Solutions due by 10.30am Friday 26th February.

1. Show that if a square matrix is lower triangular or upper trianguar, then its determinant is the product of its diagonal entries.

   Recall lower triangular matrices have zeros below the main diagonal: $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$.

   Solution:

   Prove by induction on the size of the matrix. For the base case, $n = 1$, the determinant is the product of the diagonal since there's only one entry. For the induction step, assume the proposition is true for any $(n-1) \times (n-1)$ matrix and try to show it's true for an $n \times n$ matrix.

   Let A be an $n \times n$ lower triangular matrix. Expanding along the first column, $\det(A) = a_{11}M_{11}$, since $a_{j1} = 0$ for $j \geq 1$. But M_{11} is the determinant of an $(n-1) \times (n-1)$ matrix. By the induction hypothesis its determinant is the product of its main diagonal which is, $M_{11} = a_{22}a_{33}...a_{nn}$. Therefore $\det(A) = a_{11}a_{22} \dots a_{nn}$.Proof for upper triangular follows by transposition. ■

2. Suppose A is diagonalizable. Show that $A^n = \underbrace{A \times A \times \dots A}_{n}$ can be expressed more simply as $A^n = P^{-1}D^nP$.

   Solution:
   Since it's diagonalizable, it can be represented by $A = P^{-1}DP$, where D is diagonal. Note that for a diagonal matrix, $\underbrace{D \times D \times \dots D}_{n} = D^n$ is just the matrix where $(D^n)_{ii} = (D_{ii})^n$.

   $A^n = \underbrace{A \times A \times \dots A}_{n} = \underbrace{(P^{-1}DP) \times (P^{-1}DP) \times \dots (P^{-1}DP)}_{n} = P^{-1}D(PP^{-1})DP \cdots = P^{-1}D^nP$ since $PP^{-1} = I$.

3. Suppose A is diagonalizable. Show that

   a) $trace(A) = \lambda_1 + \cdots + \lambda_n$.

   b) $\det(A) = \lambda_1 \cdot \cdots \cdot \lambda_n$.

Note that this is true in general for square matrices, whether they're diagonalizable or not, but the proof is harder. [Hint: you can use $\det(AB) = \det(A)\det(B)$.]

Solution:
a) $trace(A) = trace(P^{-1}DP) = trace(PP^{-1}D)$ (by cyclic property) $= trace(ID) = trace(D) = \sum_i \lambda_i$

b) $\det(A) = \det(P^{-1}DP) = \det(P^{-1})\det(D)\det(P) = \det(D)$ (since $\det(P^{-1}) = 1/\det(P)$) $= \Pi_i \lambda_i$ (since diagonal matrices are lower triangular).

4. Consider an economy where agents can be in one of the states 'employed' or 'unemployed'. Suppose the probability that an employed person stays employed from one period to the next is 0.95. Suppose the probability that an unemployed person becomes employed in the next period is 0.7.

   a) Write this setup as a system of first order difference equations.
   b) Solve the system in terms of eigenvalues and eigenvectors.
   c) What is the long run rate of unemployment is this economy?

   Solution:
   a) $x_{t+1} = Ax_t$, where $x_t = \begin{bmatrix} e_t \\ u_t \end{bmatrix}$ and $A = \begin{pmatrix} 0.95 & 0.7 \\ 0.05 & 0.3 \end{pmatrix}$.

   b) $trace(A) = 1.25 = \lambda_1 + \lambda_2$. $\det(A) = (0.95 \cdot 0.3 - 0.7 \cdot 0.05) = .25 = \lambda_1\lambda_2$. So $\lambda_1 = 1$, $\lambda_2 = 1/4$.

$$\left(\begin{array}{cc|c} 1-0.95 & -0.7 & 0 \\ -0.05 & 1-0.3 & 0 \end{array}\right) \to \left(\begin{array}{cc|c} 0.05 & -0.7 & 0 \\ 0 & 0 & 0 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & 14 \\ 0 & 0 & 0 \end{array}\right) \implies v_1 = \begin{bmatrix} 14 \\ 1 \end{bmatrix}$$
$$\left(\begin{array}{cc|c} 0.25-0.95 & -0.7 & 0 \\ -0.05 & 0.25-0.3 & 0 \end{array}\right) \to \left(\begin{array}{cc|c} -0.7 & -0.7 & 0 \\ -0.05 & -0.05 & 0 \end{array}\right) \to \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 0 & 0 \end{array}\right) \implies v_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$x_t = a_1 1^t \begin{bmatrix} 14 \\ 1 \end{bmatrix} + a_2 (0.25)^t \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

   c) In the long run, as $t \to \infty$, $(0.25)^t \to 0$. $e_\infty = 14a_1$, $u_\infty = a_1$. The long run unemployment rate is $\frac{u_\infty}{u_\infty + e_\infty} = \frac{1}{1+14} \approx 6.6\%$

5. We defined $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$. Show that $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Solution:
$\mathbb{E}[(X-\mu)^2] = \mathbb{E}[X^2-2\mu X+\mu^2] = \mathbb{E}[X^2]-2\mu\mathbb{E}[X]+\mu^2 = \mathbb{E}[X^2]-2\mu^2+\mu^2 = \mathbb{E}[X^2]-\mu^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

6. The pdf of the exponential distribution is

$$f(x,\lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

its cdf is

$$F(x,\lambda) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

a) Verify that f is a valid density function.

b) Verify that F is the distribution function corresponding to this density.

Solution:
a) $\int_0^\infty \lambda e^{-\lambda x} \mathrm{d}x = \left[-e^{-\lambda x}\right]_0^1 = \lim_{x\to\infty} -e^{-\lambda x} - -e^{-\lambda \cdot 0} = 0 - -1 = 1$. Where the limit goes to zero since $\lambda > 0$.

b) $F'(x;\lambda) = \frac{\mathrm{d}}{\mathrm{d}x}[1 - e^{-\lambda x}] = 0 + \lambda e^{-\lambda x} = f(x;\lambda)$.

7. Show that $M_{aX+b}(t) = e^{bt}M_X(at)$.

Solution:
$M_{aX+b}(t) = \mathbb{E}[e^{t(aX+b)}] = \mathbb{E}[e^{(at)X}e^{bt}] = e^{bt}\mathbb{E}[e^{(at)X}] = e^{bt}M_X(at)$.

8. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu,\sigma^2)$. Show that $\sqrt{n}(\bar{X}-\mu) \sim N(0,\sigma^2)$.

Solution:

$M_{X_i}(t) = \exp\{\mu t + \sigma^2 t^2/2\}$

$\sqrt{n}(\bar{X}-\mu) = \sqrt{n}(\frac{1}{n}\sum X_i - \mu) = \sqrt{n}(\frac{1}{n}\sum(X_i-\mu)) = \sum n^{-1/2}(X_i-\mu)$.

Let $Z_i = n^{-1/2}(X_i-\mu)$. Note that $Z_i = n^{-1/2}X_i - n^{-1/2}\mu$ which is of the form $aX_i + b$. From the question above, we know how to find the mgf.

$M_{Z_i}(t) = \exp\{-n^{-1/2}\mu t\}\exp\{\mu(n^{-1/2}t) + \sigma^2(n^{-1/2}t)^2/2\} = \exp\{-n^{-1/2}\mu t + n^{-1/2}\mu t + n^{-1}\sigma^2 t^2/2\} = \exp\{n^{-1}\sigma^2 t^2/2\}$

$M_{\sqrt{n}(\bar{X}-\mu)}(t) = M_{\sum Z_i}(t) = [M_{Z_i}(t)]^n = \exp\{n^{-1}\sigma^2 t^2/2\}^n = \exp\{\sigma^2 t^2/2\}$, which is the mgf of a $N(0, \sigma^2)$.

9. An economic agent needs to make a decision based on an unknown random variable $\theta \sim N(0, \sigma_\theta^2)$. The agent knows the means and variances of these distributions. They would like to know the realization of θ exactly but are only given a noisy signal $y = \theta + \varepsilon$, where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is iid noise. Suppose the agent's strategy is to guess θ using a multiple of the signal: $\hat{\theta} = cy$. What constant c should the agent pick in order to minimize expected squared error: $\mathbb{E}[(\hat{\theta} - \theta)^2]$?

   Solution:
   $\text{MSE} = \mathbb{E}[(\hat{\theta} - \theta)^2] = \mathbb{E}[\hat{\theta}^2 - 2\hat{\theta}\theta + \theta^2] = \mathbb{E}[c^2y^2 - 2cy\theta + \theta^2] = c^2\mathbb{E}[y^2] - 2c\mathbb{E}[y\theta] + \mathbb{E}[\theta^2] = c^2(\sigma_\theta^2 + \sigma_\varepsilon^2) - 2c\sigma_\theta^2 + \sigma^2\theta$

   First order condition for a minimum is $\frac{\mathrm{d}MSE}{\mathrm{d}c} = 0 \implies 2c(\sigma_\theta^2 + \sigma_\varepsilon^2) - 2\sigma_\theta^2 = 0 \implies c = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2}$.

   Second order condition is $\frac{\mathrm{d}^2 MSE}{\mathrm{d}c^2} = 2(\sigma_\theta^2 + \sigma_\varepsilon^2) > 0$, so we have a necessary and sufficient condition for a local minimum. Since the objective function is strictly convex, we have a global minimum.