Math 415 - Lecture 24

Least squares

Wednesday October 21st 2015

Textbook reading: Chapter 3.3

Suggested practice exercises: Exercises 3, 5, 6, 13, 24, 25

Khan Academy video: Least Squares Approximation, Least Squares Examples, Another Least Squares Example

1 Review

Let W be a subspace of $\mathbb{R}^n$ and $\mathbf{x}$ in $\mathbb{R}^n$ (but maybe not in W). Let $\mathbf{x}_W$ be the orthonormal projection of $\mathbf{x}$ onto W. (vector in W as close as possible to $\mathbf{x}$)

- If $\mathbf{v}_1, \cdots, \mathbf{v}_m$ is an orthogonal basis of W then

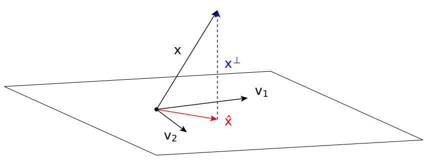
$$\mathbf{x}_W = \underbrace{\left(\frac{\mathbf{x} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v_1}}_{\text{proj. of } \mathbf{x} \text{ onto } \mathbf{v}_1} + \cdots + \underbrace{\left(\frac{\mathbf{x} \cdot \mathbf{v}_m}{\mathbf{v}_m \cdot \mathbf{v}_m}\right)\mathbf{v_m}}_{\text{proj. of } \mathbf{x} \text{ onto } \mathbf{v}_m}.$$

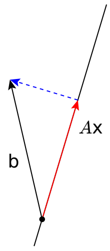
- The decomposition $\mathbf{x} = \underbrace{\mathbf{x}_W}_{\text{in } W} + \underbrace{\mathbf{x}^\perp}_{\text{in } W^\perp}$ is unique.

2 Least squares

2.1 Least squares

Definition. $\hat{\mathbf{x}}$ is a **least squares solution** of the system $A\mathbf{x} = \mathbf{b}$ if $\hat{\mathbf{x}}$ is such that $A\hat{\mathbf{x}} - \mathbf{b}$ is as small as possible.





- If $A\mathbf{x} = \mathbf{b}$ is consistent, then a least squares solution $\hat{\mathbf{x}}$ is just an ordinary solution. (in that case, $A\hat{\mathbf{x}} - \mathbf{b} = 0$)
- Interesting case: $A\mathbf{x} = \mathbf{b}$ is inconsistent. (in other words: the system is overdetermined)

Idea. $A\mathbf{x} = \mathbf{b}$ is consistent $\iff$ $\mathbf{b}$ is in $\mathrm{Col}(A)$

So if $A\mathbf{x} = \mathbf{b}$ is inconsistent we

- replace $\mathbf{b}$ with its projection $\hat{\mathbf{b}}$ onto $\mathrm{Col}(A)$,
- and solve $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$.(consistent by construction!)

Example 1. Find the least squares solution to $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

Solution.

3 The normal equations

Theorem 1. $\hat{\mathbf{x}}$ *is a least squares solution of* $A\mathbf{x} = \mathbf{b} \iff A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$

Proof. $\square$

Example 2 (again). Find the least squares solution to $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}.$$

Solution.

Example 3. Find the least squares solution to $A\mathbf{x} = \mathbf{b}$, where

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

What is the projection of $\mathbf{b}$ onto $\mathrm{Col}(A)$?

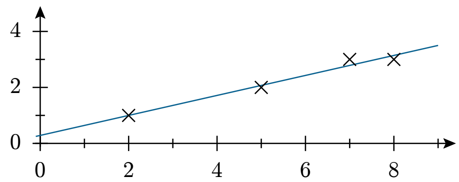
Note that the columns of A are not orthogonal.

Solution.

4 Applications

4.1 Least square regression lines

Experimental data: (x_i, y_i), for $i = 1, 2, 3, \ldots..$ Wanted: parameters β_1, β_2 such that $y_i \approx \beta_1 + \beta_2 x_i$ for all i



The approximation should be so that

$$\mathrm{SS}_{\mathrm{res}} = \underbrace{\sum_i [y_i - (\beta_1 + \beta_2 x_i)]^2}_{\text{residue sum of squares}} \text{ is as small as possible.}$$

Example 4. Find β_1, β_2 such that the line $y = \beta_1 + \beta_2 x$ best fits the data points $(2, 1), (5, 2), (7, 3), (8, 3)$.

Solution.

Example 5. Blood is drawn form volunteers to determine the effects of a new experimental drug designed to lower cholesterol levels, The following data shows the results of varying the dosage from 0 unit to 1 units in step of 0.2 of a unit. Find a line $C = \beta_1 D + \beta_2$ that best fits the data. What drug usage would you recommend if you want to accomplish a Cholesterol level of 215?

Drug Dosage: D	0.0	0.2	0.4	0.6	0.8	1
Cholesterol: C	289	273	254	226	213	189

Solution.