Math 265
Professor Priyam Patel
4/5/16

Class Handout #16

Least Squares Approximation

Our goal: When $A\mathbf{x} = \mathbf{b}$ is inconsistent, find the closest thing we can to a solution, that is find an $\widehat{\mathbf{x}} \in \mathbb{R}^n$ such that $A\widehat{\mathbf{x}}$ is as close as possible to $\mathbf{b}$.

We find all *least squares solutions* $\widehat{\mathbf{x}}$ by solving the *normal system* $A^T A\widehat{\mathbf{x}} = A^T\mathbf{b}$.

Theorem 5.14 When rank $A = n$, the least squares solution $\widehat{\mathbf{x}}$ to the normal system is unique and $A\widehat{\mathbf{x}} = \text{proj}_{\text{Col}A}\mathbf{b}$.

Exercise 1: (Set up the following problem.) Find the least squares solution to $A\mathbf{x} = \mathbf{b}$ where

$$A = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 1 & 1 & 2 \\ -2 & 3 & 4 & 1 \\ 4 & 2 & 1 & 0 \\ 0 & 2 & 1 & 3 \\ 1 & -1 & 2 & 0 \end{bmatrix} \text{ and } \mathbf{b} = \begin{bmatrix} 1 \\ 5 \\ -2 \\ 1 \\ 3 \\ 5 \end{bmatrix}.$$

Most common application of least squares approximation:

You are given a data set and need to find the best fit line, parabola, function. In general, we are looking for the coefficients $x_1, x_2, \ldots, x_n$ so the $y(t) = x_1 f_1(t) + x_2 f_2(t) + \cdots + x_n f_n(t)$ is the function that best fits the data set.

Example: The following data show atmospheric pollutants y_i at half hour intervals t_i:

t_i	1	1.5	2	2.5	3	3.5	4	4.5	5
y_i	-0.15	0.24	0.68	1.04	1.21	1.15	0.86	0.41	-0.08

Exercise 1: For the data above, set up an inconsistent system $A\mathbf{x} = \mathbf{y}$ for which you would like to find the least squares solution.

Example: In the manufacturing of a product Z, the amount of compound A present depends on the amount of ingredient B used in the refining process. The following data was obtained:

B used	2	4	6	8	10
A present	3.5	8.2	10.5	12.9	14.6

Exercise 2: For the data above, set up an inconsistent system $A\mathbf{x} = \mathbf{y}$ for which you would like to find the least squares solution.

Section 6.1:

Definition: Let V and W be vector spaces. A function $L : V \longrightarrow W$ is called a *linear transformation* of V into W if:

1. $L(\mathbf{u}+\mathbf{v}) = L(\mathbf{u}) + L(\mathbf{v})$ for any $\mathbf{u}$ and $\mathbf{v}$ in V.
2. $L(c\mathbf{u}) = cL(\mathbf{u})$ for any $\mathbf{u}$ in V and any scalar c.

Exercise 3: Let $L : \mathbb{R}_3 \longrightarrow \mathbb{R}_3$ be defined by $L\left(\begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}\right) = \begin{bmatrix} 2u_1 & 2u_2 & 2u_3 \end{bmatrix}$. Is L a linear transformation?

Exercise 4: Let $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be defined by $L\left(\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}\right) = \begin{bmatrix} u_1 + 1 \\ 2u_2 \\ u_3 \end{bmatrix}$. Is L a linear transformation?

The nice thing about linear transformations is that once you know $L(\mathbf{u})$ and $L(\mathbf{v})$ you know how L transforms any linear combination of $\mathbf{u}$ and $\mathbf{v}$. For example, if $L(\mathbf{u}) = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ and $L(\mathbf{v}) = \begin{bmatrix} -3 \\ 1 \\ 5 \end{bmatrix}$, what is $L(3\mathbf{u} - 2\mathbf{v})$?

What this means is that is $L : V \longrightarrow W$ is a linear transformation, and $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is a basis for V, then once we know $L(\mathbf{v}_1), \ldots, L(\mathbf{v}_n)$, we know $L(\mathbf{v})$ for any $\mathbf{v} \in V$.

Let's examine a special case of this. Let $L : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. Then for a random vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$, we know that $L\left(\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}\right) = v_1 L\left(\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) + v_2 L\left(\begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}\right) + \cdots + v_n L\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}\right)$.

This means $L(\mathbf{v}) = A\mathbf{v}$ where $A =$

The matrix A above is called the *standard matrix* for L.

Exercise 5: Let $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ be the linear transformation defined by $L\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 \\ 3x_2 - 2x_3 \end{bmatrix}$. Find the standard matrix A for L.

Exercise 6: Let $L : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ be the linear transformation defined by $L(\mathbf{u}) = 5\mathbf{u}$. Find the standard matrix A for L.