Vectors

MATH 3512, BCIT

Matrix Methods and Statistics for Geomatics

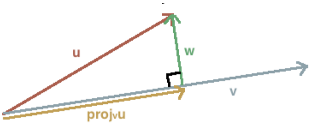
October 8, 2018

Projection

The projection u_H of a vector u onto a hyperplane H is the vector in the hyperplane that is "most similar" to u. The formal definition for u_H requires that

1. u is in H
2. $(u - u_H)$ is orthogonal to all basis vectors of H

Example 1: Finding a Projection. Let H be the line spanned by $\vec{v} = (-1, 1)^\intercal$ in $\mathbb{R}^2$. What is the projection $\vec{w}$ of $\vec{u} = (3, -2)^\intercal$?



Let $\vec{w} = (w_1, w_2)^\intercal$. Then (1) $\vec{u} - \vec{w}$ is orthogonal to $\vec{v}$ and (2) $\vec{w} = \alpha\vec{v}$ for some $\alpha \in \mathbb{R}$.

$$\begin{array}{ccccc} w_1 & - & w_2 & = & 5 \\ w_1 & + & w_2 & = & 0 \end{array} \tag{1}$$

Cramer's rule tells us that $\vec{w} = (2.5, -2.5)^\intercal$.

Projection

Let $u = (u_1, \ldots, u_n)^\intercal$ be a vector and H be a k-dimensional hyperplane in the vector space $\mathbb{R}^n$. Let $x_1, \ldots, x_k$ be a basis for H. Then it is true for all vectors v in the hyperplane that

$$\|u - v\| \geq \|u - u_H\| \tag{2}$$

Proof: use the theorem of Pythagoras for

$$\|u - v\|^2 = \|u - u_H\|^2 + \|u_H - v\|^2 \geq \|u - u_H\|^2 \tag{3}$$

The claim follows. It illustrates what I mean when I say that u_H is the vector in H that is most similar to u.

Example 2: Finding Another Projection. What is the projection of $\vec{u} = (5, 2, 10)^\intercal$ onto the plane T characterized by $2x + y + 3z = 0$?

First we find two linearly independent vectors in H to form a basis of H, for example $\vec{v_1} = (1, 1, -1)^\intercal$ and $\vec{v_2} = (0, -3, 1)^\intercal$. The conditions

1. $u_H \in T$
2. $(u - u_H) \perp v_1$
3. $(u - u_H) \perp v_2$

give us the system of linear equations

$$\begin{bmatrix} 2 & 1 & 3 \\ -1 & 1 & 1 \\ 0 & 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ -4 \end{bmatrix} \tag{4}$$

for which the solution is $u_H = (\hat{x}, \hat{y}, \hat{z})^\intercal = (-1, -1, 1)^\intercal$.

Least Squares Method

Let there be two linearly independent vectors u and v in $\mathbb{R}^n$. Then the formula for the projection u_v of u onto the line spanned by v is

$$u_v = \left(\frac{u \cdot v}{v \cdot v}\right) v \tag{5}$$

To verify the formula, note that $u_v = av$ for some real number a. Therefore

$$(u - av) \perp v \tag{6}$$

Isolate a in the equation $(u - av) \cdot v = 0$ to yield the formula.

Least Squares Method

Formula (5) only works when the hyperplane is a line. You can scale up the idea in terms of dimensions by the following theorem.

Formula for Projection Onto Plane with Orthogonal Basis

Let $\{u, v\}$ be an orthogonal basis for H. Then the projection of w onto H is the sum of w_u and w_v, the projections of w onto the lines spanned by u and v, respectively.

Proof: check the following

1. $(w_u + w_v) \in H$ (trivial)
2. $(w - (w_u + w_v)) \perp u$ (use the fact that $u \perp v$)
3. $(w - (w_u + w_v)) \perp v$ (same idea)

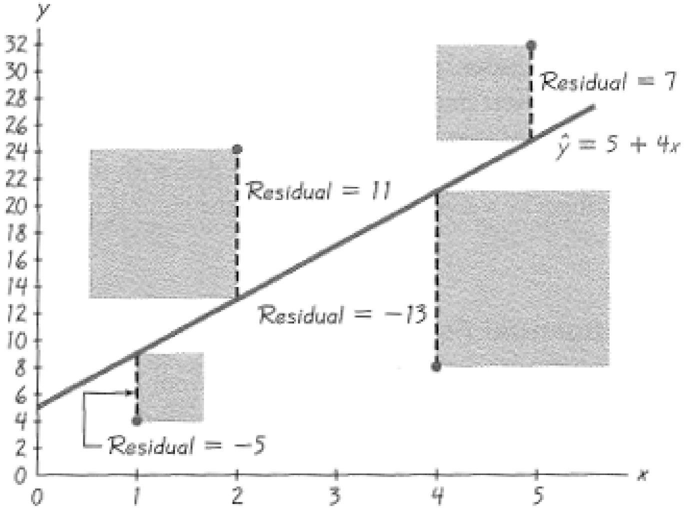
Least Squares Method

Consider the following table of measurements for the length of shoe prints and the height of the person wearing the shoes.

Shoe Print (cm)	Height (cm)
29.7	175.3
29.9	177.8
31.4	185.4
31.8	175.3
27.6	172.7

In the statistics portion of this course, we will learn whether the paired data provide evidence of a linear relationship. In the linear algebra portion, we will learn how to find the line which is closest to the data points in the least squares sense.

Least Squares Method



Least Squares Method

Least Squares Method

If L is a given line, the error for each data point is the vertical distance from that point to the line. The squared error is the sum of the squares of the errors. The line that best fits the data in the least squares sense is the line that minimizes the squared error.

You can find the regression line using calculus optimization. However, there is also an elegant method using linear algebra.

Let L be a line with slope m and y-intercept b. Let $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ be a set of paired data. Then the following equations hold:

$$\begin{array}{ccccccc} y_1 & = & mx_1 & + & b & + & \epsilon_1 \\ y_2 & = & mx_2 & + & b & + & \epsilon_2 \\ & & & \vdots & & & \\ y_n & = & mx_n & + & b & + & \epsilon_n \end{array} \tag{7}$$

where the ϵ_i are the errors $(i = 1, \ldots, n)$. This system is equivalent to the following vector equation,

$$Y = AV + E \tag{8}$$

where Y, A, V, E are defined on the next slide.

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, A = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}, V = \begin{bmatrix} m \\ b \end{bmatrix}, E = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

E is called the error vector. According to (8), it is

$$E = Y - AV \tag{9}$$

We are trying to choose m, b so that

$$\|E\|^2 = \|Y - AV\|^2 \tag{10}$$

is minimal.

Least Squares Method

Let

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \tag{11}$$

Then $AV = mX + bB$. The set $S = \{AV | m, b \in \mathbb{R}\}$ is a plane in n-dimensional space. The ordered pair (m, b) that minimizes the squared error corresponds to the projection Y_S of Y onto S.

Least Squares Method

Let C be $B - B_X$, where B_X is the projection of B onto the line spanned by X. Then

$$C = B - \left(\frac{B \cdot X}{X \cdot X}\right) X \tag{12}$$

Note that $X \perp C$. X and C form an orthogonal basis for S, therefore

$$Y_S = Y_X + Y_C = \left(\frac{Y \cdot X}{X \cdot X}\right) X + \left(\frac{Y \cdot C}{C \cdot C}\right) C \tag{13}$$

Least Squares Method

Replace the rightmost C by $B - \left(\frac{B \cdot X}{X \cdot X}\right) X$ for

$$Y_S = \left(\frac{Y \cdot X}{X \cdot X} - \left(\frac{Y \cdot C}{C \cdot C}\right)\left(\frac{B \cdot X}{X \cdot X}\right)\right) X + \left(\frac{Y \cdot C}{C \cdot C}\right) B \tag{14}$$

or alternatively

$$m = \frac{Y \cdot X}{X \cdot X} - \left(\frac{Y \cdot C}{C \cdot C}\right)\left(\frac{B \cdot X}{X \cdot X}\right) \tag{15}$$

$$b = \frac{Y \cdot C}{C \cdot C} \tag{16}$$

with

$$C = B - \left(\frac{B \cdot X}{X \cdot X}\right) X \tag{17}$$

Example 3: Angles at Gray Cliff. This example is from Oscar S. Adams's *Application of the Theory of Least Squares to the Adjustment of Triangulation* (1915), a "working manual for the computer in the office." You measure the following angles.

from	to	angle
Boulder	Tower	$65^\circ 6' 29.3''$
Tower	Tyonek	$19^\circ 46' 26.9''$
Tyonek	Round Point	$8^\circ 39' 14.2''$
Round Point	Boulder	$266^\circ 27' 47.9''$

Angles at Gray Cliff

Notice that the angles do not add up to 360°. We are missing $1.7''$. How should we adjust these numbers?

Basic assumptions underlying least squares theory in surveying are

1. mistakes and systematic errors have been eliminated
2. the number of observations being adjusted is large
3. the frequency distribution of the errors is normal

Angles at Gray Cliff

Convert the angles to real numbers

$$\hat{a} = 65.108, \hat{b} = 19.774, \hat{c} = 8.6539, \hat{d} = 266.46 \tag{18}$$

The sum is 359.999527778. Here is a system of equations with measurement errors, exploiting the fact that d is supposed to be $360^\circ - (a + b + c)$

$$\begin{aligned} a &= 65.108 + \epsilon_1 \\ b &= 19.774 + \epsilon_2 \\ c &= 8.6539 + \epsilon_3 \\ 360 - (a + b + c) &= 266.46 + \epsilon_4 \end{aligned} \tag{19}$$

Angles at Gray Cliff

The system of equations is equivalent to the following matrix equation.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 65.108 \\ 19.774 \\ 8.6539 \\ -93.537 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \end{bmatrix} \tag{20}$$

In symbols,

$$AV = Y + E \tag{21}$$

Again, we want to minimize

$$\|Y - AV\|^2 = \|E\|^2 \tag{22}$$

End of Lesson

Next Lesson: Eigenvalues and Eigenvectors