MATH3027: Optimization 2022

Week 3 computer lab: least squares

Prof. Richard Wilkinson
Please send any comments or mistakes to r.d.wilkinson@nottingham.ac.uk

Polynomial fit and denoising

We will replicate the linear regression and regularized linear least squares examples from this week, with a different model. First, we will generate a noisy dataset of 200 samples coming from

$$v_i = u_i^2 + \mathcal{N}(0, 0.04)\,, \quad i = 1, \ldots, 200\,,$$

where the u_i's are uniformly sampled in $[-1, 1]$, and $\mathcal{N}(0, 0.04)$ means adding Gaussian noise of mean 0 and variance 0.04 for **each** sample.

1. Generate the pairs (u_i, v_i) using suitable random generators. Make a plot illustrating (u_i, v_i).
2. Describe the least-squares regression problem for finding the optimal parameters in the model
$$v(u) = au^2 + bu + c$$
3. Compute the least squares solution and compute the total least squares error in the ℓ_2 norm.
4. Now, instead of solving a regression problem, use the v_i values to recover a denoised signal using regularized least squares using the same total variation regularization described in the lecture notes.

Conditioning

Consider the linear systems

$$\begin{aligned} x + y &= 2 \\ x + 1.001y &= 2 \end{aligned} \tag{1}$$

and

$$\begin{aligned} x + y &= 2 \\ x + 1.001y &= 2.001 \end{aligned} \tag{2}$$

- Find the solution of both systems. Check your answer by using the ‘solve‘ command in R.

  Note that a small change in the RHS ($\mathbf{b}$ in the terminology of the notes) has led to a large change in the solutions ($\mathbf{x}$).

- We can measure the sensitivity of the linear system

  $$\mathbf{A}\mathbf{x} = \mathbf{b}$$

  (where $\mathbf{A}$ is a positive definite $n \times n$ matrix) by measuring how much the output $\mathbf{x}$ changes as $\mathbf{b}$ changes. Suppose $\mathbf{b}$ changes to $\mathbf{b} + \delta\mathbf{b}$, and that this results in the solution $\mathbf{x}$ changing to $\tilde{\mathbf{x}} = \mathbf{x} + \delta\mathbf{x}$. Show that the relative change in $\mathbf{x}$ compared to the relative change to $\mathbf{b}$ can be expressed as follows:

  $$\frac{||\delta\mathbf{x}||/||\mathbf{x}||}{||\delta\mathbf{b}||/||\mathbf{b}||} = \frac{||\mathbf{A}^{-1}\delta\mathbf{b}||}{||\delta\mathbf{b}||}\frac{||\mathbf{A}\mathbf{x}||}{||\mathbf{x}||}$$

- The matrix norm induced by the L_2 norm is defined to be

  $$||\mathrm{A}||_2 = \max\{||\mathbf{A}\mathbf{x}||_2 : ||\mathbf{x}|| \leq 1\} = \max_{\mathbf{x}\neq 0} \frac{||\mathbf{A}\mathbf{x}||_2}{||\mathbf{x}||_2}$$

  Show that

  $$\max_{\mathbf{b},\delta\mathbf{b}} \frac{||\delta\mathbf{x}||/||\mathbf{x}||}{||\delta\mathbf{b}||/||\mathbf{b}||} = ||\mathbf{A}^{-1}|| \cdot ||\mathbf{A}||.$$

- Thus show that for the L_2 norm, we have

  $$\max_{\mathbf{b},\delta\mathbf{b}} \frac{||\delta\mathbf{x}||/||\mathbf{x}||}{||\delta\mathbf{b}||/||\mathbf{b}||} = \frac{\lambda_{max}(\mathbf{A})}{\lambda_{min}(\mathbf{A})}$$

  where $\lambda_{max}(\mathbf{A})$ and $\lambda_{min}(\mathbf{A})$ are the largest and smallest eigenvalues of $\mathbf{A}$.

  Hint: The two norm of the matrix is equal to the largest eigenvalue:

  $$||\mathbf{A}||_2 = \lambda_{max}(\mathbf{A})$$

- This is known as the condition number of $\mathbf{A}$. Compute $\mathbf{A}$ for the linear systems above.