Economics 204
Fall 2011
Problem Set 6 Suggested Solutions

1. Consider the following quadratic forms:

$$\begin{aligned} f(x, y) &= 2x^2 - 4xy + 5y^2, \\ g(x, y) &= x^2 + 6xy + y^2, \\ h(x, y) &= 16xy. \end{aligned}$$

Answer the following questions for each of these forms:

(a) Find a symmetric matrix M such that the form equals $[x\ y]\ M \begin{bmatrix} x \\ y \end{bmatrix}$.

Solution. We seek a, b, c, and d such that:

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 2x^2 - 4xy + 5y^2$$

The diagonal elements of the matrix are the coefficients of the squared terms and the off diagonal elements are half of the crossed terms. This gives us

$$M_f = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 5 \end{bmatrix}.$$

For the second quadratic form, $g(x, y) = 3x^2 - 2xy + y^2$, we find:

$$M_g = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix}.$$

For the third quadratic form, $h(x, y) = 16xy$, we find:

$$M_h = \begin{bmatrix} 0 & 8 \\ 8 & 0 \end{bmatrix}.$$

(b) Find the eigenvalues of matrix M.

Solution. We compute the characteristic polynomial for each matrix, set it equal to zero, and solve. We find that the eigenvalues of M_f are $1, 6$ of M_g are -2, 4 and for M_h we have $-8, 8$.

(c) Find an orthonormal basis of eigenvectors.

Solution. We find that $(2, 1)$ and $(-1, 2)$ form a basis of eigenvectors for M_f. To normalize, we divide both vectors by their respective lengths, which yields an orthonormal basis of eigenvectors:

$$\{v_1, v_2\} = \left\{ \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix} \right\}.$$

Repeating the same process for matrix M_g yields

$$\{w_1, w_2\} = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right\}.$$

and for matrix M_h

$$\{u_1, u_2\} = \left\{ \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right\}.$$

(d) Find a unitary matrix S such that $M = S^{-1}DS$, where D is a diagonal matrix.

Solution. We do this first for matrix M_f. Note that $S^{-1} = (Mtx)_{U,V}(id)$, where U is the standard basis, so the columns are just the eigenvectors:

$$S^{-1} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}.$$

Since the columns of S^{-1} are orthonormal, it follows that S^{-1} is a unitary matrix. Since S^{-1} is unitary, $S = (S^{-1})^{-1} = (S^{-1})^T$, so

$$S = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}.$$

Similarly, for matrices M_g and M_h we have

$$S^{-1} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

and $S = (S^{-1})^T$ as well.

(e) Describe the level sets of the form and state whether the form has a local maximum, local minimum, or neither at $(0, 0)$. (Level sets are solutions to $f(x, y) = c$ for some $c \in \mathbf{R}$.)

Solution. The quadratic form f is associated with matrix M_f which has two positive eigenvalues. This means that the level sets are ellipses and that there is a local minimum at the origin.

The maximum value of the form on the unit circle is simply the norm of M, which is equal to the largest of the absolute values of the eigenvalues, which is 6. Similarly, the minimum value of the form on the unit circle is 1. To obtain level sets of the form, we convert to the basis of orthonormal eigenvectors v_1 and v_2 and write

$$f(\gamma_1, \gamma_2) = (\gamma_1,\ \gamma_2)\ D \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}$$

where D is the diagonal matrix of eigenvalues and $(\gamma_1, \gamma_2)^T$ are the coordinates of $(x,\, y)^T$ in the basis $\{v_1,\, v_2\}$, i.e.

$$\begin{pmatrix}\gamma_1\\ \gamma_2\end{pmatrix} = S\begin{pmatrix}x\\ y\end{pmatrix}$$

Thus, we have ellipses associated with the level c described by the equation $6\gamma_1^2 + \gamma_2^2 = c$. For these eclipses major axis is along the line formed by v_2 and a minor axis is along the line formed by v_1. The ellipse crosses the major axis at $\pm\sqrt{c}$, and crosses the minor axis at $\pm\sqrt{c/6}$. Note, therefore, that the ellipse is longest in the direction of the eigenvector corresponding to the smallest eigenvalue.

The form g is associated with M_g, which has one positive and one negative eigenvalue. This means that there is neither a maximum or a minimum at the origin. The maximum value of the form on the unit circle is 4, and the minimum value is -2. Level sets are given by $r(\gamma_1, \gamma_2) = -2\gamma_1^2 + 4\gamma_2^2 = c$, which generates a hyperbola, where (γ_1, γ_2) are coordinates in the eigenvector basis. Rearranging this to $\gamma_1 = \pm\sqrt{2\gamma_2^2 - c/4}$ informs us that the slopes of the asymptotes in the (γ_1, γ_2) plane are plus and minus $\sqrt{2}$. Taking first the case that $c < 0$ and solving for $\gamma_2 = 0$, we see that the hyperbola crosses the γ_2-axis at $\gamma_1 = \pm\sqrt{-c/4}$. For $c > 0$, we can no longer have $\gamma_2 = 0$, and we instead solve for $\gamma_1 = 0$ to learn that the hyperbola intersects the γ_2-axis at the points $\gamma_2 = \pm\sqrt{c/2}$ (and for c = 0, the hyperbola reduces to the asymptotes).

The form g is associated with M_g, which has one positive and one negative eigenvalue. Again, this means that there is neither a maximum or a minimum at the origin. The maximum value of the form on the unit circle is 8, and the minimum value is -8.
Level sets are given by $g(\gamma_1, \gamma_2) = -8\gamma_1^2 + 8\gamma_2^2 = c$, which generates a hyperbola. Rearranging this to $\gamma_1 = \pm\sqrt{\gamma_2^2 - c/8}$ informs us that the slopes of the asymptotes in the (γ_1, γ_2) plane are plus and minus one. Taking first the case that $c < 0$ and solving for $\gamma_2 = 0$, we see that the hyperbola crosses the γ_1-axis at $\gamma_1 = \pm\sqrt{-c/8}$. For $c > 0$, we can no longer have $\gamma_2 = 0$, and we instead solve for $\gamma_1 = 0$ to learn that the hyperbola intersects the γ_2-axis at the points $\gamma_2 = \pm\sqrt{c/8}$ (and for c = 0, the hyperbola reduces to the asymptotes).

2. Suppose $\Psi_1,\, \Psi_2 : X \to 2^Y$ are compact-valued, upper hemicontinuous correspondences, where $X \subset \mathbf{R}^n$, $Y \subset \mathbf{R}^m$ for some n, m. Suppose that $\Psi_1 \cap \Psi_2 \neq \emptyset$ for each $x \in X$.

   (a) Show that $\Psi_1 \cap \Psi_2$ is upper hemicontinuous, where $\Psi_1 \cap \Psi_2$ is defined by

$$(\Psi_1 \cap \Psi_2)(x) = \Psi_1(x) \cap \Psi_2(x),\ \forall x \in X$$

Solution. We will appeal to the sequential characterization of upper hemicontinuity (Theorem 12 in lecture 7), since $\Psi_1 \cap \Psi_2$ is clearly a compact-valued.

So, lets denote by $\Psi_1 \cap \Psi_2$ by Φ, fix $x_0 \in X$ and consider an arbitrary sequence $\{x_n\} \subset X$, with $x_n \to x_0$. Let $\{y_n\}$ be a companion sequence of $\{x_n\}$, i.e. we have $y_n \in \Phi(x_n)$ for all n. We have to show that there is a convergent subsequence $\{y_{n_k}\}$ such that $\lim y_{n_k} \in \Phi(x_0)$.

Since $y_n \in \Phi(x_n)$ for all n, it must be the case that $y_n \in \Psi_1(x_n)$ and $y_n \in \Psi_2(x_n)$ for all n. We know that Ψ_2 is compact-valued and upper-hemicontinuous, therefore there exists a subsequence $\{y_{n_k}\}$ of $\{y_n\}$ such that $y_{n_k} \to y_0$ with $y_0 \in \Psi_2(x_0)$. Now, since Ψ_1 is also uhc and compact-valued, there exists a further subsequence $\{y_{n_{k_\ell}}\}$ of $\{y_{n_k}\}$ such that $y_{n_{k_\ell}} \to y'$ and $y' \in \Psi_1(x_0)$. Since $\{y_{n_{k_\ell}}\}$ is a subsequence of $\{y_{n_k}\}$ and $y_{n_k} \to y_0$, $y_{n_{k_\ell}} \to y_0$ as well. By uniqueness of limits $y_0 = y'$, thus $y_0 \in \Psi_1(x_0)$. So $y_0 \in \Psi_1(x_0) \cap \Psi_2(x_0)$ and we are done.

(b) Lets now weaken our assumptions a bit: lets assume that Ψ_1 is only closed-valued, rather then compact-valued. Show that $\Psi_1 \cap \Psi_2$ is still upper hemicontinuous.

Solution. Note that we can again use Theorem 12 to give us sufficient conditions for $\Phi(x)$ be upper hemicontinuous. So, as before lets fix $x_0 \in X$, and consider an arbitrary sequence $\{x_n\} \subset X$, with $x_n \to x_0$. Let $\{y_n\}$ be a companion sequence of $\{x_n\}$, i.e. $y_n \in \Phi(x_n)$ for all n which by definition implies $y_n \in \Psi_1(x_n)$ and $y_n \in \Psi_2(x_n)$ for all n.

Because by our assumption $\Psi_2(x)$ is compact valued and upper hemicontinuous, $\{y_n\}$ has a convergent subsequence $\{y_{n_k}\} \in \Psi_2(x_{n_k})$. Let y_0 be the limit of that subsequence, so that we have $y_0 \in \Psi_2(x_0)$.

Now, lets consider the sequence of ordered pairs $\{x_{n_k}, y_{n_k}\}$ that converges to (x_0, y_0). By our assumption $\Psi_1(x)$ is closed-valued and upper hemicontinuous and thus has closed graph. Therefore, our limit (x_0, y_0) is in the graph of $\Psi_1(x)$, which implies that $y_0 \in \Psi_1(x_0) \implies y_0 \in \Psi_2(x_0)$, and we get the result we seek.

Finally, you might have already noticed that 204 final exam for 2010 further weakened the assumptions of this problem asking for the proof when *both* $\Psi_1(x)$ and $\Psi_2(x)$ are closed. While the final question also asked you to supply proof that directly uses the definition of upper hemicontinuity, it is a good exercise to think through one that relies on alternative, i.e. sequential, characterization.

3. Let $f : \mathbf{R}^2 \to \mathbf{R}^2$ be a C^1 function and define $F : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}^2$ by

$$F(x, \omega) = f(x_1, x_2) + \left(5\omega_1 + \omega_1^3,\ 5\omega_2 + \omega_1(1 + 3\omega_1\omega_2)\right).$$

Show that there is a set of Lebesgue measure zero, $\Omega_0 \subset \mathbf{R}^2$, such that if $\omega \notin \Omega_0$, then for each x_0 satisfying $F(x_0, \omega_0) = 0$ there is an open set U containing x_0, an open set V containing ω_0, and a C^1 function $h : V \to U$ such that for all $\omega \in V$, $x = h(\omega)$ is the unique element of U satisfying $F(x, \omega) = 0$.

Solution. If we can show that the Jacobian of F with respect to all of its arguments has rank 2 whenever $F(x, \omega) = 0$, then the Transversality Theorem guarantees that there is a set of Lebesgue measure zero, $\Omega_0 \subset \mathbf{R}^2$, such that if $\omega \notin \Omega_0$, then for each x_0 satisfying $F(x_0, \omega_0) = 0$, $D_x F(x_0, \omega_0)$ has rank 2 as well.

In this setup, the Jacobian of F is given by

$$DF(x, \omega) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & 5 + 3\omega_1^2 & 0 \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & 1 + 6\omega_1\omega_2 & 5 + 3\omega_1^2 \end{bmatrix},$$

since $\frac{\partial F_1}{\partial \omega_1} = 5 + 3\omega_1^2$, $\frac{\partial F_2}{\partial \omega_1} = 1 + 6\omega_1\omega_2$, $\frac{\partial F_1}{\partial \omega_2} = 0$, and $\frac{\partial F_2}{\partial \omega_2} = 5 + 3\omega_1^2$. The matrix

$$\begin{bmatrix} 5 + 3\omega_1^2 & 0 \\ 1 + 6\omega_1\omega_2 & 5 + 3\omega_1^2 \end{bmatrix}$$

has rank 2 for all ω, and therefore the matrix $DF(x, \omega)$ must also have rank 2 whenever $F(x, \omega) = 0$. It follows that the Transversality Theorem applies to this function F, and hence there is a set of Lebesgue measure zero, $\Omega_0 \subset \mathbf{R}^2$, such that if $\omega \notin \Omega_0$, then for each x_0 satisfying $F(x_0, \omega_0) = 0$, $D_x F(x_0, \omega_0)$ has rank 2 as well. Of course, this implies that $|D_x F(x_0, \omega_0)| \neq 0$.

We now complete the proof by using the Implicit Function Theorem. It states that whenever $F(x_0, \omega_0) = 0$ and $|D_x F(x_0, \omega_0)| \neq 0$, there is an open set U containing x_0, an open set V containing ω_0, and a C^1 function $h : V \to U$ such that for all $\omega \in V$, $x = h(\omega)$ is the unique element of U satisfying $F(x, \omega) = 0$.

4. *The Minimax Theorem* is used for proving quite a few important results in economics, for instance, about an outcome of zero-sum games in noncooperative game theory or in analyzing Bayesian estimators in statistical decision theory. Now you have a chance to prove this Minimax Theorem yourself.

   Let X and Y be non-empty, closed, bounded and convex subsets of any two Euclidean spaces. Prove that if $f : X \times Y \to \mathbf{R}$ is continuous, and if the sets $\{\, z \in X \mid f(z, y) \geq \alpha\}$ and $\{\, w \in Y \mid f(x, w) \leq \alpha\}$ are convex for each $(x,\ y,\ \alpha) \in X \times Y \times \mathbf{R}$, then

$$\max_{x \in X} \min_{y \in Y} f(x, y) = \min_{y \in Y} \max_{x \in X} f(x, y).$$

   (Hint: Start by defining two self-correspondences $\Phi(y) : Y \to 2^Y$ and $\Pi(x) : X \to 2^X$ as

$$\Phi(y) = \underset{x \in X}{\operatorname{argmax}} f(x, y)$$
$$\Pi(x) = \underset{y \in Y}{\operatorname{argmin}} f(x, y)$$

Then, define self-correspondence $\Psi : X \times Y \to 2^{X\times Y}$ by

$$\Psi(x, y) = \Pi(x) \times \Phi(y).$$

Use Kakutani's Fixed Point theorem).

Solution. First, observe that our requirement that the sets $\{ z \in X \mid f(z, y) \geq \alpha\}$ and $\{ w \in Y \mid f(x, w) \leq \alpha\}$ be convex for each $(x, y, \alpha) \in X \times Y \times \mathbf{R}$ essentially, tells us that $f(x, y)$ is quasi-concave in x and quasi-convex in y. Graphically, you can imagine that function $f(x, y)$ has a "saddle" at some point $(x^*, y^*) \in \mathbf{R}^2$. Because we have such saddle point, we can switch the order of min and max functions. This is the "heart" of the Minimax Theorem.

One can show that a necessary and sufficient condition for the existence of a saddle point is

$$f(x, y^*) \leq f(x^*, y^*) \leq f(x^*, y) \quad \text{for every } x \in Y \text{ and } y \in Y.$$

This is the insight that we will use in our proof. In what follows, we will define a self-correspondence on $X \times Y$ and show that its fixed point is actually the saddle point that we seek.

To begin, note that inequality in one direction is rather straightforward:

$$\max_{x\in X} \min_{y\in Y} f(x, y) \leq \min_{y\in Y} \max_{x\in X} f(x, y).$$

This is true because for all $x \in X$ and $y \in Y$ we must have

$$f(x, y) \leq \max_{x\in X} f(x, y).$$

This implies that for all $x \in X$ and $y \in Y$

$$\min_{y\in Y} f(x, y) \leq \max_{x\in X} f(x, y),$$

and so on. Proving inequality in other direction is much more difficult and is, in fact, an application of the Kakutani Fixed Point argument.

Lets first define two self-correspondences $\Phi(y) : Y \to 2^Y$ and $\Pi(x) : X \to 2^X$ as

$$\Phi(y) = \underset{x\in X}{\text{argmax}} f(x, y)$$
$$\Pi(x) = \underset{y\in Y}{\text{argmin}} f(x, y)$$

and, finally, define self-correspondence $\Psi : X \times Y \to 2^{X\times Y}$ by

$$\Psi(x, y) = \Pi(x) \times \Phi(y).$$

Observe that Φ and Π are both non-empty, compact-valued and upper-hemicontinuous by the Berge's Theorem of Maximum. By our assumption of quasi-concavity and quasi-convexity, they are both convex-valued. Thus, Ψ is also an upper-hemicontinuous correspondence with non-empty, convex and compact values and Kakutani's Fixed Point Theorem guarantee's the existence of $(x^*, y^*) \in X \times Y$ with $(x^*, y^*) \in \Psi(x^*, y^*)$. In particular, this means that

$$x^* \in \underset{x \in X}{\operatorname{argmax}} f(x, y^*) \quad \text{and} \quad y^* \in \underset{y \in Y}{\operatorname{argmin}} f(x^*, y)$$

In, other words, we have a saddle with

$$f(x, y^*) \leq f(x^*, y^*) \leq f(x^*, y) \quad \text{for every } x \in Y \text{ and } y \in Y.$$

To get the result we seek, denote by $V = f(x^*, y^*)$. By definition of (x^*, y^*) we have

$$\begin{aligned} f(x, y^*) \leq V \quad \text{for all } x \in X \implies \\ \max_{x \in X} f(x, y^*) \leq V \end{aligned}$$

Therefore,

$$\min_{y \in Y} \max_{x \in X} f(x, y) \leq \max_{x \in X} f(x, y^*) \leq V.$$

Similarly, we obtain

$$\max_{x \in X} \min_{y \in Y} f(x, y) \geq V,$$

and we are done.

5. Show that the closure of a convex set is convex.

Solution. Let C be a convex set and let $\bar{C}$ denote its closure. We wish to show that $\bar{C}$ is convex. Let $\lambda \in [0, 1]$, and let $x, y \in \bar{C}$. We will show that $\lambda x + (1 - \lambda) y \in \bar{C}$. Since $x, y \in \bar{C}$ we can find convergent sequences x_n, y_n in C converging to x and y, respectively. Moreover, letting $z_n = \lambda x_n + (1 - \lambda) y_n$, we obtain, by the convexity of C, that $z_n \in C$. Note, moreover, that $z_n \to \lambda x + (1 - \lambda) y = z$. Hence, z is a limit point of C which implies that $z = \lambda x + (1 - \lambda) y \in \bar{C}$. We are done

6. One of the most useful versions of the Separating Hyperplane Theorem is the one on *strong separation* of convex sets. We say that two sets A and B are strongly separated by a hyperplane if there exists $p \in \mathbf{R}^n$ with $p \neq 0$ such that

$$\sup p \cdot A < \inf p \cdot B$$

(In other words, sets are strongly separated if they are contained in the closed halfspaces that are $\epsilon > 0$ away from each other. Notice that another way to show strong separation is to demonstrate existence of two constants c and d together with non-zero vector p such that $p \cdot a \leq c < d \leq p \cdot b$ for all $a \in A$ and for all $b \in B$. Please check Theorem 8 in lecture 13 to make sure you understand how strict separation is different from strong one.)

Of course, strong separation requires a stronger initial assumptions.

(a) Let A and B be non-empty, disjoint, convex subsets of $\mathbf{R}^n$ with A being *compact* and B closed. Show directly, without invoking Theorem 7 in Lecture 13, that A and B can be strongly separated.[1]

(Hint: Look at the set $Y = B - A$. Is it compact? Closed?)

Solution. Consider the set $Y = B - A$. We claim that Y is non-empty, closed and convex. Non-emptiness and convexity is immediate (but make sure you know why!). Now, lets show Y is closed, so take a sequence $\{y_n\}$ of elements in Y, converging to point y.

By definition of Y, for every n there exists $a_n \in A$ and $b_n \in B$ such that $y_n = b_n - a_n$. Since A is compact, there exists a subsequence a_{n_k} that is converges to $a \in A$. Note that $b_{b_k} = y_{n_k} + a_{n_k}$, so b_{n_k} is a convergent subsequence. Let $\lim b_{n_k} = b$ and note that $b \in B$ because B is closed. By continuity, $b_{n_k} - a_{n_k} \to b - a$ and by our assumption $b_n - a_n = y_n \to y \implies y \in Y$.

Now, note that $\{0\} \notin Y$ because A and B are disjoint. By Theorem 7 in Lecture 13, we can separate them with a hyperplane, i.e. there exists a non-zero $p \in \mathbf{R}^n$ such that $p \cdot 0 \leq p \cdot Y$ or $p \cdot y \geq 0$ for all $y \in Y$.

Theorem 7 assumes for the clarity of exposition that Y is compact. In our case, Y is just closed. However, by doing some extra work, we will strengthen our conclusions by showing essentially that our sets are $\epsilon > 0$ away. We will demonstrate that, in addition, to non-zero p vector, there exists a non-zero constant c such that $p \cdot y \geq c > 0$ for all $y \in Y$.

Lets imagine for a moment, we have proved that already. Then, the way Y is defined, for all $a \in A$ and $b \in B$

$$p \cdot b - p \cdot a \geq c > 0$$

This implies

$$p \cdot a + c \leq p \cdot b$$

or

$$\sup p \cdot A + c \leq \inf p \cdot B$$

and we get the result we seek.

Now, it just remains to prove our assertion about existence of non-zero constant c. We will show this by first proving an intermediate step that given a non-empty, convex, closed set Y and a point outside of it, $x \notin Y$, we can always find a unique point $y_0 \in Y$ that is "closest" to x, in a sense that[2]

$$||y_0 - x|| \leq ||y - x||, \text{ for all } y \in Y.$$

[1]Although you can't invoke theorem directly, going carefully over its proof will get you far in showing this result.

[2]Since uniqueness is not essential to our argument, we will not dwell on it here

Moreover, for all $y \in Y$

$$(y_0 - x)^T \cdot (y - y_0) \geq 0.$$

To see this, first, lets pick some point $\hat{y} \in Y$ and define the set $\hat{Y}$ of all points that are closer to x then $\hat{y}$

$$\hat{Y} = \{\, y \in Y \mid ||y - x|| \leq ||\hat{y} - x||\}$$

Clearly, $\hat{Y}$ is compact, as it is closed and bounded. Because norm is a continuous function, $g(y) = ||y - x||$ attains a minimum on $\hat{Y}$ at some point y_0. Also, notice that for all $y \in Y$, $||y_0 - x|| \leq ||y - x||$, that is, y_0 is a closest point to x in Y.
Secondly, take any $y \in Y$. By convexity of Y, the line segment $\alpha y + (1 - \alpha)y_0 \in Y$ and by the argument given above

$$\begin{aligned}
||y_0 - x||^2 \leq & ||(\alpha y + (1-\alpha)y_0) - x||^2 = \\
= & ||(y_0 + \alpha(y - y_0)) - x||^2 = \\
= & ||(y_0 - x) + \alpha(y - y_0)||^2 = \\
= & ||y_0 - x||^2 + 2\alpha(y_0 - x)^T \cdot (y - y_0) + \alpha^2||y - y_0||^2.
\end{aligned}$$

This implies that

$$2(y_0 - x)^T \cdot (y - y_0) + \alpha||y - y_0||^2 \geq 0.$$

Finally, letting $\alpha \to 0$ we get the result we seek.
Now, we define our separating hyperplane as $(y_0 - x)^T \cdot y$ (i.e. as in the theorem, we are taking $p = y_0 - x$; notice that it is a shortest distance from x to the set Y and, thus, a normal vector to our hyperplane). We also set $c = (y_0 - x)^T \cdot y_0$.[3] Then for all $y \in Y$ we have

$$(y_0 - x)^T \cdot y - (y_0 - x)^T \cdot y_0 = (y_0 - x)^T \cdot (y - y_0) \geq 0$$

and therefore

$$(y_0 - x)^T \cdot y \geq (y_0 - x)^T \cdot y_0 = c.$$

Moreover,

$$(y_0 - x)^T \cdot y_0 - (y_0 - x)^T \cdot x = ||y_0 - x||^2 > 0, \text{ because } y_0 \neq x$$

As a result, we have $(y_0 - x)^T \cdot y_0 > (y_0 - x)^T \cdot x$ and

$$(y_0 - x)^T \cdot x < c \leq (y_0 - x)^T \cdot y \text{ for all } y \in Y,$$

thus, completing the proof.

[3] You can think about this constant c as a projection of the vector y_0 onto the normal vector p.

(b) Demonstrate by means of an example that the requirement that A is compact is essential, it can't be just closed. In your example, are the sets A and B strictly separated?

Solution. Consider following two sets:

$$\begin{aligned} A =& \{(x, y) \in \mathbf{R}^2 \mid xy \geq 1\} \\ B =& \{(x, y) \in \mathbf{R}^2 \mid y \leq 0\} \end{aligned}$$

Clearly, A and B can't be strongly separated. Intuitively, A and B come arbitrarily "close" to each other. At the same time, those two sets can be strictly separated with $p = (0, 1)$.

7. Consider the following inhomogeneous linear differential equation

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

(a) Write down the corresponding homogeneous equation.

(b) Find the general solution of the homogeneous equation.

(c) Find a particular solution of the original inhomogeneous equation satisfying the initial condition $y(0) = (1,\ 1)^T$.
(Hint: The integrals can be solved by integrating by parts twice.)

(d) Find the general solution of the original inhomogeneous equation.

Solution. This exercise is worked out in the lecture 15 notes on page 5. In case, you wondered — no, that was not intended. But you are welcome anyway. Good luck on the exam!