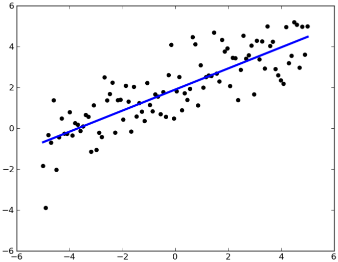
Section 6.5

Least Squares Problems

Least Squares problems

What if we can't solve $Ax = b$? How can we solve it as closely as possible?



To solve $Ax = b$ as closely as possible, we orthogonally project b onto $\mathrm{Col}(A)$; call the result $\widehat{b}$. Then solve $Ax = \widehat{b}$. This is the *least squares solution* to $Ax = b$.

Outline of Section 6.5

- The method of least squares
- Application to best fit lines/planes
- Application to best fit curves

Least squares solutions

$A = m \times n$ matrix.

A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b.

The error is $\|A\widehat{x} - b\|$.

Demo

Least squares solutions

A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b.

The error is $\|A\widehat{x} - b\|$.

Theorem. The least squares solutions to $Ax = b$ are the solutions to

$$(A^T A)x = (A^T b)$$

So this is just like what we did before when we were finding the projection of b onto $\text{Col}(A)$. But now we just solve and don't (necessarily) multiply the solution by A.

Least squares solutions

Example

Theorem. The least squares solutions to $Ax = b$ are the solutions to

$$(A^T A)x = (A^T b)$$

Find the least squares solutions to $Ax = b$ for this A and b:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

What is the error?

Least squares solutions

Example

Formula: $(A^T A)x = (A^T b)$

Find the least squares solution/error to $Ax = b$:

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$$

Least squares solutions

Theorem. Let A be an $m \times n$ matrix. The following are equivalent:

1. $Ax = b$ has a unique least squares solution for all b in $\mathbb{R}^n$
2. The columns of A are linearly independent
3. $A^T A$ is invertible

In this case the least squares solution is $(A^T A)^{-1}(A^T b)$.

Application

Best fit lines

Problem. Find the best-fit line through $(0, 6)$, $(1, 0)$, and $(2, 0)$.

Demo

Best fit lines

Poll

What does the best fit line minimize?

1. the sum of the squares of the distances from the data points to the line
2. the sum of the squares of the vertical distances from the data points to the line
3. the sum of the squares of the horizontal distances from the data points to the line
4. the maximal distance from the data points to the line

Least Squares Problems

More applications

Determine the least squares problem $Ax = b$ to find the best parabola $y = Cx^2 + Dx + E$ for the points:

$$(0,0),(2,0),(3,0),(0,1)$$

Demo

Least Squares Problems

More applications

Determine the least squares problem $Ax = b$ to find the best fit ellipse $Cx^2 + Dxy + Ey^2 + Fx + Gy + H = 0$ for the points:

$$(0,0), (2,0), (3,0), (0,1)$$

Gauss invented the method of least squares to predict the orbit of the asteroid Ceres as it passed behind the sun in 1801.

Least Squares Problems

Best fit plane

Determine the least squares problem $Ax = b$ to find the best fit linear function $f(x, y) = Cx + Dy + E$

x	y	$f(x,y)$
1	0	0
0	1	1
-1	0	3
0	-1	4

Summary of Section 6.5

- A least squares solution to $Ax = b$ is an $\widehat{x}$ in $\mathbb{R}^n$ so that $A\widehat{x}$ is as close as possible to b.
- The error is $\|A\widehat{x} - b\|$.
- The least squares solutions to $Ax = b$ are the solutions to $(A^T A)x = (A^T b)$.
- To find a best fit line/parabola/etc. write the general form of the line/parabola/etc. with unknown coefficients and plug in the given points to get a system of linear equations in the unknown coefficients.

Typical Exam Questions 6.5

- Find the best fit line through $(1,0)$, $(2,1)$, and $(3,1)$. What is the error?
- Find the best fit parabola through $(1,0)$, $(2,1)$, $(3,1)$, and $(3,0)$. What is the error?
- True/false. For every set of three points in $\mathbb{R}^2$ there is a unique best fit line.
- True/false. If $\widehat{x}$ is the least squares solution to $Ax = b$ for an $m \times n$ matrix A, then $\widehat{x}$ is the closest point in $\mathbb{R}^n$ to b.
- True/false. If $\widehat{x}$ and $\widehat{y}$ are both least squares solutions to $Ax = b$ then $\widehat{x} - \widehat{y}$ is in the null space of A.