Math 50 Final – Additional practice problems

November 14, 2025

Exercise 1: Suppose X_1 and X_2 are Normal random variables and

$$\Sigma = \begin{bmatrix} 2 & -1/2 \\ -1/2 & 3 \end{bmatrix} \tag{1}$$

Find the conditional distributions of $X_1|X_2$ and $X_2|X_1$.

Exercise 2: Gaussian process models come from marginalizing over the prior distribution in a nonlinear regression model of the form

$$\beta_j \sim \text{Normal}(0, \tau_j^2), \quad E[\beta_i \beta_j] = 0, \quad i \neq j \tag{2}$$

$$f(x) = \sum_{j=1}^{\infty} \beta_j \phi_j(x) \quad \text{(assuming the sum converges with probability 1)} \tag{3}$$

For two points x_1 and x_2, find the distribution of $f(x_1)|f(x_2)$.

Exercise 3: Which of the following indicates that the assumptions of a linear regression model may be violated? Explain your answer

1. A low R^2
2. Large confidence intervals relative to the values of the fitted coefficients
3. Cross validation is performed and finds the model has a high variance but a low bias
4. None of the above

Exercise 4: Consider the regularized least squares estimator $\hat{\beta}_R$ which minimizes

$$\sum_{i=1}^{N} (Y_i - \beta X_i)^2 + \lambda(\beta - b)^2 \tag{4}$$

Write $\hat{\beta}_R$ in terms of the least squares estimator $\hat{\beta}$.

Exercise 5: You are using Laplace's Rule of Succession to estimate the outcome of an election with two candidates based on polling data. Suppose the true fraction of voters supporting candidate A is 0.7. Approximately how many people must be pooled in order to ensure the 95% interval does not overlap with the prior odds of 1/2?

Exercise 6: Consider the model with interaction term

$$Y = X_1 - X_2 + 3X_3 + 2X_1X_2 + +0.5X_1X_3 + \epsilon \tag{5}$$

If $X_1 \sim \text{Normal}(0, 2)$, what is $Y|X_2, X_3$? Is this satisfy the assumptions of a linear regression model?

Exercise 7: Consider a linear regression model with no interacts ($\beta_0 = 0$) and mean zero predictor ($E[X] = 0$). Derive the sample distribution of $\hat{\beta}_1$ in terms of σ_ϵ^2 and $\hat{\sigma}_X^2$.

Exercise 8: Suppose that a certain parameter θ in a model has a sample distribution

$$\hat{\theta} \sim \text{Normal}(\theta, 0.1) \tag{6}$$

where θ is the true value. Now suppose Bayesian inference is performed with priors

$$\theta \sim \text{Normal}(0, 2) \tag{7}$$

What is the posterior distribution of θ?