Statistical Modelling 2: Homework #1

Professor James Scott

Abhishek Sinha

Problem 1

Give an appropriate positive constant c such that $f(n) \leq c \cdot g(n)$ for all $n > 1$.

1. $f(n) = n^2 + n + 1$, $g(n) = 2n^3$
2. $f(n) = n\sqrt{n} + n^2$, $g(n) = n^2$
3. $f(n) = n^2 - n + 1$, $g(n) = n^2/2$

Solution
We solve each solution algebraically to determine a possible constant c.

Part One

$$\begin{aligned} n^2 + n + 1 &= \\ &\leq n^2 + n^2 + n^2 \\ &= 3n^2 \\ &\leq c \cdot 2n^3 \end{aligned}$$

Thus a valid c could be when $c = 2$.

Part Two

$$\begin{aligned} n^2 + n\sqrt{n} &= \\ &= n^2 + n^{3/2} \\ &\leq n^2 + n^{4/2} \\ &= n^2 + n^2 \\ &= 2n^2 \\ &\leq c \cdot n^2 \end{aligned}$$

Thus a valid c is $c = 2$.

Part Three

$$\begin{aligned} n^2 - n + 1 &= \\ &\leq n^2 \\ &\leq c \cdot n^2/2 \end{aligned}$$

Thus a valid c is $c = 2$.

Problem 2

Part A

$$\begin{aligned}
cov(x) &= E\{(x-\mu)(x-\mu)^T\} \\
&= E\{(x-\mu)(x^T-\mu^T)\}, \text{using property of transpose} \\
&= E\{xx^T - x\mu^T - \mu x^T + \mu\mu^T\}, \text{using distributive property of matrix multiplication} \\
&= E(xx^T) - E(x\mu^T) - E(\mu x^T) + E(\mu\mu^T), \text{using Linearity of Expectation} \\
&= E(xx^T) - E(x)\mu^T - \mu E(x^T) + \mu\mu^T, \mu \text{ being a constant can be pushed out of the expectation} \\
&= E(xx^T) - \mu\mu^T - \mu\mu^T + \mu\mu^T \\
&= E(xx^T) - \mu\mu^T
\end{aligned}$$

$$\begin{aligned}
cov(Ax+b) &= E\{(Ax+b)(Ax+b)^T\} - E(Ax+b)E(Ax+b)^T, \text{using previous result} \\
&= E\{(Ax+b)(x^TA^T+b^T)\} - \{E(Ax)+b\}\{E(Ax)+b\}^T, \text{using property of transpose} \\
&\stackrel{1}{=} E(Axx^TA^T) + E(Axb^T) + E(bx^TA^T) + E(bb^T) - (A\mu+b)(A\mu+b)^T \\
&\stackrel{2}{=} AE(xx^T)A^T + AE(x)b^T + bE(x^T)A^T + bb^T - (A\mu+b)(\mu^TA^T+b^T) \\
&\stackrel{3}{=} AE(xx^T)A^T + A\mu b^T + b\mu^TA^T + bb^T - A\mu\mu^TA^T - A\mu b^T - b\mu^TA^T - bb^T \\
&= AE(xx^T)A^T - A\mu\mu^TA^T \\
&= Acov(x)A^T, \text{using previous result}
\end{aligned}$$

1 follows from linearity of expectation and distributive property of matrix multiplication. 2 follows from the fact that A and μ being constants can be pushed out of the expectation. 3 follows from the definition of μ

Part D

Let us try to find the moment generating function for the random variable z.

$$\begin{aligned}
MGF_x(t) &= E\{\exp(t^Tx)\}, \text{using definition of MGF} \\
&= E\{\exp(t^T(Lz+\mu))\}, \text{substituting for } x \text{ in terms of } z \\
&= E\{\exp(t^TLz)\}E\{\exp(t^T\mu)\}, \text{distributing the powers} \\
&= E\{\exp(t^TLz)\}\exp(t^T\mu), t^T\mu \text{ is constant for a given } t \\
&= E\{\exp((L^Tt)^Tz)\}\exp(t^T\mu), \text{using basic property of transpose} \\
&= \exp((L^Tt)^TI(L^Tt))\exp(t^T\mu), \text{using MGF for standard multivariate normal} \\
&= \exp(t^TLL^Tt)\exp(t^T\mu) \\
&= \exp(t^T\mu + t^T(LL^T)t), \text{combining the powers}
\end{aligned}$$

Now using the if part of the result in part C, we can easily see that $x \sim N(u, LL^T)$

Part E

Suppose $x \sim N(\mu, \Sigma)$. We know that the matrix Σ is symmetric since for any i, j we have that $cov(x_i, x_j) = cov(x_j, x_i)$. From the spectral theorem, there exists an orthonormal matrix U and a diagonal matrix D such that $\Sigma = UDU^T$. We can rewrite this decomposition as $\Sigma = UD^{\frac{1}{2}}(UD^{\frac{1}{2}})^T$ where $D^{\frac{1}{2}}$ is the matrix obtained by taking the square roots of the the diagonal entries of D.

Problem 3

Write part of **Quick-Sort(**$list, start, end$**)**

```
function Quick-Sort(list, start, end)
    if start ≥ end then
        return
    end if
    mid ← Partition(list, start, end)
    Quick-Sort(list, start, mid − 1)
    Quick-Sort(list, mid + 1, end)
end function
```

Algorithm 1: Start of QuickSort

Problem 4

Suppose we would like to fit a straight line through the origin, i.e., $Y_i = \beta_1 x_i + e_i$ with $i = 1, \ldots, n$, $\mathrm{E}[e_i] = 0$, and $\mathrm{Var}[e_i] = \sigma_e^2$ and $\mathrm{Cov}[e_i, e_j] = 0, \forall i \neq j$.

Part A

Find the least squares esimator for $\hat{\beta}_1$ for the slope β_1.

Solution

To find the least squares estimator, we should minimize our Residual Sum of Squares, RSS:

$$\begin{aligned} RSS &= \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 \\ &= \sum_{i=1}^{n} \left(Y_i - \hat{\beta}_1 x_i\right)^2 \end{aligned}$$

By taking the partial derivative in respect to $\hat{\beta}_1$, we get:

$$\frac{\partial}{\partial \hat{\beta}_1}(RSS) = -2\sum_{i=1}^{n} x_i(Y_i - \hat{\beta}_1 x_i) = 0$$

This gives us:

$$\begin{aligned} \sum_{i=1}^{n} x_i(Y_i - \hat{\beta}_1 x_i) &= \sum_{i=1}^{n} x_i Y_i - \sum_{i=1}^{n} \hat{\beta}_1 x_i^2 \\ &= \sum_{i=1}^{n} x_i Y_i - \hat{\beta}_1 \sum_{i=1}^{n} x_i^2 \end{aligned}$$

Solving for $\hat{\beta}_1$ gives the final estimator for β_1:

$$\hat{\beta}_1 = \frac{\sum x_i Y_i}{\sum x_i^2}$$

Problem 4 continued on next page. . .

Part B

Calculate the bias and the variance for the estimated slope $\hat{\beta_1}$.

Solution

For the bias, we need to calculate the expected value $\mathrm{E}[\hat{\beta_1}]$:

$$\begin{aligned}
\mathrm{E}[\hat{\beta_1}] &= \mathrm{E}\left[\frac{\sum x_i Y_i}{\sum x_i^2}\right] \\
&= \frac{\sum x_i \mathrm{E}[Y_i]}{\sum x_i^2} \\
&= \frac{\sum x_i(\beta_1 x_i)}{\sum x_i^2} \\
&= \frac{\sum x_i^2 \beta_1}{\sum x_i^2} \\
&= \beta_1 \frac{\sum x_i^2 \beta_1}{\sum x_i^2} \\
&= \beta_1
\end{aligned}$$

Thus since our estimator's expected value is β_1, we can conclude that the bias of our estimator is 0.

For the variance:

$$\begin{aligned}
\mathrm{Var}[\hat{\beta_1}] &= \mathrm{Var}\left[\frac{\sum x_i Y_i}{\sum x_i^2}\right] \\
&= \frac{\sum x_i^2}{\sum x_i^2 \sum x_i^2} \mathrm{Var}[Y_i] \\
&= \frac{\sum x_i^2}{\sum x_i^2 \sum x_i^2} \mathrm{Var}[Y_i] \\
&= \frac{1}{\sum x_i^2} \mathrm{Var}[Y_i] \\
&= \frac{1}{\sum x_i^2} \sigma^2 \\
&= \frac{\sigma^2}{\sum x_i^2}
\end{aligned}$$

Problem 5

Prove a polynomial of degree k, $a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0$ is a member of $\Theta(n^k)$ where $a_k \ldots a_0$ are nonnegative constants.

Proof. To prove that $a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0$, we must show the following:

$$\exists c_1 \exists c_2 \forall n \geq n_0,\ c_1 \cdot g(n) \leq f(n) \leq c_2 \cdot g(n)$$

For the first inequality, it is easy to see that it holds because no matter what the constants are, $n^k \leq a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0$ even if $c_1 = 1$ and $n_0 = 1$. This is because $n^k \leq c_1 \cdot a_k n^k$ for any nonnegative constant, c_1 and a_k.

Taking the second inequality, we prove it in the following way. By summation, $\sum_{i=0}^{k} a_i$ will give us a new constant, A. By taking this value of A, we can then do the following:

$$\begin{aligned} a_k n^k + a_{k-1} n^{k-1} + \ldots + a_1 n^1 + a_0 n^0 = & \\ & \leq (a_k + a_{k-1} \ldots a_1 + a_0) \cdot n^k \\ & = A \cdot n^k \\ & \leq c_2 \cdot n^k \end{aligned}$$

where $n_0 = 1$ and $c_2 = A$. c_2 is just a constant. Thus the proof is complete. $\square$

Problem 18

Evaluate $\sum_{k=1}^{5} k^2$ and $\sum_{k=1}^{5}(k-1)^2$.

Problem 19

Find the derivative of $f(x) = x^4 + 3x^2 - 2$

Problem 6

Evaluate the integrals $\int_0^1 (1-x^2)\mathrm{d}x$ and $\int_1^\infty \frac{1}{x^2}\mathrm{d}x$.