Solutions
Math 291-3
Homework 1
April 8, 2022

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, $\mathrm{tr}(A)$, to be $\mathrm{tr}(A) \stackrel{def}{=} a_{11} + a_{22} + \cdots + a_{nn}$ (that is, the sum of the entries along the main diagonal of A).

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

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

(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) Let $p(\lambda) = \det(A - \lambda I_{n\times n})$ be the characteristic polynomial of A. Because A is diagonalizable, A has exactly n eigenvalues $\lambda_1, \ldots, \lambda_n$ (counted according to geometric multiplicity). Because the algebraic multiplicity of an eigenvalue is at least as great as the geometric multiplicity of the eigenvalue, the sum of the algebraic multiplicities of the eigenvalues of A is n and the characteristic polynomial of A factors as $p(\lambda) = (\lambda_1 - \lambda)\cdots(\lambda_n - \lambda)$. Note that $\det(A) = p(0) = \lambda_1 \cdots \lambda_n$, establishing the first claim. For the second claim, note that we can expand $p(\lambda)$ as

$$p(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}(\lambda_1 + \cdots + \lambda_n)\lambda^{n-1} + (\text{lower order terms}),$$

so that the coefficient of λ^{n-1} is $(-1)^{n-1}(\lambda_1 + \cdots + \lambda_n)$. On the other hand, when computing $p(\lambda) = \det(A - \lambda I_{n\times n})$ using the pattern definition of the determinant, note the pattern $P = \{(1,1), \ldots, (n,n)\}$ identifying the locations on the main diagonal has no inversions and contributes the term $(-1)^0(a_{11} - \lambda)\cdots(a_{nn} - \lambda)$ to $\det(A - \lambda I_{n\times n})$. Expanding this term gives $(-1)^n\lambda^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})\lambda^{n-1} + (\text{lower order terms})$. On the other hand, if Q is any other pattern in $A - \lambda I_n$, then there are at least two locations in this pattern that do not lie on the main diagonal, and therefore $\mathrm{prod}(Q(A - \lambda I_{n\times n}))$ is a polynomial in λ of degree no more than $n-2$. Therefore

$$p(\lambda) = \det(A - \lambda I_{n\times n}) = (-1)^n\lambda^n + (-1)^{n-1}(a_{11} + \cdots + a_{nn})\lambda^{n-1} + (\text{lower order terms}),$$

so that the coefficient of λ^{n-1} is $(-1)^{n-1}\mathrm{tr}(A)$. Therefore we have $\mathrm{tr}(A) = \lambda_1 + \cdots + \lambda_n$.

(b) Let $A = \begin{bmatrix} 3 & 2 \\ 2 & 2 \end{bmatrix}$, and let λ_1, λ_2 be the eigenvalues of A (which we know exist because A is symmetric). By (a), $\lambda_1 + \lambda_2 = \mathrm{tr}(A) = 5$ and $\lambda_1\lambda_2 = \det(A) = 2$. Because $\lambda_1\lambda_2 > 0$, λ_1 and

λ_2 are either both positive, or both negative. Because their sum is positive, they are both positive. It follows that A is positive definite.

Let $B = \begin{bmatrix} 6 & 4 \\ 4 & 2 \end{bmatrix}$, and let λ_1, λ_2 be the eigenvalues of B (which we know exist because A is symmetric). By (a), $\lambda_1\lambda_2 = \det(B) = -4$. Because $\lambda_1\lambda_2 < 0$, one of λ_1 and λ_2 is positive and the other is negative. Therefore B is indefinite.

■

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

Note that f is C^1 (and therefore differentiable) on $\mathbb{R}^2$, so that the only critical points that f may have must satisfy

$$\begin{bmatrix} 0 & 0 \end{bmatrix} = Df(x, y) = \begin{bmatrix} 4x - 2y & 2y - 2x - 1 \end{bmatrix}, \quad \text{or rather} \quad \begin{bmatrix} 4 & -2 \\ -2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Therefore we have $\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1 \end{bmatrix}$, so that the only critical point of f is $(\frac{1}{2}, 1)$. Because this point lies in the triangular region, $f(\frac{1}{2}, 1) = \frac{5}{2}$ may be a global max or min value of f on the triangular region.

We now identify points on the boundary of the region where f may have a local min or max. Note that the boundary of this region consists of the three line segments C_1, C_2, and C_3 connecting (respectively) $(0, 2)$ to $(0, 0)$ and $(0, 0)$ to $(2, 0)$ and $(0, 2)$ to $(2, 0)$. We inspect the values of f on each of these three line segments by constraining the inputs of f to these line segments and thinking of f as a single-variable function.

On C_1 we have $g(y) = f(0, y) = y^2 - y + 3$ for $y \in [0, 2]$. This function may have extreme values when $y = 0$, $y = 2$, or where $0 = g'(y) = 2y - 1$, so $y = \frac{1}{2}$. Therefore, possible extreme values of f are $f(0, 0) = 3$, $f(0, 2) = 5$ and $f(0, \frac{1}{2}) = \frac{11}{4}$.

On C_2 we have $g(x) = f(x, 0) = 2x^2 + 3$ for $x \in [0, 2]$. This function may only have extreme values when $x = 0$ or $x = 2$ (since $0 = g'(x) = 4x$ when $x = 0$, which is already an endpoint of the interval). Therefore, possible extreme values of f are $f(0, 0) = 3$ and $f(2, 0) = 11$.

On C_3 we have $g(x) = f(x, 2 - x) = 2x^2 - 2x(2 - x) + (2 - x)^2 - (2 - x) + 3 = 5x^2 - 7x + 5$ for $x \in [0, 2]$. This function may have extreme values when $x = 0$, $x = 2$, or when $0 = g'(x) = 10x - 7$, or $x = \frac{7}{10}$. Therefore, possible extreme values of f are $f(0, 2) = 5$, $f(2, 0) = 11$, or $f(\frac{7}{10}, \frac{13}{10}) = \frac{51}{20}$.

Comparing all of these values, we see that f has a global minimum value of $\frac{5}{2}$ at $(\frac{1}{2}, 1)$, and a global maximum value of 11 at $(2, 0)$. ■

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

Note that f is C^2 (and therefore differentiable) on $\mathbb{R}^2$, so the only critical points of f satisfy

$$\begin{bmatrix} 0 & 0 \end{bmatrix} = Df(x, y) = \begin{bmatrix} -2y^2(xy^2 - y - 1) & -2(2xy - 1)(xy^2 - y - 1) - 4y(y^2 - 1) \end{bmatrix}.$$

From the equation $0 = f_x(x, y)$, we see that either $y = 0$ or $xy^2 - y - 1 = 0$. Note that $f_y(x, 0) = -2 \neq 0$ for every x, so that $y \neq 0$. Therefore we have $y \neq 0$ and $xy^2 - y - 1 = 0$, so that

$x = \frac{1+y}{y^2}$. From the equation $0 = f_y(x, y)$ and the condition that $xy^2 - y - 1 = 0$, we have $0 = -4y(y^2 - 1) = -4y(y-1)(y+1)$. Because $y \neq 0$, we conclude that either $y = 1$ or $y = -1$. If $y = 1$ then $x = \frac{1+1}{1^2} = 2$. If $y = -1$ then $x = \frac{1-1}{(-1)^2} = 0$. Therefore the only possible critical points of f are $(2, 1)$ and $(0, -1)$. One can quickly verify that $(2, 1)$ and $(0, -1)$ are indeed critical points of f.

We now compute that

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

so that

$$D^2 f(2, 1) = \begin{bmatrix} -32 & -24 \\ -24 & -26 \end{bmatrix} \quad \text{and} \quad D^2 f(0, -1) = \begin{bmatrix} -2 & 2 \\ 2 & -10 \end{bmatrix}.$$

Because $D^2 f(2, 1)$ is symmetric, it has two eigenvalues λ_1, λ_2. Because $-58 = \mathrm{tr}(D^2 f(2, 1)) = \lambda_1 + \lambda_2$, at least one of the eigenvalues is negative. Because $\lambda_1 \lambda_2 = \det(D^2 f(2, 1)) = 256 > 0$, it must be that both of the eigenvalues are negative, so that $D^2 f(2, 1)$ is negative definite. The Second Derivative Test implies that f has a local maximum at $(2, 1)$.

Because $D^2 f(0, -1)$ is symmetric, it has two eigenvalues λ_1, λ_2. Because $-12 = \mathrm{tr}(D^2 f(0, -1)) = \lambda_1 + \lambda_2$, at least one of the eigenvalues is negative. Because $\lambda_1 \lambda_2 = \det(D^2 f(0, -1)) = 16 > 0$, it must be that both of the eigenvalues are negative, so that $D^2 f(0, -1)$ is negative definite. The Second Derivative Test implies that f has a local maximum at $(0, -1)$. ■

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.)

Fix $s > 0$. Let T denote the portion of the plane $2s = x + y + z$ in the first octant (i.e. where $x \geq 0$, $y \geq 0$, and $z \geq 0$. Then T is closed an bounded, and therefore compact. Moreover, the function $A(x, y, z) \stackrel{def}{=} \sqrt{s(s-x)(s-y)(s-z)}$ is continuous on T, and therefore (by the Extreme Value Theorem) has a global maximum value on T.

We also note that $x, y, z \leq s$. To show this, suppose that $\vec{P}, \vec{Q}, \vec{R}$ are the vertices of the triangle T and $x = \|\vec{P} - \vec{Q}\|$ and $y = \|\vec{Q} - \vec{R}\|$ and $z = \|\vec{P} - \vec{R}\|$, then

$$2x = x + \|\vec{P} - \vec{Q}\| \leq x + \|\vec{P} - \vec{R}\| + \|\vec{R} - \vec{Q}\| = x + z + y = 2s,$$

so that $x \leq s$. A similar argument shows that $y \leq s$ and $z \leq s$ as well.

The global maximum value of A on T either occurs on the edge of T (i.e. where $x = 0$ or $y = 0$ or $z = 0$) or is a constrained local maximum value of A on T. If (say) $x = 0$, then we have $2s = y + z$. But since $y \leq s$ and $z \leq s$, we must have $y = z = s$. Therefore $A(0, y, z) = 0$. A similar argument shows that $A(x, 0, z) = 0$ and $A(x, y, 0) = 0$ on the other portions of the edge of T. Because $A(x, y, z) > 0$ when $x > 0$ and $y > 0$ and $z > 0$, the global maximum of $A(x, y, z)$ on T must be a constrained local maximum value of A on T.

On the other hand, by the method of Lagrange Multipliers a constrained local extreme value of A on T must occur at a point (x, y, z) (with $x > 0$ and $y > 0$ and $z > 0$) that satisfies (for some

$\lambda \in \mathbb{R}$)

$$\begin{cases} \nabla A(x,y,z) & \lambda \nabla[x+y+z] \\ x+y+z=2s \end{cases} \qquad \Leftrightarrow \qquad \begin{cases} \frac{-s(s-y)(s-z)}{2A(x,y,z)} = \lambda \\ \frac{-s(s-x)(s-z)}{2A(x,y,z)} = \lambda \\ \frac{-s(s-x)(s-y)}{2A(x,y,z)} = \lambda \\ x+y+z=2s \end{cases}$$

Note that since $s>0$ and $s-x>0$ and $s-y>0$ and $s-z>0$, $\lambda \neq 0$. Multiplying each fo the first three equations (respectively) by the nonzero quantities $-2(s-x)$ and $-2(s-y)$ and $-2(s-z)$ yields

$$A(x,y,z) = \lambda(s-x) \quad \text{and} \quad A(x,y,z) = \lambda(s-y) \quad \text{and} \quad A(x,y,z) = \lambda(s-z),$$

whence it follows that $\lambda(s-x) = \lambda(s-y) = \lambda(s-z)$. Because $\lambda \neq 0$, we have $(s-x) = (s-y) = (s-z)$, so that $-x=-y=-z$, so that $x=y=z$, and since $x+y+z=2s$ we have $x=y=z=\frac{2s}{3}$. Therefore there is a single choice of x,y,z for which $A(x,y,z)$ might have a global extreme value on T, and (by our comments above) this choice must yield the global maximum value of A on T. That is, the triangle with perimeter $2s$ that has maximum area is an equilateral triangle. ■

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

(a) Fix $\vec{x}_0 \in S^{n-1}$ and define $f : \mathbb{R}^n \to \mathbb{R}$ 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) Because f is C^1 (and therefore continuous) on $\mathbb{R}^n$ and S^{n-1} is compact, f does attain maximum and minimum values on S^{n-1}. Writing S^{n-1} as the level set $g(\vec{y}) = \|\vec{y}\|^2 = y_1^2 + \cdots + y_n^2 = 1$, we see that these global extreme values of f must be constrained extreme values, and therefore (since $\nabla g(\vec{y}) = 2\vec{y} \neq \vec{0}$ on S^{n-1}) satisfy, for some $\lambda \in \mathbb{R}$,

$$\begin{cases} \nabla f(\vec{y}) = \lambda \nabla g(\vec{y}) \\ g(\vec{y}) = 1 \end{cases} \qquad \Leftrightarrow \qquad \begin{cases} \vec{x}_0 = \lambda 2\vec{y} \\ \|\vec{y}\|^2 = 1 \end{cases}$$

Because $\vec{x}_0 \neq 0$, $\lambda \neq 0$ and therefore $\vec{y} = \frac{1}{2\lambda}\vec{x}_0$. Because $1 = \|\vec{y}\| = \frac{1}{2|\lambda|}\|\vec{x}_0\| = \frac{1}{2|\lambda|}$, we must have $\lambda = \pm 2$ and therefore the candidates for the extreme values of f on S^{n-1} are $f(-\vec{x}_0) = -\|\vec{x}_0\|^2 = -1$ and $f(\vec{x}_0) = \|\vec{x}_0\|^2 = 1$. Therefore the global maximum value of f on S^{n-1} is 1 (at $\vec{x}_0$), and the global minimum value of f on S^{n-1} is -1 (at $-\vec{x}_0$).

(b) Let $\vec{x}, \vec{y} \neq \vec{0}$, and let $\vec{x}' = \frac{1}{\|\vec{x}\|}\vec{x}$ and $\vec{y}' = \frac{1}{\|\vec{y}\|}\vec{y}$. By part (a) (with $\vec{x}_0 = \vec{x}'$), we have

$$-1 \leq \vec{x}' \cdot \vec{y}' \leq 1, \quad \text{so that} \quad |\vec{x}' \cdot \vec{y}'| \leq 1.$$

But then we multiply both sides of this inequality by $\|\vec{x}\|\|\vec{y}\|$ to obtain $|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\|\|\vec{y}\|$.

■

Exercise 6 Let $A \in M_{m \times n}(\mathbb{R})$, and define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(\vec{x}) \stackrel{def}{=} \|A\vec{x}\|$. 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!)

Consider the function $g : \mathbb{R}^2 \to \mathbb{R}$ defined by $g(\vec{x}) = (f(\vec{x}))^2 = \|A\vec{x}\|^2 = (A\vec{x}) \cdot (A\vec{x}) = \vec{x} \cdot (A^T A\vec{a})$. Note that because $f(\vec{x}) \geq 0$ for all $\vec{x}$, if f achieves a global maximum at $\vec{x} \in S^{n-1}$ then so does g. Note that $A^T A \in M_{n\times n}(\mathbb{R})$ is symmetric, and therefore (by a theorem proved in class) the global maximum value of g on S^{n-1} is the largest eigenvalue λ of $A^T A$, and is attained at an eigenvector $\vec{x} \in S^{n-1}$ of $A^T A$ associated to the eigenvalue λ. Then the global maximum value of $f(\vec{x}) = \sqrt{g(\vec{x})}$ is $\sqrt{\lambda}$. ∎

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

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

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

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

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

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

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

(Note: $D^2 L(\vec{\lambda}, \vec{x})$ 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) Note that $(\lambda_1, \dots, \lambda_k, \vec{x})$ is a critical point of L if and only if

$$0_{(k+n)\times 1} = \nabla L(\lambda_1, \dots, \lambda_k, \vec{x}) = \begin{bmatrix} -(g_1(\vec{x}) - c_1) \\ \vdots \\ -(g_k(\vec{x}) - c_k) \\ \nabla f(\vec{x}) - \lambda_1 \nabla g_1(\vec{x}) - \cdots - \lambda_k \nabla g_k(\vec{x}) \end{bmatrix}$$

which is equivalent to

$$\begin{cases} \nabla f(\vec{x}) = \lambda_1 \nabla g_1(\vec{x}) + \cdots + \lambda_k \nabla g_k(\vec{x}) \\ g_1(\vec{x}) = c_1 \\ \vdots \\ g_k(\vec{x}) = c_k \end{cases}$$

(b) Let $i, j = 1, \dots, k$ and $r, s = 1, \dots, n$. Then

$$L_{\lambda_i \lambda_j}(\lambda, \vec{x}) = 0, \quad L_{\lambda_i x_r}(\lambda, \vec{x}) = -(g_i)_{x_r}(\vec{x}) = L_{x_r \lambda_i}(\lambda, \vec{x}),$$

$$L_{x_r x_s}(\lambda, \vec{x}) = f_{x_r x_s}(\vec{x}) - \lambda_1 (g_1)_{x_r x_s}(\vec{x}) - \cdots - \lambda_k (g_k)_{x_r x_s}(\vec{x}).$$

Then $D^2L(\lambda, \vec{x})$ can be written (in block form) as

$$D^2L(\lambda, \vec{x}) = \begin{bmatrix} 0_{k\times k} & -D\vec{g}(\vec{x}) \\ -D\vec{g}(\vec{x})^T & D^2 f(\vec{x}) - \lambda_1 D^2 g_1(\vec{x}) - \cdots - \lambda_k D^2 g_k(\vec{x}) \end{bmatrix}$$

■

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 $f(x,y) \le g(x,y)$ for all $(x,y) \in R$, then
$$\iint_R f\,dA \le \iint_R g\,dA.$$
4. $|f|$ is also integrable on R and
$$\left|\iint_R f\,dA\right| \le \iint_R |f|\,dA.$$

(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.)

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

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

so that

$$\lim_{\|\mathcal{P}\|\to 0} R(cf, \mathcal{P}, \mathcal{C}) = \lim_{\|\mathcal{P}\|\to 0} cR(f, \mathcal{P}, \mathcal{C}) = c\int_R f\,dV_n.$$

The conclusion of Part 2 follows.

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

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

Because this inequality holds for every partition $\mathcal{P}$ and every choice of sample points $\mathcal{C}$, and because f and g are integrable on R (and therefore the limits of these Riemann sums as $\|\mathcal{P}\| \to 0$ exist), we have

$$\int_R f\,dV_n = \lim_{\|\mathcal{P}\|\to 0} R(f, \mathcal{P}, \mathcal{C}) \le \lim_{\|\mathcal{P}\|\to 0} R(g, \mathcal{P}, \mathcal{C}) = \int_R g\,dV_n.$$

(Part 4.) This part is a bit more sophisticated, as we must first argue that $|f|$ is integrable. Note that $\vec{x} \mapsto |f(\vec{x})|$ is the composition of $\vec{x} \mapsto f(\vec{x})$ and $t \mapsto |t|$. The latter function is continuous on $\mathbb{R}$, and (because f is integrable on RB) the former is continuous everywhere on R except on a set a measure zero (by Lebesgue's Criterion for Riemann Integrability). Because the composition of continuous functions is continuous, $\vec{x} \mapsto |f(\vec{x})|$ is continuous on R except on a set of measure zero, and is therefore intergrable on R.

Let $\mathcal{P}$ be a partition of R, and let $\mathcal{C}$ be a choice of sample points for $\mathcal{P}$. Then the triangle inequality (and the fact that $\mathrm{Vol}_n(R_i) \geq 0$ for each i) implies that

$$|R(f, \mathcal{P}, \mathcal{C})| = \left| \sum_i f(\vec{c}_i)\mathrm{Vol}_n(R_i) \right| \leq \sum_i |f(\vec{c}_i)|\mathrm{Vol}_n(R_i) = R(|f|, \mathcal{P}, \mathcal{C}).$$

By the same reasoning used in the proof of Part 3., we have

$$\left| \int_R f\, dV_n \right| = \lim_{\|\mathcal{P}\| \to 0} |R(f, \mathcal{P}, \mathcal{C})| \leq \lim_{\|\mathcal{P}\| \to 0} R(|f|, \mathcal{P}, \mathcal{C}) = \int_R |f|\, dV_n.$$

■

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

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

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

Note that because polynomials are continuous, f is continuous at each point (x, y) in the open set $\{(x, y) : \|(x, y)\| \neq 2\}$. Therefore f may only be discontinuous on the circle $C = \{(x, y) : \|(x, y)\| = 2\}$. Because C is the level set of the C^1 function $g(x, y) = x^2 + y^2$ with $g(x, y) = 4$, and because $\nabla g(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \neq \vec{0}$ on C, the Measure Zero Theorem implies that C has measure zero. Because the set of discontinuities of f in a rectangle of the form $[-a, a] \times [-b, b]$ is a (possibly empty) subset of the circle C, and therefore has measure zero by the Measure Zero Theorem, the Lebesgue Criterion for Riemann Integrability implies that f is integrable on $[-a, a] \times [-b, b]$. ■

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

Because a finite nonempty subset of $\mathbb{R}^n$ can be written as the finite union of singleton sets, by the Measure Zero Theorem it suffices to prove the result for a set of the form $A = \{\vec{a}\}$.

Let $\epsilon > 0$, and define $B = [a_1, a_1 + \frac{\sqrt[n]{\epsilon}}{2}] \times \cdots \times [a_n, a_n + \frac{\sqrt[n]{\epsilon}}{2}]$. Then $A \subset B$ and

$$\mathrm{Vol}_n(B) = \left(\frac{\sqrt[n]{\epsilon}}{2} \right)^n = \frac{\epsilon}{2^n} < \epsilon.$$

■