

Problem 1 (15 points; 3 points each). Decide if each of the following are true or false and provide a justification or counterexample in each case. A justification could consist of a theorem from the text. All vector spaces are assumed to be finite-dimensional here.

- (a) _____ There is a unique least squares solution to $A\mathbf{x} = \mathbf{b}$.

FALSE: This was one of your DQs. You showed that the set of least square solutions to $A\mathbf{x} = \mathbf{b}$ is exactly $\hat{\mathbf{x}} + \text{NS}(A)$, where $\hat{\mathbf{x}}$ is any fixed least squares solution, that is, $\hat{\mathbf{x}}$ satisfies, $A^T A \hat{\mathbf{x}} = A^T \mathbf{b}$.

In general, there is no reason for $A^T A$ to be invertible, so $(A^T A)^{-1}$ need not even exist. If it does exist, then $\hat{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b}$ is unique.

What you do know is that there is a unique $\hat{\mathbf{b}}$ so that $\hat{\mathbf{b}}$ is the closest thing of the form $A\mathbf{x}$ to \mathbf{b} , in other words, $\|\hat{\mathbf{b}} - \mathbf{b}\|_2^2 = \min\{\|A\mathbf{x} - \mathbf{b}\|_2^2 \mid \mathbf{x} \in \mathbb{R}^n\}$ and a least-square solution is a solution to $A\mathbf{x} = \hat{\mathbf{b}}$.

- (b) _____ There is a unique \mathbf{y} so that $\|\mathbf{y} - \mathbf{b}\|$ is minimal and $A\mathbf{x} = \mathbf{y}$.

TRUE: This is the conclusion of the main theorem about the existence of "least-square" solutions. This is covered in the notes and in the book.

- (c) _____ If $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal basis for V with respect to an inner product $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ and $\mathbf{v} = \sum_{i=1}^n \alpha_i \mathbf{u}_i$, then $\|\mathbf{v}\|_2^2 = \sum_{i=1}^n |\alpha_i|^2$.

This is true and is essentially the Pythagorean Theorem. This is a computation

$$\begin{aligned} \|\mathbf{v}\|_2^2 &= \langle \mathbf{v}, \mathbf{v} \rangle = \left\langle \sum_{i=1}^n \alpha_i \mathbf{u}_i, \sum_{i=1}^n \alpha_i \mathbf{u}_i \right\rangle \\ &= \sum_{i=1}^n \alpha_i \left\langle \mathbf{u}_i, \sum_{i=1}^n \alpha_i \mathbf{u}_i \right\rangle \\ &= \sum_{i=1}^n \alpha_i \left(\sum_{j=1}^n \bar{\alpha}_j \langle \mathbf{u}_i, \mathbf{u}_j \rangle \right) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \bar{\alpha}_j \delta_{i,j} \\ &= \sum_{i=1}^n \alpha_i \bar{\alpha}_i = \sum_{i=1}^n |\alpha_i|^2 \end{aligned}$$

Here

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

- (d) _____ All norms $\|\cdot\| : \mathbb{R}^n \rightarrow [0, \infty)$ on \mathbb{R}^n come from an inner product.

This is false. The book provides several norms. For a norm $\|\cdot\|$ to be given by an inner product it must satisfy the parallelogram law $\|\mathbf{u} - \mathbf{v}\|^2 + \|\mathbf{u} + \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$.

Of all of the norms $\|\cdot\|_p$ for $1 \leq p \leq \infty$, the only one that satisfies the parallelogram law is $p = 2$, this is the only one given by an inner product.

For example, $\|(a, b)\|_\infty = \max\{|a|, |b|\}$ and clearly we can choose a, b, c , and d so that

$$\max\{|a - c|, |b - d|\} + \max\{|a + c|, |b + d|\} \neq 2 \max\{|a|, |b|\} + 2 \max\{|c|, |d|\}$$

Let $(a, b) = (1, 3)$ and $(c, d) = (2, 1)$, then

$$\begin{aligned} \max\{|1 - 2|, |3 - 1|\} + \max\{|1 + 2|, |3 + 1|\} &= 2 + 4 \\ &\neq 2 \max\{|1|, |3|\} + 2 \max\{|2|, |1|\} = 6 + 4 \end{aligned}$$

- (e) _____ If $\mathcal{C} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal basis for V with respect to an inner product $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$ and $\mathbf{v} \in V$, then for any $(c_1, \dots, c_n) = [\mathbf{v}]_{\mathcal{C}}$, $c_i = \langle \mathbf{v}, \mathbf{u}_i \rangle$.

This is another computation. Say $(c_1, \dots, c_n) = [\mathbf{v}]_{\mathcal{C}}$, then $\mathbf{v} = \sum_{i=1}^n c_i \mathbf{u}_i$. Now just compute

$$\langle \mathbf{v}, \mathbf{u}_j \rangle = \left\langle \sum_{i=1}^n c_i \mathbf{u}_i, \mathbf{u}_j \right\rangle = \sum_{i=1}^n c_i \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \sum_{i=1}^n c_i \delta_{i,j} = c_j$$

Problem 2 (25 points). You are given some data points $\{(x_i, y_i) \mid i = 1, \dots, N\}$ and want to model the data by a function of the form $f(x) = a + bx + c \cos(x) + d \sin(x)$. This involves setting up a matrix A and finding a least-squares solution to $A\mathbf{x} = \mathbf{b}$.

- a) (5 points) What is \mathbf{b} ? (In terms of the data.)

\mathbf{b} is the vector of y_i 's

$$\mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

- b) (8 points) What is A ? (Again, in terms of the data.)

Let \mathbf{X} be the vector of x_i 's.

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

A is the matrix given my columns is

$$A = \begin{bmatrix} \mathbf{1} & \mathbf{X} & \cos(\mathbf{X}) & \sin(\mathbf{X}) \end{bmatrix} = \begin{bmatrix} 1 & x_1 & \cos(x_1) & \sin(x_1) \\ 1 & x_2 & \cos(x_2) & \sin(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_N & \cos(x_N) & \sin(x_N) \end{bmatrix}$$

- c) (7 points) Suppose you have the least-squares solution $\hat{\mathbf{x}}$. What is $f(x)$? (In terms of $\hat{\mathbf{x}}$)

Let

$$\hat{\mathbf{x}} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

then $A\hat{\mathbf{x}} = a\mathbf{1} + b\mathbf{X} + c\cos(\mathbf{X}) + d\sin(\mathbf{X}) = \hat{\mathbf{b}}$ is the vector closest to \mathbf{b} that is of the form $A\mathbf{x}$ for some \mathbf{x} . The model is

$$f(x) = a + bx + c\cos(x) + d\sin(x)$$

So $f(\mathbf{X}) = \hat{\mathbf{b}}$.

- d) (5 points) What is the relationship between $\hat{\mathbf{b}} = A\hat{\mathbf{x}}$ and \mathbf{b} ?

This is described above, $\hat{\mathbf{b}}$ is the vector in \mathbb{R}^N closest to \mathbf{b} where $\hat{\mathbf{b}}$ is of the form $A\mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^4$.