspace of $\mathbf{R}^n$.

Facts:

1. $W^\perp$ is also a subspace of $\mathbf{R}^n$
2. $(W^\perp)^\perp = W$
3. $\dim W + \dim W^\perp = n$
4. If $A = \begin{pmatrix} v_1 & v_2 & \cdots & v_m \end{pmatrix}$ and $W = \mathrm{Col}\ A$, then $W^\perp = \mathrm{Nul}(A^T)$ since

$$\begin{aligned} W^\perp &= \text{all vectors orthogonal to each } v_1, v_2, \ldots, v_m \\ &= \left\{ x \text{ in } \mathbf{R}^n \mid x \cdot v_i = 0 \text{ for all } i = 1, 2, \ldots, m \right\} \\ &= \mathrm{Nul} \begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix} = \mathrm{Nul}(A^T). \end{aligned}$$

Let's check 1.

- Is 0 in $W^\perp$? Yes: $0 \cdot w = 0$ for any w in W.
- Suppose x, y are in $W^\perp$. So $x \cdot w = 0$ and $y \cdot w = 0$ for all w in W. Then $(x + y) \cdot w = x \cdot w + y \cdot w = 0 + 0 = 0$ for all w in W. So $x + y$ is also in $W^\perp$.
- Suppose x is in $W^\perp$. So $x \cdot w = 0$ for all w in W. If c is a scalar, then $(cx) \cdot w = c(x \cdot 0) = c(0) = 0$ for any w in W. So cx is in $W^\perp$.

Problem: if $W = \text{Span}\left\{\begin{pmatrix}1\\1\\-1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$, compute $W^\perp$.

By property 4, we have to find the null space of the matrix whose rows are $\begin{pmatrix}1 & 1 & -1\end{pmatrix}$ and $\begin{pmatrix}1 & 1 & 1\end{pmatrix}$, which we did before:

$$\text{Nul}\begin{pmatrix}1 & 1 & -1\\1 & 1 & 1\end{pmatrix} = \text{Span}\left\{\begin{pmatrix}-1\\1\\0\end{pmatrix}\right\}.$$

[interactive]

$$\text{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \text{Nul}\begin{pmatrix}— v_1^T —\\— v_2^T —\\\vdots\\— v_m^T —\end{pmatrix}$$

Definition

The **row space** of an $m \times n$ matrix A is the span of the *rows* of A. It is denoted $\operatorname{Row} A$. Equivalently, it is the column space of A^T:

$$\operatorname{Row} A = \operatorname{Col} A^T.$$

It is a subspace of $\mathbf{R}^n$.

We showed before that if A has rows $v_1^T, v_2^T, \ldots, v_m^T$, then

$$\operatorname{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \operatorname{Nul} A.$$

Hence we have shown:

Fact: $(\operatorname{Row} A)^\perp = \operatorname{Nul} A$.

Replacing A by A^T, and remembering $\operatorname{Row} A^T = \operatorname{Col} A$:

Fact: $(\operatorname{Col} A)^\perp = \operatorname{Nul} A^T$.

Using property 2 and taking the orthogonal complements of both sides, we get:

Fact: $(\operatorname{Nul} A)^\perp = \operatorname{Row} A$ and $\operatorname{Col} A = (\operatorname{Nul} A^T)^\perp$.

Dimension of the row space

Even though $\text{Row}(A)$ lives in $\mathbf{R}^n$ and $\text{Col}(A)$ lives in $\mathbf{R}^m$ if A is an $m \times n$ matrix, both subspaces have the same dimension.

Theorem

If A is an $m \times n$ matrix, then $\dim(\text{Row}\ A) = \dim(\text{Col}\, A)$.

Orthogonal Complements of Most of the Subspaces We've Seen

For any vectors $v_1, v_2, \ldots, v_m$:

$$\mathrm{Span}\{v_1, v_2, \ldots, v_m\}^\perp = \mathrm{Nul}\begin{pmatrix} — v_1^T — \\ — v_2^T — \\ \vdots \\ — v_m^T — \end{pmatrix}$$

For any matrix A:

$$\mathrm{Row}\,A = \mathrm{Col}\,A^T$$

and

$$(\mathrm{Row}\,A)^\perp = \mathrm{Nul}\,A \qquad \mathrm{Row}\,A = (\mathrm{Nul}\,A)^\perp$$

$$(\mathrm{Col}\,A)^\perp = \mathrm{Nul}\,A^T \qquad \mathrm{Col}\,A = (\mathrm{Nul}\,A^T)^\perp$$

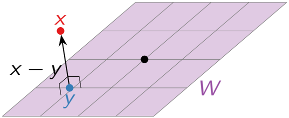
For any other subspace W, first find a basis $v_1, \ldots, v_m$, then use the above trick to compute $W^\perp = \mathrm{Span}\{v_1, \ldots, v_m\}^\perp$.

Section 6.3

Orthogonal Projections (will finish in next set of slides)

Best Approximation

Suppose you measure a data point x which you know for theoretical reasons must lie on a subspace W.



Due to measurement error, though, the measured x is not actually in W. Best approximation: y is the *closest* point to x on W.

How do you know that y is the closest point? The vector from y to x is orthogonal to W: it is in the *orthogonal complement* $W^{\perp}$.

Orthogonal Decomposition

Theorem

Every vector x in $\mathbf{R}^n$ can be written as

$$x = x_W + x_{W^\perp}$$

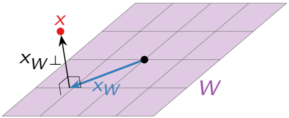
for unique vectors x_W in W and $x_{W^\perp}$ in $W^\perp$.

The equation $x = x_W + x_{W^\perp}$ is called the **orthogonal decomposition** of x (with respect to W).

The vector x_W is the **orthogonal projection** of x onto W.

The vector x_W *is the closest vector to* x *on* W.

[interactive 1] [interactive 2]



Theorem

Every vector x in $\mathbf{R}^n$ can be written as

$$x = x_W + x_{W^\perp}$$

for unique vectors x_W in W and $x_{W^\perp}$ in $W^\perp$.

Why?

Uniqueness: suppose $x = x_W + x_{W^\perp} = x'_W + x'_{W^\perp}$ for x_W, x'_W in W and $x_{W^\perp}, x'_{W^\perp}$ in $W^\perp$. Rewrite:

$$x_W - x'_W = x'_{W^\perp} - x_{W^\perp}.$$

The left side is in W, and the right side is in $W^\perp$, so they are both in $W \cap W^\perp$. But the only vector that is perpendicular to itself is the zero vector! Hence

$$0 = x_W - x'_W \implies x_W = x'_W$$
$$0 = x_{W^\perp} - x'_{W^\perp} \implies x_{W^\perp} = x'_{W^\perp}$$

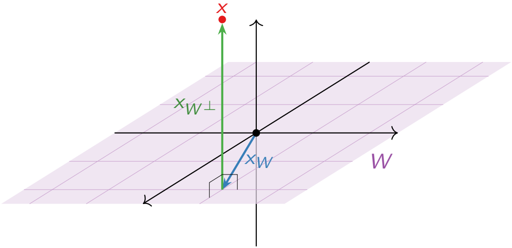
Existence: We will compute the orthogonal decomposition later using orthogonal projections.

Let W be the xy-plane in $\mathbf{R}^3$. Then $W^\perp$ is the z-axis.

$$x = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} \implies x_W = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \qquad x_{W^\perp} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}.$$

$$x = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \implies x_W = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} \qquad x_{W^\perp} = \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}.$$

This is just decomposing a vector into a "horizontal" component (in the xy-plane) and a "vertical" component (on the z-axis).



[interactive]

Problem: Given x and W, how do you compute the decomposition $x = x_W + x_{W^\perp}$?

Observation: It is enough to compute x_W, because $x_{W^\perp} = x - x_W$.

The $A^T A$ Trick

Theorem (The $A^T A$ Trick)

Let W be a subspace of $\mathbf{R}^n$, let $v_1, v_2, \ldots, v_m$ be a spanning set for W (e.g., a basis), and let

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \end{pmatrix}.$$

Then for any x in $\mathbf{R}^n$, the matrix equation

$$A^T A v = A^T x \quad \text{(in the unknown vector } v\text{)}$$

is consistent, and $x_W = Av$ for any solution v.

Recipe for Computing $x = x_W + x_{W^\perp}$

- Write W as a column space of a matrix A.
- Find a solution v of $A^T A v = A^T x$ (by row reducing).
- Then $x_W = Av$ and $x_{W^\perp} = x - x_W$.

Problem: Compute the orthogonal projection of a vector $x = (x_1, x_2, x_3)$ in $\mathbf{R}^3$ onto the xy-plane.

First we need a basis for the xy-plane: let's choose

$$e_1 = \begin{pmatrix}1\\0\\0\end{pmatrix} \qquad e_2 = \begin{pmatrix}0\\1\\0\end{pmatrix} \qquad \rightsquigarrow \qquad A = \begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}.$$

Then

$$A^T A = \begin{pmatrix}1&0\\0&1\end{pmatrix} = I_2 \qquad A^T\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}1&0&0\\0&1&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}x_1\\x_2\end{pmatrix}.$$

Then $A^T Av = v$ and $A^T x = \binom{x_1}{x_2}$, so the only solution of $A^T Av = A^T x$ is $v = \binom{x_1}{x_2}$. Therefore,

$$x_W = Av = A\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}1&0\\0&1\\0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}x_1\\x_2\\0\end{pmatrix}.$$

Problem: Let

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ in } \mathbf{R}^3 \;\middle|\; x_1 - x_2 + x_3 = 0 \right\}.$$

Compute the distance from x to W.

The distance from x to W is $\|x_{W^\perp}\|$, so we need to compute the orthogonal projection. First we need a basis for $W = \operatorname{Nul}\begin{pmatrix} 1 & -1 & 1 \end{pmatrix}$. This matrix is in RREF, so the parametric form of the solution set is

$$\begin{aligned} x_1 &= x_2 - x_3 \\ x_2 &= x_2 \\ x_3 &= \phantom{x_2 - {}} x_3 \end{aligned} \quad \xrightarrow{\text{PVF}} \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Hence we can take a basis to be

$$\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \rightsquigarrow \quad A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

The A^TA Trick

Another Example, Continued

Problem: Let

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \qquad W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ in } \mathbf{R}^3 \;\middle|\; x_1 - x_2 + x_3 = 0 \right\}.$$

Compute the distance from x to W.

We compute

$$A^TA = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A^Tx = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

To solve $A^TAv = A^Tx$ we form an augmented matrix and row reduce:

$$\left(\begin{array}{cc|c} 2 & -1 & 3 \\ -1 & 2 & 2 \end{array}\right) \xrightarrow{\text{RREF}} \begin{pmatrix} 1 & 0 & 8/3 \\ 0 & 1 & 7/3 \end{pmatrix} \rightsquigarrow v = \frac{1}{3}\begin{pmatrix} 8 \\ 7 \end{pmatrix}.$$

$$x_W = Av = \frac{1}{3}\begin{pmatrix} 1 \\ 8 \\ 7 \end{pmatrix} \qquad x_{W^\perp} = x - x_W = \frac{1}{3}\begin{pmatrix} 2 \\ -2 \\ 2 \end{pmatrix}.$$

The distance is $\|x_{W^\perp}\| = \frac{1}{3}\sqrt{4+4+4} \approx 1.155.$

[interactive]

Theorem (The A^TA Trick)

Let W be a subspace of $\mathbf{R}^n$, let $v_1, v_2, \ldots, v_m$ be a spanning set for W (e.g., a basis), and let

$$A = \begin{pmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_m \\ | & | & & | \end{pmatrix}.$$

Then for any x in $\mathbf{R}^n$, the matrix equation

$$A^TAv = A^Tx \qquad \text{(in the unknown vector } v\text{)}$$

is consistent, and $x_W = Av$ for any solution v.

Proof: Let $x = x_W + x_{W^\perp}$. Then $x_{W^\perp}$ is in $W^\perp = \mathrm{Nul}(A^T)$, so $A^Tx_{W^\perp} = 0$. Hence

$$A^Tx = A^T(x_W + x_{W^\perp}) = A^Tx_W + A^Tx_{W^\perp} = A^Tx_W.$$

Since x_W is in $W = \mathrm{Span}\{v_1, v_2, \ldots, v_m\}$, we can write

$$x_W = c_1v_1 + c_2v_2 + \cdots + c_mv_m.$$

If $v = (c_1, c_2, \ldots, c_m)$ then $Av = x_W$, so

$$A^Tx = A^Tx_W = A^TAv.$$

Orthogonal Projection onto a Line

Problem: Let $L = \text{Span}\{u\}$ be a line in $\mathbf{R}^n$ and let x be a vector in $\mathbf{R}^n$. Compute x_L.

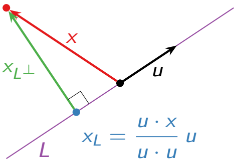
We have to solve $u^T uv = u^T x$, where u is an $n \times 1$ matrix. But $u^T u = u \cdot u$ and $u^T x = u \cdot x$ are scalars, so

$$v = \frac{u \cdot x}{u \cdot u} \qquad \Longrightarrow \qquad x_L = uv = \frac{u \cdot x}{u \cdot u} u.$$

Projection onto a Line

The projection of x onto a line $L = \text{Span}\{u\}$ is

$$x_L = \frac{u \cdot x}{u \cdot u} u \qquad x_{L^\perp} = x - x_L.$$

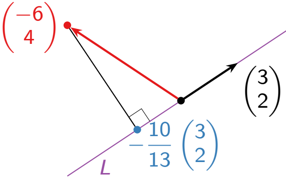


Problem: Compute the orthogonal projection of $x = \binom{-6}{4}$ onto the line L spanned by $u = \binom{3}{2}$, and find the distance from u to L.

$$x_L = \frac{x \cdot u}{u \cdot u}\, u = \frac{-18+8}{9+4}\begin{pmatrix}3\\2\end{pmatrix} = -\frac{10}{13}\begin{pmatrix}3\\2\end{pmatrix} \qquad x_{L^\perp} = x - x_L = \frac{1}{13}\begin{pmatrix}-48\\72\end{pmatrix}.$$

The distance from x to L is

$$\|x_{L^\perp}\| = \frac{1}{13}\sqrt{48^2 + 72^2} \approx 6.656.$$



[interactive]

Summary

Let W be a subspace of $\mathbf{R}^n$.

- The **orthogonal complement** $W^\perp$ is the set of all vectors orthogonal to everything in W.
- We have $(W^\perp)^\perp = W$ and $\dim W + \dim W^\perp = n$.
- $\operatorname{Row} A = \operatorname{Col} A^T$, $(\operatorname{Row} A)^\perp = \operatorname{Nul} A$, $\operatorname{Row} A = (\operatorname{Nul} A)^\perp$, $(\operatorname{Col} A)^\perp = \operatorname{Nul} A^T$, $\operatorname{Col} A = (\operatorname{Nul} A^T)^\perp$.
- **Orthogonal decomposition:** any vector x in $\mathbf{R}^n$ can be written in a unique way as $x = x_W + x_{W^\perp}$ for x_W in W and $x_{W^\perp}$ in $W^\perp$. The vector x_W is the **orthogonal projection** of x onto W.
- The vector x_W is the *closest point to x in W*: it is the *best approximation*.
- The *distance* from x to W is $\|x_{W^\perp}\|$.
- If $W = \operatorname{Col} A$ then to compute x_W, solve the equation $A^T Av = A^T x$; then $x_W = Av$.
- If $W = L = \operatorname{Span}\{u\}$ is a line then $x_L = \frac{u\cdot x}{u\cdot u}\, u$.