Economics 204
Fall 2013
Problem Set 5 Suggested Solutions

1. Let A and B be $n \times n$ matrices such that $A^2 = A$ and $B^2 = B$. Suppose that A and B have the same rank. Prove that A and B are similar.

   Solution. To begin, we note that from the last problem set we know that if λ is an eigenvalue of A then λ^2 is an eigenvalue of A^2. Since $A = A^2$ we must have $\lambda = \lambda^2$ for any eigenvalue of A. Therefore, the eigenvalues of A and B can only be either 0 or 1.

   Now, observe that eigenspace of the eigenvalue $\lambda = 0$ is just $\ker(A)$. We claim that that eigenspace corresponding to the eigenvalue $\lambda = 1$ is $\text{Im}(A)$. Lets pick any $x \in \text{Im}(A)$, i.e. there exists some $y \in X$ such that $x = Ay$

   $$Ax = A \cdot Ay = A^2 y = Ay = x$$

   which implies that x is an eigenvector associated with eigenvalue $\lambda = 1$. Now, observing that any eigenspace of A is a subset of $\text{Im}(A)$ we immediately get the result we desire.

   Let k be the rank of A and B, which is, given what we said above, is dimension of eigenspace corresponding to eigenvector $\lambda = 1$. By the Rank Nullity Theorem we know that

   $$\text{rank} A + \dim \ker(A) = n$$

   thus, the dimension of eigenspace corresponding to eigenvector $\lambda = 0$ is $n - k$ and we know that there n independent eigenvalues that span the whole space. Therefore, A and B are both diagonalizable, implying that there are non-singular matrices C_1 and C_2 such that

   $$C_1 A C_1^{-1} = \begin{bmatrix} I_{k,k} & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad C_2 B C_2^{-1} = \begin{bmatrix} I_{k,k} & 0 \\ 0 & 0 \end{bmatrix}$$

   Thus, we have $C_1 A C_1^{-1} = C_2 B C_2^{-1} \iff A = (C_1^{-1} C_2) B (C_1^{-1} C_2)^{-1}$.

2. Identify which of the following matrices can be diagonalized and provide the diagonalization. If you claim that a diagnoalization does not exist, prove it.

   $$A = \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 3 \\ 1 & 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 3 & -1 & -2 \\ 2 & 0 & -2 \\ 2 & -1 & -1 \end{bmatrix}$$

   Solution. Note that matrix A has two distinct eigenvalues: 1 and -4, and is therefore diagonalizable (these can be found either by inspection using the fact that the sum of the eigenvalues equals the trace of the matrix, or by solving the characteristic polynomial). Two eigenvectors associated with these eigenvalues are $[3, 1]^T$ and $[1, 2]^T$ are respectively. A diagonalization therefore is

$$\begin{bmatrix} 1 & 0 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 2 & -3 \\ 2 & -5 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

Note that, for the matrix on the far right, the columns of the matrix are eigenvectors.

Matrix B has only one eigenvalue: 1, repeated twice. This eigenvalue does not generate two linearly independent eigenvectors: the eigenvector $[1, 0]^T$ spans the eigenspace. Since the dimension of the eigenspace is therefore less than 2, matrix B is not diagonalizable.

Matrix C has three distinct eigenvalues: 1, 2 and -2, and is therefore diagonalizable. Three eigenvectors associated with these eigenvalues are $[3, 1, 2]^T$, $[0, 3, 1]^T$ and $[0, -1, 1]^T$ respectively. A diagonalization therefore is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & -1 \\ 2 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 3 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 1 & 3 & -1 \\ 2 & 1 & 1 \end{bmatrix}$$

Matrix D has only two eigenvalues, 1 and 0, with 1 repeated. However, we are in luck, as the eigenvalue of 1 is associated with two linearly independent eigenvectors: $[1, 0, 1]^T$ and $[1, 2, 0]^T$. The eigenvalue of 0 is associated with the eigenvector $[1, 1, 1]^T$. Thus, the matrix D is diagonalizable and a diagonalization is:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 & -1 & -2 \\ 2 & 0 & -2 \\ 2 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

3. Let A and B be $n \times n$ matrices. Prove or disprove each of the following statements:

   (a) If A and B are diagonalizable, so is $A + B$.

   Solution. Take

$$A = \begin{bmatrix} 1 & 0 \\ -1 & -1 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$$

   they are both diagonalizable (they are either lower- or upper-triangular matrices with district diagonal term), but

$$A + B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

   is NOT diagonalizable.

(b) If A and B are diagonalizable, so is $A \cdot B$.

Solution. Take

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

then A and B are clearly diagonalizable (B is already in diagonal form and A is upper-triangular). However,

$$A \cdot B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

and as you have seen in previous exercise, it is NOT diagonalizable.

(c) If $A^2 = A$ then A is diagonalizable.

Solution. We have just proved that in problem 1. In case, you wondered — no, that was not intended. But you are welcome anyway.

4. Prove that if $f : \mathbf{R} \to \mathbf{R}$ is differentiable, then $\mathrm{Im}\,(f')$ is an interval (possibly a singleton).

Solution. Lets start by observing that $\mathrm{Im}\,(f')$ has to be at least a singleton (i.e. can't be an empty set) because by assumption f is differentiable on all of $\mathbf{R}$. Lets consider following function $F : \Omega \to \mathbf{R}$ defined as

$$F(x, y) = \frac{f(x) - f(y)}{x - y}$$

on open half plane $\Omega = \{(x, y) \in \mathbf{R}^2 |\, x > y\}$. Let J be $\mathrm{Im}\,(F)$. Note that F is continuous and Ω is connected set (we will prove that in a bit!), so J must be a connected set in $\mathbf{R}$, i.e. an interval.

By Mean Value Theorem, for any two point $(x, y) \in \Omega$ there is a point $z \in \mathbf{R}$, such that

$$f(x) - f(y) = f'(z)(x - y).$$

By construction of F, we must have $F(x, y) = f'(z)$. Therefore, it must be the case that $J \subset \mathrm{Im}(f')$. Moreover, for a given $x \in \mathbf{R}$ we have

$$f'(x) = \lim_{y \to x} F(x, y),$$

therefore, $\mathrm{Im}\,(f')$ must be in the closure of J. Thus, we have

$$J \subset \mathrm{Im}(f') \subset \overline{J}$$

and it remains to show that Ω is a connected set. It is easy to see that Ω is a convex set, where we call a set S convex if for any $x,\, x' \in S$ and $\alpha \in [0, 1]$, the point $(1-\alpha)x + \alpha x' \in S$. So, lets prove more general result that any convex set $\Omega \in \mathbf{R}^k$ is connected. As with most proofs of "connectedness" we proceed by contraction. Suppose that there is a separation of Ω, where $\Omega = A \cup B$ and A and B are separated sets. We will prove this claim by showing that separation of Ω implies that a line segment connecting any two points in A and B is also separated. But that would be impossible as it would contradict the definition of Ω being convex.

So, lets pick any $a \in A$ and $b \in B$, and lets define following function:

$$p(t) = (1-t)a + b$$

for $t \in \mathbf{R}$. Lets also define two following sets $A_0 = p^{-1}(A)$, and $B_0 = p^{-1}(B)$. We want to demonstrate that $\overline{A_0} \cap B_0 = \emptyset$ and that $A_0 \cap \overline{B_0} = \emptyset$. Assume to contrary that $A_0 \cap \overline{B_0} \neq \emptyset$ then pick $t_0 \in A_0 \cap \overline{B_0}$. If $t_0 \in B_0$ then we have that $p(t_0) \in A \cap B$ which is a contradiction to the fact that A and B are separated. So, it must be the case that $t_0 \in \overline{B_0}$. Now, we want to show that this implies $p(t_0) \in \overline{B}$ because again, this will lead us to contraction with our assumption that A and B are separated. Note that that for any $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $t \in B_\delta(t_0)$ we have $p(t) \in B_\varepsilon(p(t_0))$ (that is just a continuity of $p(t)$ and, for instance, $\delta = \frac{\varepsilon}{|b-a|}$ will work). Now, since t_0 is a limit point, there is some $t_1 \in B_\delta(t_0) \cap B_0$ such that $t_1 \neq t_0$ and, thus, we have that $p(t_1) \in p(B_\delta(t_0))$ which implies that $p(t_1) \in B_\varepsilon(p(t_0))$. This means that there is $p(t_1) \neq p(t_0)$ for any $\varepsilon-$ball around $p(t_0)$. Thus, $p(t_0) \in \overline{B}$. We arrive at contradiction, because this implies $p(t_0) \in A \cap \overline{B}$. Thus, $A_0 \cap \overline{B_0} = \emptyset$. Similar reasoning can be employed to show that $\overline{A_0} \cap B_0 = \emptyset$.

Now, lets demonstrate that there is a $t_0 \in (0,1)$ such that $p(t_0) \notin A \cup B$. Again, lets reason by contradiction, and suppose that the there is no such t_0 such that $p(t_0) \notin A \cup B$. Then, for all t_0 we have $p(t_0) \in A$ or $p(t_0) \in B$ but not in both because A and B are separated. So this means that $(0,1) \subset p^{-1}(A \cup B) = A_0 \cup B_0$. Since (0,1) is connected it must be completely contained in either A_0 or in B_0. The proof of that by contradiction is short and left as an exercise. So assume that $(0,1) \subset A_0$. This means that $[0,1] \in \overline{A_0}$, but $p(1) \in B$ so that is a contradiction because we have just shown that A_0 and B_0 are separated. Therefore, there must be $t_0 \in (0,1)$ such that $p(t_0) \notin A \cup B$. But this is immediate contradiction to the fact that Ω is a convex set! Therefore, Ω must be connected. We are done.

5. Some practice with implicit function theorem

   (a) Define $f : \mathbf{R}^3 \to \mathbf{R}$ by

   $$f(x,\, y_1,\, y_2) = x^2 y_1 + e^x + y_2.$$

   Show that $f(0,\, 1,\, -1) = 0$ and $D_1 f(0,\, 1,\, -1) \neq 0$. Thus, there exists a differentiable function g in some neighborhood of $(1,-1)$ of $\mathbf{R}^2$, such that

$g(1, -1) = 0$ and

$$f(g(y_1,\, y_2),\, y_1,\, y_2) = 0.$$

Compute $Dg(1,\, -1)$.

Solution. Direct computations immediately verify that $f(0,\, 1,\, -1) = 0$ and $D_1 f(0,\, 1,\, -1) = e^0 = 1$, as well as $D_2 f(0,\, 1,\, -1) = 0$ and $D_3 f(0,\, 1,\, -1) = 1$. Since $D_1 f(0,\, 1,\, -1)$ is non-singular, the Implicit Function Theorem guarantees the existence of C^1 mapping g, defined in the neighborhood of $(y_1,\, y_2) = (1, -1)$, such that $g(y_1,\, y_2) = 0$ and $f(g(y_1,\, y_2),\, y_1,\, y_2) = 0$ as required.

Also, it allows us to compute

$$D_1 g(1,\, -1) = \frac{D_2 f}{D_1 f} = 0 \quad \text{and} \quad D_2 g(1,\, -1) \frac{D_3 f}{D_1 f} = -1.$$

(b) Let $x = x(y, z)$, $y = y(x, z)$ and $z = z(x, y)$ be functions that are implicitly defined by equation $F(x,\, y,\, z) = 0$. Prove that

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = -1.$$

Solution. Since we are not given a particular point $(x_0,\, y_0,\, z_0)$ and since the problem already presumes the existence of implicit functions $x = x(y, z)$, $y = y(x, z)$ and $z = z(x, y)$, we just appeal directly to Implicit Function Theorem to compute the partial derivatives that we need:

$$\frac{\partial x}{\partial y} = -\frac{F_y}{F_x}, \quad \frac{\partial y}{\partial z} = -\frac{F_z}{F_y}, \quad \text{and} \quad \frac{\partial z}{\partial x} = -\frac{F_z}{F_y},$$

which immediately give us the result we need:

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = \left(-\frac{F_y}{F_x}\right)\left(-\frac{F_z}{F_y}\right)\left(-\frac{F_z}{F_y}\right) = -1.$$

with understanding that all partial derivatives above are evaluated at $(x_0,\, y_0,\, z_0)$.

6. Let $f : \mathbf{R} \to \mathbf{R}$ be $\mathcal{C}^1$ function and let

$$\begin{aligned} u &= f(x) \\ v &= -y + xf(x). \end{aligned}$$

If $f'(x_0) \neq 0$, show that this transformation is locally invertible in the neighborhood of $(x_0,\, y_0)$ and that the inverse has the form

$$\begin{aligned} x &= g(u) \\ y &= -v + ug(u). \end{aligned}$$

Solution. Consider $F : \mathbf{R}^2 \to \mathbf{R}^2$ given by

$$F(x,\, y) = (f(x), -y + xf(x)).$$

Computing the Jacobian of F at $(x_0,\, y_0)$ yields $-f'(x_0) \neq 0$. Thus, be the Inverse Function Theorem, F is invertible in an open neighborhood of $(x_0,\, y_0)$.

Applying Implicit Function Theorem, f has a local inverse, g in an open neighborhood of x_0. Thus, we can solve for each component of F^{-1} explicitly and get

$$g(u) = g(f(x)) = x$$

and

$$y = -v + xf(x) = -v + g(u)f(g(u)) = -v + ug(u).$$

7. Using Taylor's formula approximate following two functions

$$f(x, y) = (1+x)^m(1+y)^n \quad \text{and} \quad f(x, y) = \frac{\cos x}{\cos y}.$$

up to the second order terms, assuming x and y are small in absolute value. Estimate approximation error.

Solution. First, lets consider $f(x,y) = (1+x)^m(1+y)^n$. The second-order Taylor series expansion of $f(x,y)$ around (x_0,y_0) is given by

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ Df(x_0,y_0)\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \\ &+ \frac{1}{2!}\begin{pmatrix} x - x_0 & y - y_0 \end{pmatrix} D^2 f(x_0,y_0)\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}. \end{aligned}$$

The derivatives we need to evaluate this expression are given by

$$\begin{aligned} f_x(x,y) &= m(1+x)^{m-1}(1+y)^n \\ f_y(x,y) &= n(1+x)^m(1+y)^{n-1} \\ f_{xx}(x,y) &= m(m-1)(1+x)^{m-2}(1+y)^n \\ f_{yy}(x,y) &= n(n-1)(1+x)^m(1+y)^{n-2} \\ f_{xy}(x,y) &= nm(1+x)^{m-1}(1+y)^{n-1} \end{aligned}$$

which when evaluated at $x_0 = 0$ and $y_0 = 0$ and plugged in yield

$$f(x,y) = 1 + mx + ny + \frac{1}{2}(m(m-1)x^2 + 2mnxy + n(n-1)y^2).$$

The error estimate is the difference between the true value of $f(x,y)$ and the Taylor expansion. We write this as:

$$E_3 = (1+x)^m(1+y)^n - 1 - mx - ny - \frac{1}{2}(m(m-1)x^2 + 2mnxy + n(n-1)y^2).$$

Where the subscript 3 on E_3 indicates that this error term is a third order error term. It is also $O\left(\left|\begin{matrix} x \\ y \end{matrix}\right|^3\right)$.

Now, consider $f(x,y) = \frac{\cos x}{\cos y}$. Using Taylor's formula we have following approximations

$$\cos x = 1 - \frac{x^2}{2} + o(x^2), \quad \text{and} \quad \frac{1}{1-z} = 1 + z + z^2 + o(z^2)$$

that are true when x and z are infinitesimally small.

Thus, we have

$$\frac{\cos x}{\cos y} = \frac{1 - \frac{x^2}{2} + o(x^2)}{1 - \frac{y^2}{2} + o(y^2)} = \left(1 - \frac{x^2}{2} + o(x^2)\right)\left(1 + \frac{y^2}{2} + o(y^2)\right) =$$
$$= 1 - \frac{x^2}{2} + \frac{y^2}{2} + x^2 o(y^2) + y^2 o(x^2) \approx 1 - \frac{x^2 - y^2}{2}.$$

The estimate of the error terms is analogous to that for function $f(x,y) = (1+x)^m(1+y)^n$.