1 Ridge regression and the signal-to-noise ratio

Suppose \( x_1, \ldots, x_n \) and \( \epsilon_1, \ldots, \epsilon_n \) are independent normal random variables, that each \( x_j \) has mean zero and variance \( \tau^2 \), each \( \epsilon_i \) has mean zero and variance \( \sigma^2 \), and we are given a measurement

\[
y = Ax + \epsilon,
\]

where \( A \in \mathbb{R}^{m \times n} \) is a known, nonrandom matrix. We think of \( x \) as a random signal, and \( y \) as a noisy measurement of \( x \), where \( \epsilon \) is the noise that corrupts our measurement. The signal-to-noise ratio is defined to be \( \rho = \left( \frac{\tau}{\sigma} \right)^2 \). Throughout the problem, use the specific parameter values

\[
m = 10, \quad n = 25, \quad \sigma^2 = 3, \quad \text{and} \quad \tau^2 = 2.
\]

Note that there are more parameters than measurements, so we cannot compute a least-squares estimate of \( x \) given \( y \). Carry out \( N = 1000 \) repetitions of the following experiment.

- Generate independent random variables \( x_j, \epsilon_i, \) and \( A_{ij} \) for \( i = 1, \ldots, m \) and \( j = 1, \ldots, n \). Let each \( x_j \) be a normal random variable with mean zero and variance \( \tau^2 \), each \( \epsilon_i \) be a normal random variable with mean zero and variance \( \sigma^2 \), and each \( A_{ij} \) be a uniform random variable on the interval \([0, 1]\). Compute the measurement \( y = Ax + \epsilon \). Note that we can generate a vector of length \( n \) whose components are independent normal random variables with mean zero and variance \( \tau^2 \) using the \texttt{MATLAB} command

\[
\text{tau} \ast \text{randn} (n, 1).
\]

- For \( K = 100 \) values of \( \lambda \) uniformly spaced on a logarithmic scale between \( \frac{1}{1000\rho} \) and \( 1000\rho \), use the measurement \( y \) to compute the regularized least-squares estimate \( \hat{x}(\lambda) \). Then, compute the squared error \( e(\lambda) = \| x - \hat{x}(\lambda) \|^2 \).

Make a plot of the squared error averaged over the \( N \) repetitions of the experiment versus the regularization parameter \( \lambda \). Indicate the point \( \frac{1}{\rho} \) on your plot. Comment on the results.