ENGR 695 Advanced Topics in Engineering Mathematics Fall 2023

Lecture Outline for Monday, Sept. 4

1. Generality #1: Overdetermined systems are usually (but not always) inconsistent.

   Examples: $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \\ 1 & 1 \end{bmatrix}$ $\quad \mathbf{b} = \begin{bmatrix} 3 \\ 7 \\ 2 \end{bmatrix}$ $\quad \mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$

   Although an overdetermined system might not have an exact solution, it could still have a "best" approximate solution.

2. Generality #2: Underdetermined systems are usually (but not always) consistent:

   Examples: $A = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 3 & 5 \end{bmatrix}$ $\quad \mathbf{b} = \begin{bmatrix} 7 \\ 9 \end{bmatrix}$ $\qquad A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix}$ $\quad \mathbf{b} = \begin{bmatrix} 7 \\ 3 \end{bmatrix}$

   Underdetermined systems can have infinitely many solutions or no solution but never a unique solution because rank(A) $\leq M < N$ always.

3. Application of overdetermined systems: Curve-fitting and the method of least squares

   a. Start with an example. Consider the following small data set. How can we estimate the value of $y(3)$, that is, the value of y at $x = 3$?

i	x_i	y_i
1	1.0	1.1
2	2.0	3.2
3	4.0	5.2

   b. One possible approach: Set up a matrix expression that computes the coefficients of the quadratic expression for a curve that passes through the data points.

   $$y = c_0 + c_1 x + c_2 x^2$$

   Is the matrix equation solvable? If so, is the solution acceptable?

   c. Another possible approach: Set up a matrix expression that computes the coefficients of the linear expression for a line that passes through the data points.

   $$y = c_0 + c_1 x$$

   Is the matrix equation solvable? If so, is the solution acceptable?

(*continued on next page*)

4. Curve-fitting: the basic idea

   a. Given a data set: (x_i, y_i), $i = 1$ to M $\rightarrow$ data vectors **x** and **y**
   b. Define model: a set of functions $\{f_j(x_i)\}_{j = 1 \text{ to } N}$ and coefficients $\{c_j\}_{j = 1 \text{ to } N}$ that yield the best approximations to $\{y_i\}_{i = 1 \text{ to } M}$:

      $y(x) \approx \hat{y}(x) = \sum_{j=1}^{N} c_j f_j(x) \quad \rightarrow \quad \hat{\mathbf{y}} = F\mathbf{c}$, where $F_{ij} = f_j(x_i)$ and $\hat{\mathbf{y}}$ contains best fit
   c. Functions $\{f_j(x)\}$ (often called basis functions) can be almost anything; popular choices are 1 and x (linear fit), polynomials (including quadratic and cubic), sin/cos, exponentials, and logarithms
   d. Least squares approach:
      i. Residual vector: $\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}}$ (r_i = distance from actual y_i to approximation $\hat{y}_i$ for each data point i; **r** has M rows)
      ii. Minimize $|\mathbf{r}|^2 = \mathbf{r}^T\mathbf{r}$ or make residual orthogonal to approximation ($\mathbf{r}^T\hat{\mathbf{y}} = 0$)
      iii. Either way, the *normal equation* results (LS solution)

5. Derivation of normal equation from $\mathbf{r}^T\hat{\mathbf{y}} = 0$. Solution:

$$F^T F\mathbf{c} = F^T\mathbf{y} \quad \rightarrow \quad \mathbf{c} = \left(F^T F\right)^{-1} F^T\mathbf{y}$$

6. Practical considerations:

   a. In *Matlab*, can write `c = F\y`; automatically forms solution using normal equations (or its equivalent)
   b. F^TF is symmetric and nonsingular if there are no repeated data points
   c. F is M x N, so F^TF is $N \times N$