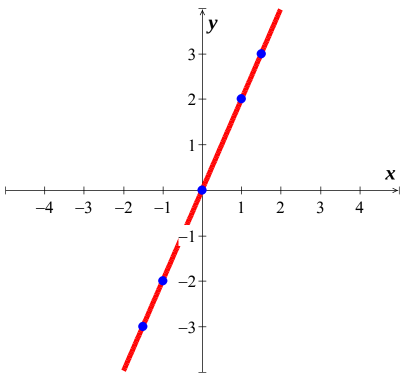
Section 3.5 – Least Squares Analysis

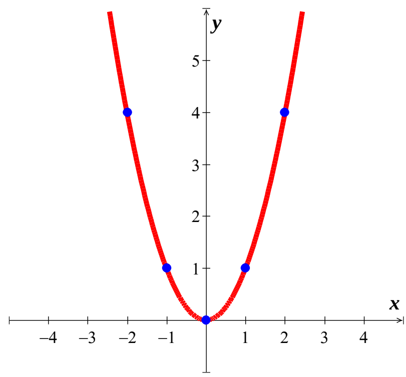
The use to ***best*** fit data, we will use results about orthogonal projections in inner product spaces to obtain a technique for fitting a line or other polynomial.

Fitting a Curve to Data

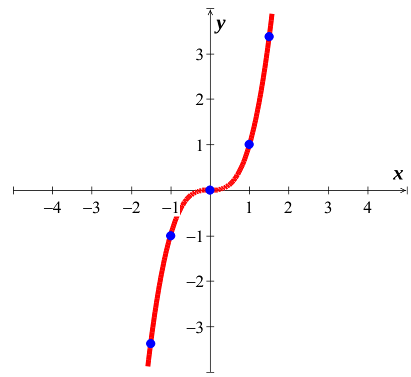
The common problem is to obtain a mathematical relationship between 2 variables x and y by ***fitting*** a curve to points in the xy-plane.
Some possibility of fitting the data



$y = ax + b$



$y = ax^2 + bx + c$



$y = ax^3 + bx^2 + cx + d$

Least Squares Fit of a Straight Line

Recall that a system of equations $A\mathbf{x} = \mathbf{y}$ is called inconsistent if it does not have a solution. Suppose we want to fit a straight line $y = mx + b$ to the determined points $\left(x_1, y_1\right), \ \ldots, \ \left(x_n, y_n\right)$

If the data points were collinear, the line would pass through all n points and the unknown coefficients m and b would satisfy the equations

$$\begin{aligned} y_1 &= mx_1 + b \\ y_2 &= mx_2 + b \\ \vdots \quad & \quad \vdots \quad \quad \vdots \\ y_n &= mx_n + b \end{aligned} \quad \Rightarrow \quad \underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} m \\ b \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\mathbf{y}}$$

The problem is to find m and b that minimize the errors is some sense.

Least Square Problem

Given a linear system $A\boldsymbol{x} = \boldsymbol{y}$ of m equations in n unknowns, find a vector $\boldsymbol{x}$ that minimizes $\|\boldsymbol{y} - A\boldsymbol{x}\|$ with respect to the Euclidean inner product on $\boldsymbol{R}^m$. We call such as $\boldsymbol{x}$ a least squares solution of the system, we call $\boldsymbol{y} - A\boldsymbol{x}$ the least squares error vectors, and we call $\|\boldsymbol{y} - A\boldsymbol{x}\|$ the least squares error.

$$A\boldsymbol{x} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{pmatrix}$$

The term "***least square solution***" results from the fact the minimizing $\|\boldsymbol{y} - A\boldsymbol{x}\| = e_1^2 + e_2^2 + \ldots + e_m^2$

Example

Find the sums of squares of the errors of (2, 4), (4, 8), (6, 6)

Solution

$$4 = 2m + b \quad \Rightarrow \quad 4 - 2m - b = e_1$$
$$8 = 4m + b \quad \Rightarrow \quad 8 - 4m - b = e_2$$
$$6 = 6m + b \quad \Rightarrow \quad 6 - 6m - b = e_3$$
$$e_1^2 + e_2^2 + \ldots + e_m^2 = (4 - 2m - b)^2 + (8 - 4m - b)^2 + (6 - 6m - b)^2$$

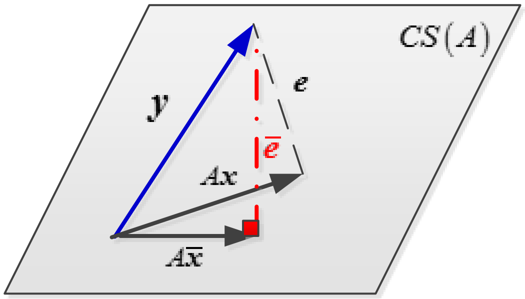
The least squares problem for this example to find the values m and b for which $e_1^2 + e_2^2 + \ldots + e_m^2$ is a minimum.

Theorem

If A is an $m \times n$ matrix, the equation $A\boldsymbol{x} = \boldsymbol{y}$ has a solution if and only if y is in the column space of A.

$$\boldsymbol{y} - A\boldsymbol{x} = \boldsymbol{e}$$

$A\boldsymbol{x}$ is a vector that is in the column space of A. For this A the column space is a plane is $\boldsymbol{R}^m$
$\boldsymbol{y}$ is a vector, not in the column space of A (otherwise $A\boldsymbol{x} = \boldsymbol{y}$ has an exact solution)
$\boldsymbol{e}$ is the error vector, the difference between $\boldsymbol{y}$ and $A\boldsymbol{x}$



The length $\|\boldsymbol{e}\|$ is a minimum exactly when $\boldsymbol{e} \perp CS(A)$

Best Approximation *Theorem*

If $CS(A)$ is a finite dimensional subspace of an inner product space, and if $\boldsymbol{y}$ is a vector in $\boldsymbol{V}$, then $proj_{CS(A)}\boldsymbol{y}$ is the best approximation to $\boldsymbol{y}$ from $CS(A)$ is the sense that

$$\left\| \boldsymbol{y} - proj_{CS(A)}\boldsymbol{y} \right\| < \left\| \boldsymbol{y} - CS(A) \right\|$$

For every vector $\boldsymbol{w}$ in $CS(A)$ that is different from $proj_{CS(A)}\boldsymbol{y}$

Theorem

For every linear system $A\boldsymbol{x} = \boldsymbol{y}$, the associated normal system

$$A^T A\boldsymbol{x} = A^T \boldsymbol{y}$$

Is consistent, and all solutions are least squares solutions of $A\boldsymbol{x} = \boldsymbol{y}$

If the columns of A are linearly independent, then $A^T A$ is invertible so has a unique solution $\overline{\boldsymbol{x}}$. This solution is often expressed theoretically as

$$\left(A^T A\right)^{-1} A^T A\overline{\boldsymbol{x}} = \left(A^T A\right)^{-1} A^T \boldsymbol{y}$$

$$\overline{\boldsymbol{x}} = \left(A^T A\right)^{-1} A^T \boldsymbol{y}$$

Proof

Let the vector $\overline{\boldsymbol{x}}$ is a least squares solution to $A\boldsymbol{x} = \boldsymbol{y} \quad \Leftrightarrow \quad (\boldsymbol{y} - A\overline{\boldsymbol{x}}) \perp CS(A)$

$$(\boldsymbol{y} - A\overline{\boldsymbol{x}}) \cdot \boldsymbol{z} = 0 \qquad \boldsymbol{z} \text{ in } CS(A) \quad \& \quad \boldsymbol{z} = A\boldsymbol{w}$$

$$(\boldsymbol{y} - A\overline{\boldsymbol{x}}) \cdot A\boldsymbol{w} = 0 \qquad \boldsymbol{w} \text{ in } \boldsymbol{R}^n$$

$$A^T(\boldsymbol{y} - A\overline{\boldsymbol{x}}) \cdot \boldsymbol{w} = 0$$

$$A^T(\boldsymbol{y} - A\overline{\boldsymbol{x}}) = 0$$

$$A^T\boldsymbol{y} - A^T A\overline{\boldsymbol{x}} = 0$$

$$A^T\boldsymbol{y} = A^T A\overline{\boldsymbol{x}}$$

Theorem

If A is an $m \times n$ matrix, then the following are equivalent

a) A has linearly independent column vectors.

b) $A^T A$ is invertible.

Example

Find the equation of the line that best fits the given points in the least-squares sense.

(40, 482), (45, 467), (50, 452), (55, 432), (60, 421)

Solution

Let $y = mx + b$ be the equation of the line that best fits the given points. Then

$$\begin{pmatrix} 40 & 1 \\ 45 & 1 \\ 50 & 1 \\ 55 & 1 \\ 60 & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 482 \\ 467 \\ 452 \\ 432 \\ 421 \end{pmatrix}$$

Where $A = \begin{pmatrix} 40 & 1 \\ 45 & 1 \\ 50 & 1 \\ 55 & 1 \\ 60 & 1 \end{pmatrix} \quad \boldsymbol{x} = \begin{pmatrix} m \\ b \end{pmatrix} \quad \boldsymbol{y} = \begin{pmatrix} 482 \\ 467 \\ 452 \\ 432 \\ 421 \end{pmatrix}$

Using the normal equation formula: $A^T Ax = A^T y$

$$\begin{pmatrix} 40 & 45 & 50 & 55 & 60 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 40 & 1 \\ 45 & 1 \\ 50 & 1 \\ 55 & 1 \\ 60 & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 40 & 45 & 50 & 55 & 60 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 482 \\ 467 \\ 452 \\ 432 \\ 421 \end{pmatrix}$$

$$\begin{pmatrix} 12,750 & 250 \\ 250 & 5 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 111,970 \\ 2,255 \end{pmatrix}$$

$X = A^{-1}B$

$$\begin{pmatrix} m \\ b \end{pmatrix} = \frac{1}{1250} \begin{pmatrix} 5 & -250 \\ -250 & 12,750 \end{pmatrix} \begin{pmatrix} 111,970 \\ 2,255 \end{pmatrix}$$

$$= \begin{pmatrix} -3.12 \\ 607 \end{pmatrix}$$

Or

$$\begin{pmatrix} 12,750 & 250 & 111,970 \\ 250 & 5 & 2,225 \end{pmatrix} \xrightarrow{rref} \begin{pmatrix} 1 & 0 & -3.12 \\ 0 & 1 & 607 \end{pmatrix}$$

Thus $\boxed{y = -3.12x + 607}$

Example

Given the system equation: $\begin{cases} x_1 - x_2 = 4 \\ 3x_1 + 2x_2 = 1 \\ -2x_1 + 4x_2 = 3 \end{cases}$

a) Find the least-squares solution of the linear system $A\boldsymbol{x} = \boldsymbol{y}$

b) Find the orthogonal projection of $\boldsymbol{y}$ on the column space of A

c) Find the error vector and the error

Solution

a) $A = \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix} \quad \boldsymbol{x} = \begin{pmatrix} m \\ b \end{pmatrix} \quad \boldsymbol{y} = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}$

$$A^T A\boldsymbol{x} = A^T \boldsymbol{y}$$

$$\begin{pmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 1 & 3 & -2 \\ -1 & 2 & 4 \end{pmatrix}\begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 14 & -3 \\ -3 & 21 \end{pmatrix}\begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 10 \end{pmatrix}$$

$$\begin{pmatrix} m \\ b \end{pmatrix} = \frac{1}{285}\begin{pmatrix} 21 & 3 \\ 3 & 14 \end{pmatrix}\begin{pmatrix} 1 \\ 10 \end{pmatrix} = \begin{pmatrix} \frac{51}{285} \\ \frac{143}{285} \end{pmatrix} \qquad X = A^{-1}B$$

$$= \begin{pmatrix} \frac{17}{95} \\ \frac{143}{285} \end{pmatrix}$$

Thus $\boxed{y = 0.1789x + 0.5018}$

b) The orthogonal projection of $\boldsymbol{y}$ on the column space of A

$$A\boldsymbol{x} = \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} \frac{17}{95} \\ \frac{143}{285} \end{pmatrix} = \begin{pmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{94}{57} \end{pmatrix}$$

c) $\boldsymbol{y} - A\boldsymbol{x} = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} - \begin{pmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{94}{57} \end{pmatrix} = \begin{pmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{4}{3} \end{pmatrix}$

The error: $\|\boldsymbol{y} - A\boldsymbol{x}\| = \sqrt{\left(\frac{1232}{285}\right)^2 + \left(-\frac{154}{285}\right)^2 + \left(\frac{4}{3}\right)^2} \boxed{\approx 4.556}$

Exercises Section 3.5 – Least Squares Analysis

1. Find the equation of the line that best fits the given points in the least-squares sense.

a) $\{(0,\ 2),\ (1,\ 2),\ (2,\ 0)\}$

b) $\{(1,\ 5),\ (2,\ 4),\ (3,\ 1),\ (4,\ 1),\ (5,\ -1)\}$

c) $\{(0,\ 1),\ (1,\ 3),\ (2,\ 4),\ (3,\ 4)\}$

d) $\{(-2,\ 0),\ (-1,\ 0),\ (0,\ 1),\ (1,\ 3),\ (2,\ 5)\}$

2. Find the orthogonal projection of the vector $\boldsymbol{u}$ on the subspace of $\boldsymbol{R}^4$ spanned by the vectors

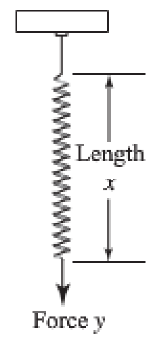
a) $\boldsymbol{u}=(-3,\ -3,\ 8,\ 9);\quad \boldsymbol{v}_1=(3,\ 1,\ 0,\ 1),\quad \boldsymbol{v}_2=(1,\ 2,\ 1,\ 1),\quad \boldsymbol{v}_3=(-1,\ 0,\ 2,\ -1)$

b) $\boldsymbol{u}=(6,\ 3,\ 9,\ 6);\quad \boldsymbol{v}_1=(2,\ 1,\ 1,\ 1),\quad \boldsymbol{v}_2=(1,\ 0,\ 1,\ 1),\quad \boldsymbol{v}_3=(-2,\ -1,\ 0,\ -1)$

c) $\boldsymbol{u}=(-2,\ 0,\ 2,\ 4);\quad \boldsymbol{v}_1=(1,\ 1,\ 3,\ 0),\quad \boldsymbol{v}_2=(-2,\ -1,\ -2,\ 1),\quad \boldsymbol{v}_3=(-3,\ -1,\ 1,\ 3)$

3. Find the standard matrix for the orthogonal projection P of $\boldsymbol{R}^2$ on the line passes through the origin and makes an angle θ with the positive x-axis.

4. Hooke's law in physics states that the length x of a uniform spring is a linear function of the force y applied to it. If we express the relationship as $y = mx + b$, then the coefficient m is called the spring constant. Suppose a particular unstretched spring has a measured length of 6.1 *inches*.(i.e., x = 6.1 when y = 0). Forces of 2 pounds, 4 pounds, and 6 pounds are then applied to the spring, and the corresponding lengths are found to be 7.6 inches, 8.7 inches, and 10.4 inches. Find the spring constant.



5. Prove: If A has a linearly independent column vectors, and if $\boldsymbol{b}$ is orthogonal to the column space of A, then the least squares solution of $A\boldsymbol{x}=\boldsymbol{b}$ is $\boldsymbol{x}=0$.

6. Let A be an $m\times n$ matrix with linearly independent row vectors. Find a standard matrix for the orthogonal projection of R^n onto the row space of A.

7. Let W be the line with parametric equations $\quad x=2t,\quad t=-t,\quad z=4t$

a) Find a basis for W.

b) Find the standard matrix for the orthogonal projection on W.

c) Use the matrix in part (b) to find the orthogonal projection of a point $P_0\left(x_0,\ y_0,\ z_0\right)$ on W.

d) Find the distance between the point $P_0(2,\ 1,\ -3)$ and the line W.

8. In R^3, consider the line l given by the equations $\quad x=t,\quad t=t,\quad z=t$

And the line m given by the equations $\quad x=s,\quad t=2s-1,\quad z=1$

Let P be the point on l, and let Q be a point on m. Find the values of t and s that minimize the distance between the lines by minimizing the squared distance $\|P-Q\|^2$

9. Determine whether the statement is true or false,

a) If A is an $m \times n$ matrix, then $A^T A$ is a square matrix.

b) If $A^T A$ is invertible, then A is invertible.

c) If A is invertible, then $A^T A$ is invertible.

d) If $A\boldsymbol{x} = \boldsymbol{b}$ is a consistent linear system, then $A^T A\boldsymbol{x} = A^T \boldsymbol{b}$ is also consistent.

e) If $A\boldsymbol{x} = \boldsymbol{b}$ is an inconsistent linear system, then $A^T A\boldsymbol{x} = A^T \boldsymbol{b}$ is also inconsistent.

f) Every linear system has a least squares solution.

g) Every linear system has a unique least squares solution.

h) If A is an $m \times n$ matrix with linearly independent columns and $\boldsymbol{b}$ is in R^m, then $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least squares solution.