CS132 Lecture 20

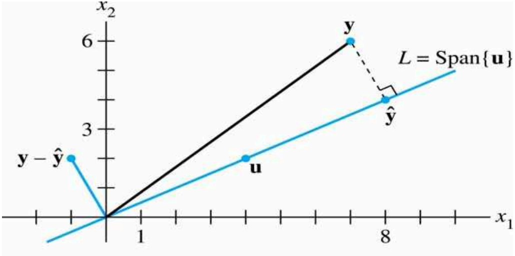
Orthogonal Projections

April 21, 2015

Outline

Orthogonal projection in $\mathbb{R}^2$

- Recall Example 3 of section 6.2: the orthogonal projection of a point in $\mathbb{R}^2$ onto a line L through the origin.

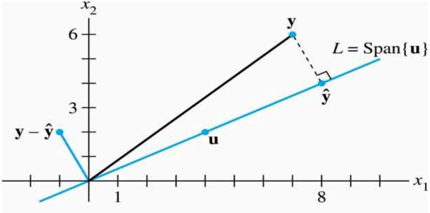


The orthogonal projection of **y** onto a line L through the origin.

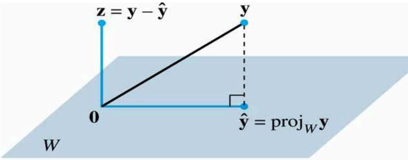
$$\hat{\mathbf{y}} = proj_L \mathbf{y} = \frac{\mathbf{y}^T\mathbf{u}}{\mathbf{u}^T\mathbf{u}}\mathbf{u}$$

Orthogonal projection in $\mathbb{R}^n$

- Analogue in $\mathbb{R}^n$: orthogonal projection of a point in $\mathbb{R}^n$ onto a subspace W (through the origin by definition)



The orthogonal projection of $\mathbf{y}$ onto a line L through the origin.



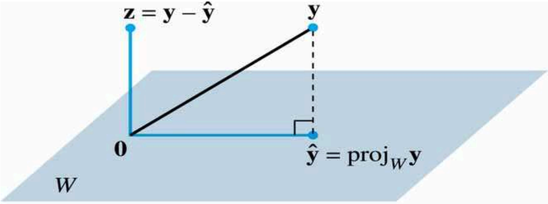
The orthogonal projection of $\mathbf{y}$ onto W.

- exists a unique decomposition $\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z}$, such that $\hat{\mathbf{y}}$ is the unique vector in W
  - for which $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W.
  - $\hat{\mathbf{y}}$ closet to $\mathbf{y}$.

Orthogonal Complements

Definition (Orthogonal Complement)

If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be **orthogonal to** W. The set of all vectors $\mathbf{z}$ that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^\perp$ (read as "W perpendicular").



The orthogonal projection of $\mathbf{y}$ onto W.

An example in $\mathbb{R}^5$

Example 1

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_5\}$ be an orthogonal basis for $\mathbb{R}^5$ and let

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

Consider the subspace $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$, and write $\mathbf{y}$ as the sum of a vector $\mathbf{z}_1$ in W and a vector $\mathbf{z}_2$ in $W^\perp$.

Solution Write

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + c_2\mathbf{u}_2}_{\mathbf{z}_1} + \underbrace{c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5}_{\mathbf{z}_2}$$

where $\mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ is in $Span\{\mathbf{u}_1, \mathbf{u}_2\}$
and $\mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5$ is in $Span\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.

An example in $\mathbb{R}^5$ (Continue)

To show that $\mathbf{z}_2$ is in $W^\perp$, it suffices to show that $\mathbf{z}_2$ is orthogonal to the vectors in the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W. (From Section 6.1.) Using properties of the inner product, compute

$$\begin{aligned} \mathbf{z}_2^T \mathbf{u}_1 &= (c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5)^T \mathbf{u}_1 \\ &= c_3\mathbf{u}_3^T \mathbf{u}_1 + c_4\mathbf{u}_4^T \mathbf{u}_1 + c_5\mathbf{u}_5^T \mathbf{u}_1 \\ &= 0 \end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_3, \mathbf{u}_4$, and $\mathbf{u}_5$. A similar calculation shows that $\mathbf{z}_2^T \mathbf{u}_2 = 0$. Thus $\mathbf{z}_2$ is in $W^\perp$.

Question?

Do we really need to know $\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$, i.e. the orthogonal basis for $W^\perp$, to compute this decomposition?

The Orthogonal Decomposition Theorem

Theorem (Theorem 8)

Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the from

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

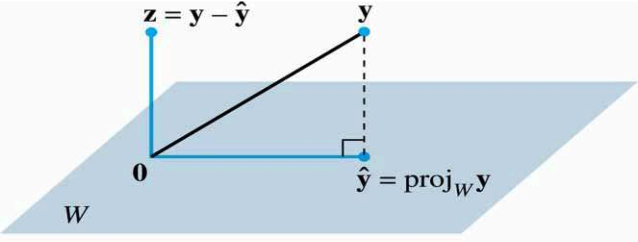
where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. In fact, if $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\hat{\mathbf{y}} = \frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}^T\mathbf{u}_p}{\mathbf{u}_p^T\mathbf{u}_p}\mathbf{u}_p \tag{2}$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

The Orthogonal Decomposition Theorem

The vector $\hat{\mathbf{y}}$ in (1) is called the **orthogonal projection of y onto** W and often is written as $proj_W\mathbf{y}$.



The orthogonal projection of $\mathbf{y}$ onto W.

The Orthogonal Decomposition Theorem

Proof.

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ be any othogonal basis for W, and define $\hat{\mathbf{y}}$ by (2). Then $\hat{\mathbf{y}}$ is in W because $\hat{\mathbf{y}}$ is a linear combination of the basis $\mathbf{u}_1, \ldots, \mathbf{u}_p$. Let $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$. Since $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \ldots, \mathbf{u}_p$, it follows from (2) that

$$\mathbf{z}^T\mathbf{u}_1 = (\mathbf{y} - \hat{\mathbf{y}})^T\mathbf{u}_1 = \mathbf{y}^T\mathbf{u}_1 - \left(\frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\right)\mathbf{u}_1^T\mathbf{u}_1 = \mathbf{y}^T\mathbf{u}_1 - \mathbf{y}^T\mathbf{u}_1 = 0$$

Thus $\mathbf{z}$ is orthogonal to $\mathbf{u}_1$. Similarly, $\mathbf{z}$ is orthogonal to each $\mathbf{u}_j$ in the basis for W. Hence $\mathbf{z}$ is orthogonal to every vector in W, i.e., $\mathbf{z}$ is in $W^\perp$. To show that the decomposition in (1) is unique, suppose $\mathbf{y}$ can also be written as $\mathbf{y} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$, with $\hat{\mathbf{y}}_1$ in W and $\mathbf{z}_1$ in $W^\perp$. Then $\hat{\mathbf{y}} + \mathbf{z} = \hat{\mathbf{y}}_1 + \mathbf{z}_1$, and so

$$\hat{\mathbf{y}} - \hat{\mathbf{y}}_1 = \mathbf{z}_1 - \mathbf{z}$$

This equality shows that the vector $\mathbf{v} = \hat{\mathbf{y}} - \hat{\mathbf{y}}_1$ is in W and in $W^\perp$. Hence $\mathbf{v}^T\mathbf{v} = 0$, which shows that $\mathbf{v} = 0$. Therefore $\hat{\mathbf{y}} = \hat{\mathbf{y}}_1$ and $\mathbf{z}_1 = \mathbf{z}$. $\square$

The Orthogonal Decomposition Theorem

Example 2

Let $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

Solution The orthogonal projection of $\mathbf{y}$ onto W is

$$\begin{aligned} \hat{\mathbf{y}} &= \frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}^T\mathbf{u}_2}{\mathbf{u}_2^T\mathbf{u}_2}\mathbf{u}_2 \\ &= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} \end{aligned}$$

The Orthogonal Decomposition Theorem

Also

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

Theorem 8 ensures that $\mathbf{y} - \hat{\mathbf{y}}$ is in $W^{\perp}$. To check?

The desired decomposition of $\mathbf{y}$

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

The Orthogonal Decomposition Theorem

Corollary (Theorem 10)

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ *is an orthonormal basis for a subspace* W *of* $\mathbb{R}^n$*, then*

$$proj_W \mathbf{y} = (\mathbf{y}^T\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}^T\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y}^T\mathbf{u}_p)\mathbf{u}_p \tag{3}$$

if $U = [\mathbf{u}_1\ \mathbf{u}_2\ \cdots\ \mathbf{u}_p]$*, then*

$$proj_W \mathbf{y} = UU^T\mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n \tag{4}$$

Proof.

Formula (3) follows immediately from (2) in Theorem 8. Also, (3) shows that $proj_W \mathbf{y}$ is a linear combination of the columns of U using the weights $\mathbf{y}^T\mathbf{u}_1, \ldots, \mathbf{y}^T\mathbf{u}_p$. The weights can be written as $\mathbf{u}_1^T\mathbf{y}, \ldots, \mathbf{u}_p^T\mathbf{y}$, showing that they are the entries in $U^T\mathbf{y}$ and justifying (4). □

Properties of orthogonal projections

Theorem (The Best Approximation Theorem)

Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \tag{5}$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

$\hat{\mathbf{y}}$ is called **the best approximation to y by elements of** W.
In particular, if $\mathbf{y} \in W$, then $\hat{\mathbf{y}} = \mathbf{y}$.

- "closeness" depends on the measure of length which is induced by **inner product**, so does the measure of angle on which "orthogonality" depends.
- $\hat{\mathbf{y}}$ is the unique minimizer of $\|\mathbf{y} - \mathbf{v}\|$ over W, independent of the particular orthogonal basis of W to compute it.

The Best Approximation Theorem

Proof.

Take $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$. See the figure. Then $\hat{\mathbf{y}} - \mathbf{v}$ is in W. By Theorem 8, $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W. In particular, $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to $\hat{\mathbf{y}} - \mathbf{v}$ (which is in W). Since

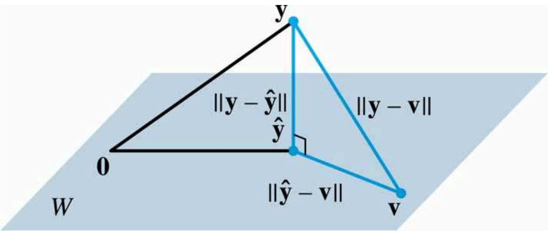
$$\mathbf{y} - \mathbf{v} = (\mathbf{y} - \hat{\mathbf{y}}) + (\hat{\mathbf{y}} - \mathbf{v})$$

the Pythagorean Theorem gives

$$\|\mathbf{y} - \mathbf{v}\|^2 = \|\mathbf{y} - \hat{\mathbf{y}}\|^2 + \|\hat{\mathbf{y}} - \mathbf{v}\|^2$$

Now $\|\hat{\mathbf{y}} - \mathbf{v}\|^2 > 0$ because $\hat{\mathbf{y}} - \mathbf{v} \neq 0$, and so inequality (5) follows immediately. $\square$

The Best Approximation Theorem



The orthogonal projection of **y** onto W is the closest point in W to **y**.

Properties of orthogonal projections

Example 3

If $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, and $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$, as in Example 2, then the closest point in W to $\mathbf{y}$ is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}^T\mathbf{u}_1}{\mathbf{u}_1^T\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}^T\mathbf{u}_2}{\mathbf{u}_2^T\mathbf{u}_2}\mathbf{u}_2 = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

Example 4

The distance from a point $\mathbf{y}$ in $\mathbb{R}^n$ to a subspace W is defined as the distance from $\mathbf{y}$ to the nearest point in W. Find the distance from $\mathbf{y}$ to $W = Span\{\mathbf{u}_1, \mathbf{u}_2\}$, where

$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \ \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \ \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

Solution By the Best Approximation Theorem, the distance from $\mathbf{y}$ to W is $\|\mathbf{y} - \hat{\mathbf{y}}\|$, where $\hat{\mathbf{y}} = proj_W \mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is a orthogonal basis for W,

$$\hat{\mathbf{y}} = \frac{15}{30}\mathbf{u}_1 + \frac{-21}{6}\mathbf{u}_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

Properties of orthogonal projections

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = 3^2 + 6^2 = 45$$

The distance from $\mathbf{y}$ to W is $\sqrt{45} = 3\sqrt{5}$.