Sections 6.2 and 6.3

Orthogonal Projections

Orthogonal Complements

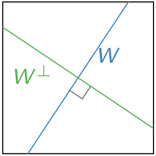
Definition

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

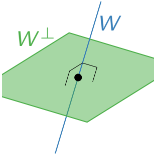
Its **orthogonal complement**, written $W^\perp$ (read "W perp"), is the set of all vectors in $\mathbf{R}^n$ which are orthogonal (perpendicular) to W. We will focus on when $n = 2$ and $n = 3$.

Pictures:

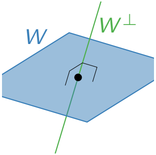
The orthogonal complement of a line in $\mathbf{R}^2$ is the perpendicular line. [interactive]



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]



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

Let $v_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$ and $v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

A vector $u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is in $W^\perp$ if and only if $u \perp v_1$ and $u \perp v_2$.

Why? If $u \perp v_1$ and $u \perp v_2$, then for any scalars c_1 and c_2:

$$u \cdot (c_1 v_1 + c_2 v_2) = c_1(u \cdot v_1) + c_2(u \cdot v_2) = c_1(0) + c_2(0) = 0,$$

Therefore, u will be orthogonal to every vector in $\mathrm{Span}\{v_1, v_2\}$.

Computation, continued

Now $u \perp v_1$ means $x + y - z = 0$ and $u \perp v_2$ means $x + y + z = 0$. This means $u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ satisfies

$$\begin{aligned} x + y - z &= 0 \\ x + y + z &= 0, \end{aligned}$$

which means u is in $\mathrm{Nul}\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}$. Therefore,

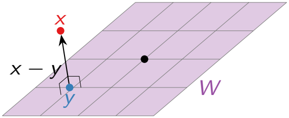
$$W^\perp = \mathrm{Nul}\begin{pmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix} = (\text{... with work ...}) = \mathrm{Span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}.$$

[interactive]

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

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

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

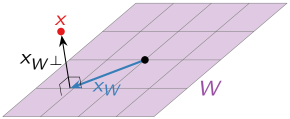
for unique vectors x_W in W and $x_{W^\perp}$ in $W^\perp$.

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]



Orthogonal Decomposition

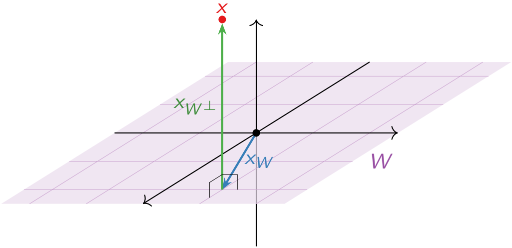
Example

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]

Orthogonal Decomposition

Computation?

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^TA Trick to compute x_W and $x_{W^\perp}$

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

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

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 $A^TAv = A^Tx$ (by row reducing).
- Then $x_W = Av$ and $x_{W^\perp} = x - x_W$.

Problem: Let

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

Find x_W, $x_{W^\perp}$, and 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 $W = \mathrm{Nul}\begin{pmatrix}1 & -1 & 1\end{pmatrix}$. 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 $A^T A$ Trick

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

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

To solve $A^T A v = A^T x$ we form an augmented matrix and row reduce:

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

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

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

[interactive]

Theorem (The $A^T A$ 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

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

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 $A^T x_{W^\perp} = 0$. Hence

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

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

$$x_W = c_1 v_1 + c_2 v_2 + \cdots + c_m v_m.$$

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

$$A^T x = A^T x_W = A^T A v.$$

Orthogonal Projection onto a Line

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

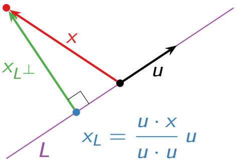
We have to solve $u^T uv = u^T x$, where u is an $n \times 1$ matrix. But $u^T u = u \cdot u$ and $u^T x = u \cdot x$ are scalars, so

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

Projection onto a Line

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

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



Orthogonal Projection onto a Line

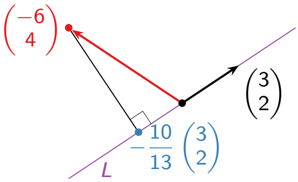
Example

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]

Projection Matrix
Method 1

Let W be a subspace of $\mathbf{R}^n$ and let $T\colon \mathbf{R}^n \to \mathbf{R}^n$ be the orthogonal projection with respect to W.

How do you compute the standard matrix A for T?

The same as any other linear transformation:

$$A = \begin{pmatrix} T(e_1) & T(e_2) & \cdots & T(e_n) \end{pmatrix}.$$

Projection Matrix

Method 1, Example 1

Problem: Let $L = \mathrm{Span}\left\{\binom{3}{2}\right\}$ and let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be the orthogonal projection onto L. Compute the matrix A for T.

It's easy to compute orthogonal projection onto a line:

$$\left.\begin{aligned} T(e_1) &= (e_1)_L = \frac{u\cdot e_1}{u\cdot u}\,u = \frac{3}{13}\begin{pmatrix}3\\2\end{pmatrix} \\ T(e_2) &= (e_2)_L = \frac{u\cdot e_2}{u\cdot u}\,u = \frac{2}{13}\begin{pmatrix}3\\2\end{pmatrix}\end{aligned}\right\} \quad\Longrightarrow\quad A = \frac{1}{13}\begin{pmatrix}9 & 6\\ 6 & 4\end{pmatrix}.$$

Projection Matrix

Method 1, Example 2

Problem: Let

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

and let $T\colon \mathbf{R}^3 \to \mathbf{R}^3$ be orthogonal projection onto W. Compute the matrix B for T.

We computed $W = \operatorname{Col} A$ for

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

To compute $T(e_i)$ we have to solve the matrix equation $A^T A v = A^T e_i$. We have

$$A^T A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \qquad A^T e_i = \text{the } i\text{th column of } A^T = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Problem: Let

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

and let $T\colon \mathbf{R}^3 \to \mathbf{R}^3$ be orthogonal projection onto W. Compute the matrix B for T.

$$\left(\begin{array}{cc|c} 2 & -1 & 1 \\ -1 & 2 & -1 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & -1/3 \end{array}\right) \implies T(e_1) = \frac{1}{3}A\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 2 & -1 & 1 \\ -1 & 2 & 0 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 2/3 \\ 0 & 1 & 1/3 \end{array}\right) \implies T(e_2) = \frac{1}{3}A\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 2 & -1 & 0 \\ -1 & 2 & 1 \end{array}\right) \xrightarrow{\text{RREF}} \left(\begin{array}{cc|c} 1 & 0 & 1/3 \\ 0 & 1 & 2/3 \end{array}\right) \implies T(e_2) = \frac{1}{3}A\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

$$\implies B = \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}.$$

Theorem

Let $\{v_1, v_2, \ldots, v_m\}$ be a *linearly independent* set in $\mathbf{R}^n$, and let

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

Then the $m \times m$ matrix $A^T A$ is invertible.

Proof: We'll show $\mathrm{Nul}(A^T A) = \{0\}$. Suppose $A^T A v = 0$. Then Av is in $\mathrm{Nul}(A^T) = \mathrm{Col}(A)^\perp$. But Av is in $\mathrm{Col}(A)$ as well, so $Av = 0$, and hence $v = 0$ because the columns of A are linearly independent.

Theorem

Let $\{v_1, v_2, \ldots, v_m\}$ be a *linearly independent* set in $\mathbf{R}^n$, and let

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

Then the $m \times m$ matrix $A^T A$ is invertible.

Let W be a subspace of $\mathbf{R}^n$ and let $T\colon \mathbf{R}^n \to \mathbf{R}^n$ be the orthogonal projection with respect to W. Let $\{v_1, v_2, \ldots, v_m\}$ be a *basis* for W and let A be the matrix with columns $v_1, v_2, \ldots, v_m$. To compute $T(x) = x_W$ you solve $A^T A v = Ax$; then $x_W = Av$.

$$v = (A^T A)^{-1}(A^T x) \implies T(x) = Av = \left[A(A^T A)^{-1} A^T\right] x.$$

If the columns of A are a *basis* for W then the matrix for T is

$$A(A^T A)^{-1} A^T.$$

Projection Matrix
Method 2, Example 1

Problem: Let $L = \mathrm{Span}\left\{\binom{3}{2}\right\}$ and let $T\colon \mathbf{R}^2 \to \mathbf{R}^2$ be the orthogonal projection onto L. Compute the matrix A for T.

The set $\left\{\binom{3}{2}\right\}$ is a basis for L, so

$$A = u(u^T u)^{-1} u^T = \frac{1}{u \cdot u} u u^T = \frac{1}{13}\begin{pmatrix}3\\2\end{pmatrix}\begin{pmatrix}3 & 2\end{pmatrix} = \frac{1}{13}\begin{pmatrix}9 & 6\\6 & 4\end{pmatrix}.$$

Matrix of Projection onto a Line

If $L = \mathrm{Span}\{u\}$ is a line in $\mathbf{R}^n$, then the matrix for projection onto L is

$$\frac{1}{u \cdot u} u u^T.$$

Projection Matrix

Method 2, Example 2

Problem: Let

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

and let $T\colon \mathbf{R}^3 \to \mathbf{R}^3$ be orthogonal projection onto W. Compute the matrix B for T.

In the slides for the last lecture we computed $W = \operatorname{Col} A$ for

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The columns are linearly independent, so they form a basis for W. Hence

$$\begin{aligned} B = A(A^TA)^{-1}A^T = A\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}^{-1} A^T &= \frac{1}{3}A\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} A^T \\ &= \frac{1}{3}\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix}. \end{aligned}$$

Poll

Let W be a subspace of $\mathbf{R}^n$ which is neither the zero subspace nor all of $\mathbf{R}^n$.

Poll

Let A be the matrix for proj_W. What is/are the eigenvalue(s) of A?

A. 0 B. 1 C. -1 D. $0, 1$ E. $1, -1$ F. $0, -1$ G. $-1, 0, 1$

The 1-eigenspace is W.

The 0-eigenspace is $W^\perp$.

We have $\dim W + \dim W^\perp = n$, so that gives n linearly independent eigenvectors already.

So the answer is D.

Theorem

Let W be an m-dimensional subspace of $\mathbf{R}^n$, let $T\colon \mathbf{R}^n \to W$ be the projection, and let A be the matrix for T. Then:

1. $\operatorname{Col} A = W$, which is the 1-eigenspace.
2. $\operatorname{Nul} A = W^\perp$, which is the 0-eigenspace.
3. $A^2 = A$.
4. A is similar to the diagonal matrix with m ones and $n - m$ zeros on the diagonal.

Proof of 4: Let $v_1, v_2, \ldots, v_m$ be a basis for W, and let $v_{m+1}, v_{m+2}, \ldots, v_n$ be a basis for $W^\perp$. These are (linearly independent) eigenvectors with eigenvalues 1 and 0, respectively, and they form a basis for $\mathbf{R}^n$ because there are n of them.

Example: If W is a plane in $\mathbf{R}^3$, then A is similar to projection onto the xy-plane:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$