ENGR 695 Advanced Topics in Engineering Mathematics Fall 2024

Lecture Outline for Wednesday, Sept. 11

1. Possible limitations of normal equation:

   a. Sometimes produces large-magnitude and oscillatory weights (coefficients) if the basis functions overlap and are highly correlated (e.g., Gaussian or exponential functions)
   b. Problem to be solved might require that all weights be positive or negative
   c. Example: See *Matlab* script `ConstrainedLSDemo.m`

2. Constrained least-squares optimization

   a. Not covered in textbook; see supplemental reading "Constrained Least-Squares Optimization Using Minimized Coefficient Magnitudes"
   b. Start with same basic idea underlying unconstrained LS:
      i. Given a data set: (x_i, y_i), $i = 1$ to M $\rightarrow$ data vectors $\mathbf{x}$ and $\mathbf{y}$
      ii. Define a set of weighted functions $\{f_j(x)\}_{j\,=\,1\text{ to }N}$ that will hopefully fit the data:

      $$y(x) \approx \hat{y}(x) = \sum_{j=1}^{N} c_j f_j(x) \qquad \hat{y}(x) \text{ is the best fit curve}$$

   c. Modifications of unconstrained LS (one of many possible approaches):
      i. Minimize $|\mathbf{c}|^2$ as well as squared error magnitude $|\mathbf{y} - \hat{\mathbf{y}}|^2$
      ii. Coefficients $\{c_j\}_{j\,=\,1\text{ to }N}$ found via

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

      where γ is called a Lagrange multiplier (derivation in supplemental reading)
      iii. In practice, start out with a very small value for γ and then increase it until the coefficients in $\mathbf{c}$ stop oscillating
   d. How it works: By adding a small value to the main diagonal of F^TF, its row vectors become more linearly independent (diagonal of F^TF becomes more dominant).
   e. 2-D analogy: $\mathbf{u}_1$ and $\mathbf{u}_2$ are basis vectors; small circle is solution that they are trying to "reach" via the linear combination $c_1\mathbf{u}_1 + c_2\mathbf{u}_2$, where c_1 and c_2 are scalars. The coefficients c_1 and c_2 must be large and have opposite algebraic signs if $\mathbf{u}_1$ and $\mathbf{u}_2$ are nearly collinear. Adding Lagrange multiplier causes $\mathbf{u}_1$ and $\mathbf{u}_2$ to "fan out."