MATH 21B, MARCH 7: LINEAR LEAST SQUARES, DATA FITTING, AND AN INTRODUCTION TO DETERMINANTS

(1) We want to find a quadratic $f(x) = c_2x^2 + c_1x + c_0$ such that $f(-1) = 1, f(0) = 0, f(1) = 2,$ and $f(2) = 5$. Write down a matrix equation to solve for $\vec{x} = \begin{bmatrix} c_2 \\ c_1 \\ c_0 \end{bmatrix}$.

Least Square Solutions: Let $A\vec{x} = \vec{b}$ be an inconsistent system of linear equations. A vector $\vec{x}^*$ which makes $A\vec{x}^*$ as close as possible to $\vec{b}$ is called a *least square solution.* Such vectors $\vec{x}^*$ are solutions to the consistent system

$$A\vec{x} = \text{proj}_{\text{im}(A)}(\vec{b}) = \vec{b} - \vec{b}^{\perp}$$

(2) Let $\vec{b}^{\perp}$ be the part of $\vec{b}$ orthogonal to $\text{im}(A)$.

(a) Explain why $A^T\vec{b}^{\perp} = 0$.

(b) Using this fact, show that $\vec{x}^*$ satisfies the equation $A^TA\vec{x}^* = A^T\vec{b}$. (The equation $A^TA\vec{x} = A^T\vec{b}$ is called the *normal equation* of $A\vec{x} = \vec{b}$.)

(c) Show that if A has linearly independent columns, then $\ker(A^TA) = 0$ and therefore A^TA is invertible. (Hint: If $A^TA\vec{v} = 0$ for some vector $\vec{v}$, then $A\vec{v}$ is orthogonal to all of the columns of A. But remember that $A\vec{v}$ is some linear combination of the columns of A.)

Unique Least Square Solution: If the columns of A are linearly independent (i.e., $\ker(A) = 0$) then the least square solution is unique, and is given by the formula

$$\vec{x}^* = (A^T A)^{-1} A^T \vec{b}$$

(3) Given a matrix A with linearly independent columns and a vector $\vec{b}$ which is not on $\mathrm{im}(A)$, use least squares on the inconsistent system $A\vec{x} = \vec{b}$ to calculate a formula for $\mathrm{proj}_{\mathrm{im}(A)}(\vec{b})$. (Note: if the columns of A were orthonormal, then we have the formula $\mathrm{proj}_{\mathrm{im}(A)}(\vec{b}) = AA^T\vec{b}$. The method above works even if the columns are *not* orthonormal.)

(4) We are going to use least squares to find the best linear fit $y = rx + s$ for the points $(1, 0), (2, 1)$, and $(3, 3)$.

(a) Write down an inconsistent system $A\vec{x} = \vec{b}$ which solves for the coefficients $\vec{x} = \begin{bmatrix} r \\ s \end{bmatrix}$.[1]

(b) Compute $A^T A$ and its inverse.

(c) Compute $\vec{x}^* = (A^T A)^{-1} A^T \vec{b}$, the least square solution.

(d) What is the best-fit line? Try plugging in $x = 1, 2, 3$.

[1] Confusing notation: The vector of coefficients, $\vec{x}$, is unrelated to the x-coordinate of the original question.

Determinant: The *determinant* of an $n \times n$ matrix A is defined as the sum

$$\det(A) = \sum_{\pi} (-1)^{|\pi|} A_{1\pi(1)} A_{2\pi(2)} \cdots A_{n\pi(n)}$$

where π varies over the $n!$ different permutations of the sequence $\{1, 2, \ldots, n\}$ and $|\pi|$ is the number of *up-crossings* in the pattern given by π. The determinant satisfies the following properties (some will only be discussed next class):

- $\det(A) = \det(A^T)$
- If A is upper triangular, then the determinant of A is the product of the diagonal entries.
- Scaling a row or column by c scales $\det(A)$ by c.
- A is invertible $\iff$ $\det(A) \neq 0$.
- Swapping two rows or two columns multiplies $\det(A)$ by -1.
- Adding a multiple of one row (or column) to another row (or column) does not change $\det(A)$.

Laplace Expansion: We can go along the entries in the first column of A and get an expression for the determinant

$$\det(A) = A_{11} \det(B_{11}) - A_{21} \det(B_{21}) + \ldots + (-1)^{n-1} A_{n1} \det(B_{n1})$$

where B_{i1} refers to the $(n-1) \times (n-1)$ matrix formed by removing the i-th row and first column of A. The same can be done with the entries in any other column (since we can swap columns and this introduces a sign) or any row (since we can transpose).

(5) Calculate the determinants.

(a) $\begin{bmatrix} 1 & 0 & 3 \\ 0 & -1 & 4 \\ 2 & 1 & -2 \end{bmatrix}$

(c) $\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \\ -1 & 5 & 3 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 2 & 0 & -1 \\ 0 & 7 & 1 & 0 \\ 1 & 2 & 4 & 1 \\ 5 & 10 & 2 & 5 \end{bmatrix}$

(d) $\begin{bmatrix} 0 & 0 & 0 & 2 \\ 1 & 2 & 4 & 5 \\ 0 & 7 & 2 & 9 \\ 0 & 0 & 6 & 4 \end{bmatrix}$

(6) For what values of λ is the matrix $\begin{bmatrix} 3-\lambda & -2 & 6 \\ 1 & -\lambda & 10 \\ 0 & 0 & 7-\lambda \end{bmatrix}$ invertible?