

Exercise 1 It is often possible to determine whether an invertible symmetric matrix is positive definite, negative definite, or indefinite without explicitly computing the eigenvalues of the matrix. Let $A \in M_{n \times n}(\mathbb{R})$. We define the **trace** of A , $\text{tr}(A)$, to be $\text{tr}(A) \stackrel{\text{def}}{=} a_{11} + a_{22} + \cdots + a_{nn}$ (that is, the sum of the entries along the main diagonal of A).

- (a) Suppose that A is diagonalizable, and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A (repeated according to algebraic multiplicity). Show that

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n \quad \text{and} \quad \text{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n.$$

(Hint: This is similar to a result that you proved that quarter in discussion for complex matrices. Here, first show that you can factor the characteristic polynomial $p(\lambda)$ of A into first-order factors.)

- (b) Determine whether each of the invertible matrices $\begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$ and $\begin{bmatrix} 6 & 4 \\ 4 & 2 \end{bmatrix}$ are positive definite, negative definite, or indefinite without explicitly computing their eigenvalues.

(Suggestion: If the eigenvalues of each matrix are λ_1 and λ_2 , use part (a) to determine whether both eigenvalues are positive, both negative, or one is positive and one is negative.)

- (a) Since A is diagonalizable, there exists diagonal matrix D with the entries being all of the eigenvalues, and invertible matrix S such that $A = SDS^{-1}$. Then,

$$\det(A) = \det(SDS^{-1}) = \det(S)\det(D)\det(S^{-1}) = \det(D)$$

Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues for A . Then,

$$\det(A) = \det(D) = \lambda_1 \cdots \lambda_n$$

Now, we will show that $\text{tr}(A) = \lambda_1 + \cdots + \lambda_n$. Consider the characteristic polynomial $\det(A - \lambda I)$, by the result from discussion last quarter, the coefficient on λ^{n-1} in $p(\lambda)$ is $(-1)^{n-1}(a_{11} + \cdots + a_{nn})$.

At the same time, since we know that $\lambda_1, \dots, \lambda_n$ are all the solutions to $p(\lambda)$, we can write $p(\lambda)$ as

$$p(\lambda) = (-1)^c(\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$$

where $c \in \mathbb{Z}$. Since we know that the coefficient on λ^n in $p(\lambda)$ is $(-1)^n$, so $c = n$. Then the coefficient on λ^{n-1} is

$$-1^c(-\lambda_1 - \cdots - \lambda_n) = -1^{c+1}(\lambda_1 + \cdots + \lambda_n) = -1^{n+1}(\lambda_1 + \cdots + \lambda_n)$$

Since the coefficient on λ^{n-1} is unique, and since $n+1$ and $n-1$ has the same parity, so $\lambda_1 + \cdots + \lambda_n = a_{11} + \cdots + a_{nn}$

(b) We have

$$\det\begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix} = 3 \times 2 - 2 \times 2 = 2, \operatorname{tr}\begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix} = 3 + 2 = 5$$

By part (a), we know that $\lambda_1 + \lambda_2 = 5, \lambda_1 \lambda_2 = 2$. Since the product is positive, so the two eigenvalues must have the same sign. And since the sum is positive, so both eigenvalues must be positive. So the matrix is positive definite.

Similarly we have

$$\det\begin{pmatrix} 6 & 4 \\ 4 & 2 \end{pmatrix} = 6 \times 2 - 4 \times 4 = -4, \operatorname{tr}\begin{pmatrix} 6 & 4 \\ 4 & 2 \end{pmatrix} = 6 + 2 = 8$$

By part (a), we know that $\lambda_1 + \lambda_2 = 8, \lambda_1 \lambda_2 = -4$. Since the product is negative, so the two eigenvalues must have different signs. So the matrix is indefinite.

■

Exercise 2 Determine the absolute minimum and maximum values of the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) \stackrel{\text{def}}{=} 2x^2 - 2xy + y^2 - y + 3$ on the closed triangular region with vertices $(0, 0)$, $(2, 0)$, and $(0, 2)$.

Let E be the triangular region enclosed by $(0, 0)$, $(2, 0)$, $(0, 2)$.

1. We will first determine the critical points of f :

$$\begin{bmatrix} 0 & 0 \end{bmatrix} = \begin{bmatrix} f_x(x, y) & f_y(x, y) \end{bmatrix} = \begin{bmatrix} 4x - 2y & -2x + 2y - 1 \end{bmatrix} \Rightarrow (x, y) = \left(\frac{1}{2}, 1\right)$$

Since $(\frac{1}{2}, 1) \in E$, we need to check

$$f\left(\frac{1}{2}, 1\right) = \frac{1}{2} - 1 + 1 - 1 + 3 = \frac{5}{2}$$

2. We will now check ∂E .

1⁰ On $y = 0, x \in [0, 2]$. Define $g(x) \stackrel{\text{def}}{=} f(x, 0) = 2x^2 + 3$. Then,

$$0 = g'(x) = 4x \Rightarrow x = 0 \in [0, 2]$$

or the other end point $(2, 0)$.

2⁰ On $x = 0, y \in [0, 2]$. Define $h(y) \stackrel{\text{def}}{=} f(0, y) = y^2 - y + 3$. Then,

$$0 = h'(y) = 2y - 1 \Rightarrow y = \frac{1}{2} \in [0, 2]$$

or the end points $(0, 0)$, $(0, 2)$.

3⁰ On $x + y - 2 = 0, x \in [0, 2]$. Define

$$l(t) \stackrel{\text{def}}{=} f(t, -t + 2) = 2t^2 - 2t(-t + 2) + (-t + 2)^2 + (-t + 2) + 3 = 5t^2 - 7t + 5$$

Then,

$$0 = l'(t) = 10t - 7 \Rightarrow t = \frac{7}{10} \in [0, 2]$$

or the end points $(0, 2)$, $(2, 0)$.

So, we need to test f at $(0, 0)$, $(0, \frac{1}{2})$, $(0, 2)$, $(\frac{7}{10}, \frac{13}{10})$, $(\frac{1}{2}, 1)$, $(2, 0)$.

$$f(0, 0) = 3, f(0, \frac{1}{2}) = \frac{11}{4}, f(0, 2) = 5, f(\frac{7}{10}, \frac{13}{10}) = \frac{51}{20}, f(\frac{1}{2}, 1) = \frac{5}{2}, f(2, 0) = 11$$

So, the global maximum of f is 11 at $(2, 0)$. The global minimum is $\frac{5}{2}$ at $(\frac{1}{2}, 1)$. ■

Exercise 3 (Colley 4.2.53(b)) Define $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by $f(x, y) \stackrel{\text{def}}{=} 2 - (xy^2 - y - 1)^2 - (y^2 - 1)^2$. Show that f has exactly two critical points, and that both of them are local maxima.

pf: Since f is C^1 on \mathbb{R}^2 (because it is polynomial), (x, y) is a critical point of f exactly when

$$\begin{bmatrix} 0 & 0 \end{bmatrix} = Df(x, y) = \begin{bmatrix} f_x(x, y) & f_y(x, y) \end{bmatrix} = \begin{bmatrix} -2(xy^2 - y - 1)y^2 & -2(xy^2 - y - 1)(2xy - 1) - 2(y^2 - 1)2y \end{bmatrix}$$

So, either $xy^2 - y - 1 = 0$ or $y^2 = 0$.

1. If $xy^2 - y - 1 = 0$, then $-2(xy^2 - y - 1)(2xy - 1) = 0$, so $(y^2 - 1)2y = 0$. Then, $y = -1, 0, 1$.

1⁰ If $y = -1$, then $x = 0$.

2⁰ If $y = 0$, then $-1 = 0$. Contradiction!

3⁰ If $y = 1$, then $x = 2$.

2. If $y^2 = 0$, then $y = 0$. Then $(-2)(-1)(-1) = 0$. Contradiction!

Thus, the two critical points are $(0, -1), (2, 1)$. Now, we'll show that they are both maximas.

$$D^2f(x, y) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} = \begin{bmatrix} -2y^4 & -2(2xy - 1)y^2 - 4y(xy^2 - y - 1) \\ -2(2xy - 1)y^2 - 4y(xy^2 - y - 1) & -2(2xy - 1)^2 - 4x(xy^2 - y - 1) - 8y^2 - 4(y^2 - 1) \end{bmatrix}$$

So,

$$D^2f(0, -1) = \begin{bmatrix} -2 & 2 \\ 2 & -2 - 8 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & -10 \end{bmatrix}$$

Then, the trace is -12, while the determinant is 16. So both of the eigenvalues are negative, f has a local maxima at $(0, -1)$.

$$D^2f(2, 1) = \begin{bmatrix} -2 & -6 \\ -6 & -18 - 8 \end{bmatrix} = \begin{bmatrix} -2 & -6 \\ -6 & -26 \end{bmatrix}$$

Then, the trace is -28, while the determinant is 16. So both of the eigenvalues are negative, f has a local maxima at $(2, 1)$. ■

Exercise 4 (Colley 4.3.28) Heron's formula for the area of a triangle whose sides have lengths x , y , and z is

$$\text{Area} = \sqrt{s(s-x)(s-y)(s-z)},$$

where $s = \frac{1}{2}(x + y + z)$ is the so-called semiperimeter of the triangle. Use Heron's formula to prove that for a fixed perimeter P , the triangle with the largest area is equilateral. (Your proof should also include a justification that there is indeed a triangle with largest area.)

■

Exercise 5 (Colley 4.3.44 and 4.3.45) Let $S^{n-1} \stackrel{\text{def}}{=} \{\vec{x} \in \mathbb{R}^n : \|\vec{x}\| = 1\}$ be the unit hypersphere in \mathbb{R}^n centered at $\vec{0}$.

(a) Fix $\vec{x}_0 \in S^{n-1}$ and define $f : \mathbb{R}^n \rightarrow \mathbb{R}$ by $f(\vec{y}) = \vec{x}_0 \cdot \vec{y}$. Determine the maximum and minimum values of f on S^{n-1} .

(b) Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ be any nonzero vectors. Use part (a) to prove that $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$.

(a)

(b)

■

Exercise 6 Let $A \in M_{m \times n}(\mathbb{R})$, and define $f : \mathbb{R}^n \rightarrow \mathbb{R}$ by $f(\vec{x}) \stackrel{\text{def}}{=} \|A\vec{x}\|$. Use the method of Lagrange multipliers to determine the maximum value of f on S^{n-1} (defined in the previous problem).

(Note: Last quarter you applied the Extreme Value Theorem to determine that f does indeed have a maximum value on S^{n-1} . The point of this problem is to actually compute this value!)

pf: Since $\|A\vec{x}\| \geq 0$ for all $\vec{x} \in \mathbb{R}^n$, so when $\|A\vec{x}\|$ reaches the maximum, $\|A\vec{x}\|^2$ will also reach the maximum. Let $h(\vec{x}) \stackrel{\text{def}}{=} \|A\vec{x}\|^2$. And let $g(\vec{x}) \stackrel{\text{def}}{=} \|\vec{x}\|^2$. Then,

$$\nabla g(\vec{x}) = D(g(\vec{x}))^T = [2x_1 \quad \cdots \quad 2x_n]^T = \begin{bmatrix} 2x_1 \\ \vdots \\ 2x_n \end{bmatrix} = 2\vec{x}$$

We will show that $\nabla h(\vec{x}) = 2A^T A\vec{x}$. Now similar to the example we did in class, fix $1 \leq k \leq n$, and write $\vec{x} = x_k \vec{e}_k + \vec{v}$ where $\vec{v} = x_1 \vec{e}_1 + \cdots + x_{k-1} \vec{e}_{k-1} + x_{k+1} \vec{e}_{k+1} + \cdots + x_n \vec{e}_n$. Then,

$$\begin{aligned} h(\vec{x}) &= A(x_k \vec{e}_k + \vec{v}) \cdot A(x_k \vec{e}_k + \vec{v}) \\ &= (Ax_k \vec{e}_k + A\vec{v}) \cdot (Ax_k \vec{e}_k + A\vec{v}) \\ &= Ax_k \vec{e}_k \cdot Ax_k \vec{e}_k + 2Ax_k \vec{e}_k \cdot A\vec{v} + A\vec{v} \cdot A\vec{v} \\ &= x_k^2 (A\vec{e}_k \cdot A\vec{e}_k) + 2x_k (A\vec{e}_k \cdot A\vec{v}) + A\vec{v} \cdot A\vec{v} \end{aligned}$$

Note that A, \vec{v} do not depend on x_k , so that

$$\begin{aligned} h_{x_k}(\vec{x}) &= 2x_k (A\vec{e}_k \cdot A\vec{e}_k) + 2(A\vec{e}_k \cdot A\vec{v}) \\ &= A\vec{e}_k \cdot 2Ax_k \vec{e}_k + 2A\vec{e}_k \cdot A\vec{v} \\ &= A\vec{e}_k \cdot (Ax_k \vec{e}_k + A\vec{v}) \\ &= A\vec{e}_k \cdot 2A\vec{x} \\ &= \vec{e}_k \cdot (2A^T A\vec{x}) \end{aligned}$$

Since h_{x_k} is the k^{th} entry of $\nabla h(\vec{x})$ and $\vec{e}_k \cdot (2A^T A\vec{x})$ is the k^{th} entry of $(2A^T A\vec{x})$, so $\nabla h(\vec{x}) = 2A^T A\vec{x}$. By the method of Lagrange Multipliers, a point \vec{x} at which h attains a constrained extreme value must satisfy, for some $\lambda \in \mathbb{R}$,

$$\begin{cases} \nabla h(\vec{x}) &= \lambda \nabla g(\vec{x}) \\ g(\vec{x}) &= 1 \end{cases}$$

So we have

$$2A^T A\vec{x} = 2\lambda \vec{x} \Rightarrow A^T A\vec{x} = \lambda \vec{x}$$

Since $\|\vec{x}\|^2 = 1$, so $\vec{x} \neq 0$, so \vec{x} is an eigenvector for $A^T A$. Then,

$$f(\vec{x}) = \|A\vec{x}\| = \sqrt{(A\vec{x}) \cdot (A\vec{x})} = \sqrt{\vec{x}^T A^T A\vec{x}} = \sqrt{\vec{x} \cdot (A^T A\vec{x})} = \sqrt{\vec{x} \cdot (\lambda \vec{x})} = \sqrt{\lambda \vec{x} \cdot \vec{x}} = \sqrt{\lambda} \|\vec{x}\| = \sqrt{\lambda}$$

So, the maximum value of f on S^{n-1} is the square root of the largest eigenvalue for $A^T A$. ■

Exercise 7 Consider the problem of optimizing a C^2 function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ subject to the constraints

$$g_1(\vec{x}) = c_1, \quad g_2(\vec{x}) = c_2, \quad \dots, \quad g_k(\vec{x}) = c_k,$$

where $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is C^2 . Define $L : \mathbb{R}^{k+n} \rightarrow \mathbb{R}$ by

$$L(\vec{\lambda}, \vec{x}) \stackrel{\text{def}}{=} f(\vec{x}) - \lambda_1(g_1(\vec{x}) - c_1) - \dots - \lambda_k(g_k(\vec{x}) - c_k),$$

where $\vec{\lambda} = (\lambda_1, \dots, \lambda_k) \in \mathbb{R}^k$ and $\vec{x} \in \mathbb{R}^n$.

(a) Show that $(\vec{\lambda}, \vec{x})$ is a critical point of L if and only if

$$\nabla f(\vec{x}) = \lambda_1 \nabla g_1(\vec{x}) + \dots + \lambda_k \nabla g_k(\vec{x}) \quad \text{and} \quad g_1(\vec{x}) = c_1, \quad g_2(\vec{x}) = c_2, \quad \dots, \quad g_k(\vec{x}) = c_k.$$

(b) Compute the Hessian $D^2L(\vec{\lambda}, \vec{x})$ of L , and express $D^2L(\vec{\lambda}, \vec{x})$ in terms of $\vec{\lambda}$, $D^2f(\vec{x})$, $D^2g_i(\vec{x})$ (for $1 \leq i \leq k$), and $D\vec{g}(\vec{x})$ (where $\vec{g} : \mathbb{R}^n \rightarrow \mathbb{R}^k$ is $\vec{g}(\vec{x}) = (g_1(\vec{x}), \dots, g_k(\vec{x}))$).

(Note: $D^2L(\vec{\lambda}, \vec{x})$ can be used to determine whether the local extrema of f subject to the given constraints are maximums, minimums, or saddle points. This is analogous to how we used Hessians in unconstrained optimization problems.)

(a)

(b)

■

Exercise 8 Suppose that f and g are both integrable on the closed rectangle $R \subset \mathbb{R}^2$. Prove that the following properties hold.

1. $f + g$ is also integrable on R and

$$\iint_R (f + g) dA = \iint_R f dA + \iint_R g dA.$$

2. For $c \in \mathbb{R}$, cf is also integrable on R and

$$\iint_R cf dA = c \iint_R f dA.$$

3. If $f(x, y) \leq g(x, y)$ for all $(x, y) \in R$, then

$$\iint_R f dA \leq \iint_R g dA.$$

4. $|f|$ is also integrable on R and

$$\left| \iint_R f dA \right| \leq \iint_R |f| dA.$$

(Note: This is Proposition 2.7 in Section 5.2. Your book proves part 1., so you need only prove parts 2., 3., and 4.)

2. Let \mathcal{P} be a partition of R , and let \mathcal{C} be a choice of sample points for \mathcal{P} . Then,

$$R(cf, \mathcal{P}, \mathcal{C}) = \sum_i cf(\vec{c}_i) \text{Vol}_n(R_i) = c \sum_i f(\vec{c}_i) \text{Vol}_n(R_i) = cR(f, \mathcal{P}, \mathcal{C})$$

So,

$$\iint_R cf dA = \lim_{\|\mathcal{P}\| \rightarrow 0} R(cf, \mathcal{P}, \mathcal{C}) = \lim_{\|\mathcal{P}\| \rightarrow 0} cR(f, \mathcal{P}, \mathcal{C}) = c \lim_{\|\mathcal{P}\| \rightarrow 0} R(f, \mathcal{P}, \mathcal{C}) = c \iint_R f dA$$

3. Let \mathcal{P} be a partition of R , and let \mathcal{C} be a choice of sample points for \mathcal{P} . Then,

$$R(f, \mathcal{P}, \mathcal{C}) - R(g, \mathcal{P}, \mathcal{C}) = R(f - g, \mathcal{P}, \mathcal{C}) = \sum_i (f(\vec{c}_i) - g(\vec{c}_i)) \text{Vol}_n(R_i)$$

Since $f(\vec{c}_i) \leq g(\vec{c}_i)$ for all $\vec{c}_i \in \mathbb{R}^2$, and since $\text{Vol}_n(R_i) \geq 0$ for all $R_i \in R$, so

$$0 \leq \sum_i (f(\vec{c}_i) - g(\vec{c}_i)) \text{Vol}_n(R_i)$$

So,

$$\begin{aligned}\iint_R f \, dA - \iint_R g \, dA &= \iint_R f \, dA + \iint_R -g \, dA \\ &= \iint_R f - g \, dA \\ &= \lim_{\|\mathcal{P}\| \rightarrow 0} R(f - g, \mathcal{P}, \mathcal{C}) \\ &\leq 0\end{aligned}$$

Thus,

$$\iint_R f \, dA \leq \iint_R g \, dA.$$

■

Exercise 9 Define $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$f(x, y) \stackrel{\text{def}}{=} \begin{cases} 10 - x^2 - y^2 & \text{if } \|(x, y)\| \leq 2, \\ x^2 + y^2 & \text{if } \|(x, y)\| > 2. \end{cases}$$

Show that f is integrable over every rectangle of the form $[-a, a] \times [-b, b]$ centered at the origin in \mathbb{R}^2 .

■

Exercise 10 Show that every finite subset of \mathbb{R}^n has measure zero.

■