INTRODUCTION TO LEAST SQUARES

David F. Gleich · CS 314 · Purdue University

October 4, 2016

1 THE 2-NORM

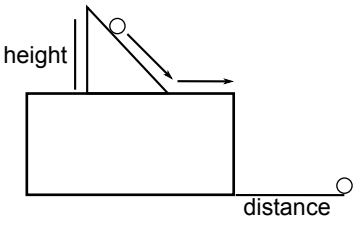
Also called the Euclidean norm

$$\|\mathbf{x}\| = \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

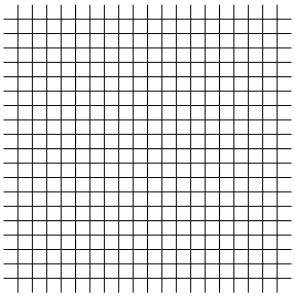
2 LEAST SQUARES PROBLEMS

FITTING DATA

Galileo wanted to find a mathematical relationship between ball height and horizontal distance in the following experiment.



height	distance
0.282	0.752
0.564	1.102
0.752	1.248
0.940	1.410



RANKING TEAMS

Suppose we have four college football teams

1. Purdue
2. IU
3. Notre Dame
4. Michigan

Purdue beats Notre Dame by 4 points: 21 to 17
Michigan beats Purdue by 9 points: 27 to 18
Purdue beats IU by 6 points: 16 to 10
Michigan betas IU by 3 points: 10 to 7
Notre Dame beats IU by 7 points: 17 to 10

Suppose we give Purdue 100 ranking points, how many points should the other teams get to predict the score differentials?

3 THE LEAST SQUARES PROBLEM

Find a vector $\mathbf{x}$ that minimizes $\|\boldsymbol{A}\mathbf{x} - \mathbf{b}\|_2^2$

"like" $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ but with too many equations.

FITTING DATA

$\boldsymbol{A} =$

$\mathbf{b} =$

RANKING TEAMS

$\boldsymbol{A} =$

$\mathbf{b} =$

4 A GEOMETRIC EXAMPLE

$$\boldsymbol{A} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

"$x = 3$"

"$2x = 4$"

5 THE NORMAL EQUATIONS

see the book $\quad \boldsymbol{A}^T\boldsymbol{A}\mathbf{x} = \boldsymbol{A}^T\mathbf{b}$

6 THE QR FACTORIZATION

orthogonal matrix $\Leftrightarrow$ orthonormal column vectors

orthogonal $\boldsymbol{Q} \Leftrightarrow \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$