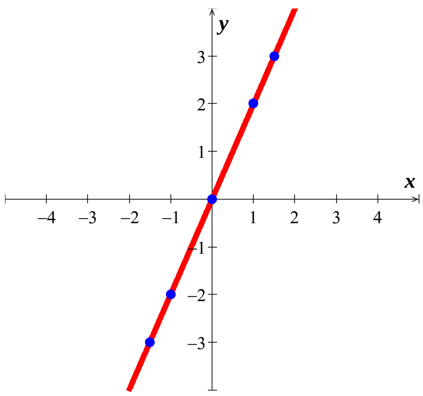
Section 3.5 – Least Squares Analysis

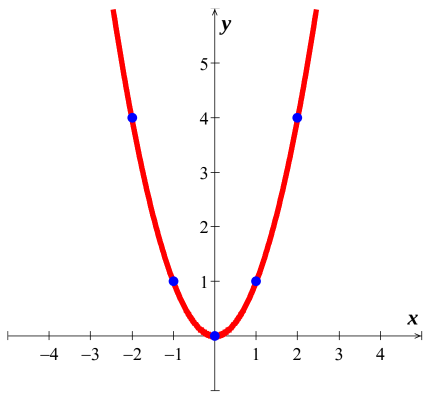
The use to ***best*** fit data, we will use results about orthogonal projections in inner product spaces to obtain a technique for fitting a line or other polynomial.

Fitting a Curve to Data

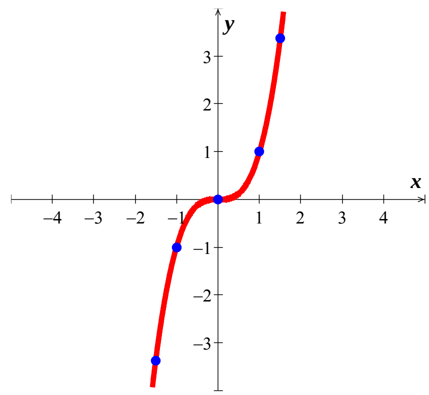
The common problem is to obtain a mathematical relationship between 2 variables x and y by ***fitting*** a curve to points in the xy-plane.
Some possibility of fitting the data



$y = ax + b$



$y = ax^2 + bx + c$



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

Least Squares Fit of a Straight Line

Recall that a system of equations $A\vec{x} = \vec{y}$ 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 $(x_1, y_1), \ldots, (x_n, y_n)$

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

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

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

Least Square Problem

Given a linear system $A\vec{x} = \vec{y}$ of m equations in n unknowns, find a vector $\vec{x}$ that minimizes $\|\vec{y} - A\vec{x}\|$ with respect to the Euclidean inner product on $\mathbb{R}^m$. We call such as $\vec{x}$ a least squares solution of the system, we call $\vec{y} - A\vec{x}$ the least squares error vectors, and we call $\|\vec{y} - A\vec{x}\|$ 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 $\|\vec{y} - A\vec{x}\| = e_1^2 + e_2^2 + \ldots + e_m^2$

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

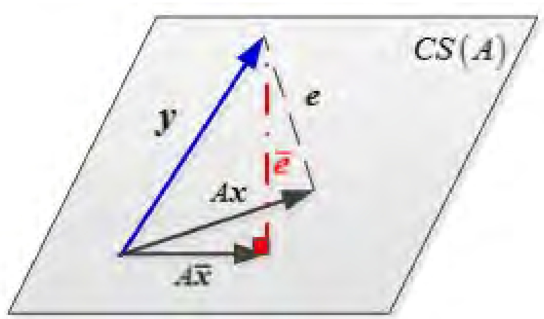
If *A* is an $m \times n$ matrix, the equation $A\vec{x} = \vec{y}$ has a solution if and only if $\vec{y}$ is in the column space of *A*.

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

$A\vec{x}$ is a vector that is in the column space of *A*. For this *A* the column space is a plane is $\mathbb{R}^m$

$\vec{y}$ is a vector, not in the column space of *A* (otherwise $A\vec{x} = \vec{y}$ has an exact solution)

$\vec{e}$ is the error vector, the difference between $\vec{y}$ and $A\vec{x}$



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

Best Approximation *Theorem*

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

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

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

Theorem

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

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

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

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

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

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

Proof

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

$$\left(\vec{y} - A\bar{x}\right) \cdot \vec{\boldsymbol{z}} = 0 \qquad\qquad \vec{\boldsymbol{z}}\ in\ CS\left(A\right) \quad \& \quad \vec{\boldsymbol{z}} = A\vec{w}$$

$$\left(\vec{y} - A\bar{x}\right) \cdot A\vec{w} = 0 \qquad\qquad \vec{w}\ in\ \mathbb{R}^n$$

$$A^T\left(\vec{y} - A\bar{x}\right) \cdot \vec{w} = 0$$

$$A^T\left(\vec{y} - A\bar{\boldsymbol{x}}\right) = 0$$

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

$$A^T\vec{y} = A^T A\bar{x}$$

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

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

$$= \frac{-3,900}{1,250}$$

$$\boxed{= -\frac{78}{25}}$$

$$b = \frac{\begin{vmatrix} 12,750 & 111,970 \\ 250 & 2,255 \end{vmatrix}}{1,250}$$

$$= \frac{758,750}{1,250}$$

$$\boxed{= 607}$$

Thus, $\boxed{y = -\frac{78}{25}x + 607}$ *or* $\boxed{y = -3.12x + 607}$

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 $A\vec{x} = \vec{y}$

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

c) Find the ***error vector*** and the ***error***

Solution

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

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

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

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

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

$$= \begin{pmatrix} \frac{51}{285} \\ \frac{143}{285} \end{pmatrix}$$

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

Thus $\boxed{y = \frac{17}{95}x + \frac{143}{285}}$ *or* $\boxed{y = 0.1789x + 0.5018}$

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

$$A\vec{x} = \begin{pmatrix} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} \frac{17}{95} \\ \frac{143}{285} \end{pmatrix}$$

$$=\begin{pmatrix} -\frac{92}{285} \\ \frac{439}{285} \\ \frac{94}{57} \end{pmatrix}$$

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

$$=\begin{pmatrix} \frac{1232}{285} \\ -\frac{154}{285} \\ \frac{4}{3} \end{pmatrix}$$

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

$$\underline{\approx 4.556}$$

Exercises *Section* 3.5 – Least Squares Analysis

(1 – 7) Find the equation of the line that best fits the given points in the least-squares sense and find the error.

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

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

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

4. $\{(-2,\ 0),\ (-1,\ 0),\ (0,\ 1),\ (1,\ 3),\ (2,\ 5)\}$

5. $\{(2,\ 3),\ (3,\ 2),\ (5,\ 1),\ (6,\ 0)\}$

6. $\{(-1,\ 0),\ (0,\ 1),\ (1,\ 2),\ (2,\ 4)\}$

7. $\{(1,\ 0),\ (2,\ 1),\ (4,\ 2),\ (5,\ 3)\}$

(8 – 10) Find the orthogonal projection of the vector $\vec{u}$ on the subspace of $\mathbb{R}^4$ spanned by the vectors

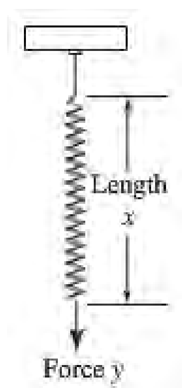
8. $\vec{u}=(-3,\ -3,\ 8,\ 9);\quad \vec{v}_1=(3,\ 1,\ 0,\ 1),\quad \vec{v}_2=(1,\ 2,\ 1,\ 1),\quad \vec{v}_3=(-1,\ 0,\ 2,\ -1)$

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

10. $\vec{u}=(-2,\ 0,\ 2,\ 4);\quad v_1=(1,\ 1,\ 3,\ 0),\quad \vec{v}_2=(-2,\ -1,\ -2,\ 1),\quad \vec{v}_3=(-3,\ -1,\ 1,\ 3)$

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

12. 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.

13. 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 $A\vec{x}=\vec{b}$ is $\vec{x}=\vec{0}$.

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

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

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 $P_0\left(2, 1, -3\right)$ and the line W.

16. In $\mathbb{R}^3$, consider the line l given by the equations $x=t, \quad y=t, \quad z=t$

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

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 $\|P-Q\|^2$

17. 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 $A\vec{x}=\vec{b}$ is a consistent linear system, then $A^T A\vec{x}=A^T\vec{b}$ is also consistent.
e) If $A\vec{x}=\vec{b}$ is an inconsistent linear system, then $A^T A\vec{x}=A^T\vec{b}$ 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 $\vec{b}$ is in R^m, then $A\vec{x}=\vec{b}$ has a unique least squares solution.

18. A certain experiment produces the data $\{(1, 1.8), (2, 2.7), (3, 3.4), (4, 3.8), (5, 3.9)\}$.

Find the function that it will fit these data in the form of $y=\beta_1 x+\beta_2 x^2$

19. According to Kepler's first law, a comet should have an ellipse, parabolic, or hyperbolic orbit (with gravitational attractions from the planets ignored). In suitable polar coordinates, the position $(r,\ \upsilon)$ of a comet satisfies an equation of the form

$$r = \beta + e(r \cdot \cos\upsilon)$$

Where β is a constant and e is the eccentricity of the orbit, with $0 \le e < 1$ for an ellipse, $e = 1$ for a parabolic, and $e > 1$ for a hyperbola.

Suppose observations of a newly discovered comet provide the data below.

υ	.88	1.10	1.42	1.77	2.14
r	3.00	2.30	1.65	1.25	1.01

Determine the type of orbit, and predict where the orbit will be when $\upsilon = 4.6$ $(\textit{radians})$?

20. To measure the takeoff performance of an airplane, the horizontal position of the plane was measured every second, from $t = 0$ to $t = 12$

The position (in *feet*) were:

$0,\ 8.8,\ 29.9,\ 62.0,\ 104.7,\ 159.1,\ 222.0,\ 294.5,\ 380.4,\ 471.1,\ 571.7,\ 686.8,$ *and* 809.2

a) Find the least square cubic curve $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ for these data.

b) Estimate the velocity of the plane when $t = 4.5$ *sec*, using the result from part (*a*).