Exercise Problem Sets 10

May 02. 2024

Problem 1. Let $m > n$ be natural numbers, and A be an $m \times n$ real matrix, $\boldsymbol{b} \in \mathbb{R}^m$ be a vector.

(1) Show that if the minimum of the function $f(x_1, \cdots, x_n) = \|A\boldsymbol{x} - \boldsymbol{b}\|$ occurs at the point $\boldsymbol{c} = (c_1, \cdots, c_n)$, then $\boldsymbol{c}$ satisfies $A^{\mathrm{T}}A\boldsymbol{c} = A^{\mathrm{T}}\boldsymbol{b}$.

(2) Find the relation between the linear regression and (1).

Problem 2. Let $\{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$ be n points with $x_i \neq x_j$ if $i \neq j$. Use the Second Partials Test to verify that the formulas for a and b given by

$$a = \frac{n\sum\limits_{i=1}^{n} x_i y_i - \Big(\sum\limits_{i=1}^{n} x_i\Big)\Big(\sum\limits_{i=1}^{n} y_i\Big)}{n\sum\limits_{i=1}^{n} x_i^2 - \Big(\sum\limits_{i=1}^{n} x_i\Big)^2} \quad \text{and} \quad b = \frac{1}{n}\Big(\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i\Big)$$

indeed minimize the function $S(a, b) = \sum\limits_{i=1}^{n} (ax_i + b - y_i)^2$.

Proof. We compute the Hessian matrix of S and obtain that

$$\begin{bmatrix} S_{aa}(a,b) & S_{ab}(a,b) \\ S_{ba}(a,b) & S_{bb}(a,b) \end{bmatrix} = \begin{bmatrix} \sum\limits_{i=1}^{n} 2x_i^2 & \sum\limits_{i=1}^{n} 2x_i \\ \sum\limits_{i=1}^{n} 2x_i & \sum\limits_{i=1}^{n} 2 \end{bmatrix} = 2\begin{bmatrix} \sum\limits_{i=1}^{n} x_i^2 & \sum\limits_{i=1}^{n} x_i \\ \sum\limits_{i=1}^{n} x_i & n \end{bmatrix}.$$

As long as $\{(x_i, y_i) \,|\, 1 \leqslant i \leqslant n\}$ is not collinear, the Cauchy inequality implies that

$$\begin{vmatrix} S_{aa}(a,b) & S_{ab}(a,b) \\ S_{ba}(a,b) & S_{bb}(a,b) \end{vmatrix} = 4\Big[n\sum_{i=1}^{n} x_i^2 - \big(\sum_{i=1}^{n} x_i\big)^2\Big] > 0\,.$$

The fact that $S_{aa}(a, b) > 0$ shows that S attains a relative minimum at (a, b). Since S is differentiable on $\mathbb{R}^2$ and $S \geqslant 0$ and $\lim\limits_{b\to\infty} S(a, b) = \infty$, if S attains its extrema at some points, the absolute extremum must be an absolute minimum. Since an absolute minimum is also a relative minimum, we conclude that (a, b) given by the formula indeed minimizes S. □

Problem 3. The Shannon index (sometimes called the Shannon-Wiener index or Shannon-Weaver index) is a measure of diversity in an ecosystem. For the case of three species, it is defined as

$$H = -p_1 \ln p_1 - p_2 \ln p_2 - p_3 \ln p_3\,,$$

where p_i is the proportion of species i in the ecosystem.

(1) Express H as a function of two variables using the fact that $p_1 + p_2 + p_3 = 1$.

(2) What is the domain of H?

(3) Find the maximum value of H. For what values of p_1, p_2, p_3 does it occur?

Problem 4. Three alleles (alternative versions of a gene) A, B, and O determine the four blood types A (AA or AO), B (BB or BO), O (OO), and AB. The Hardy-Weinberg Law states that the proportion of individuals in a population who carry two different alleles is

$$P = 2pq + 2pr + 2rq\,,$$

where p, q, and r are the proportions of A, B, and O in the population. Use the fact that $p+q+r=1$ to show that P is at most $\frac{2}{3}$.

Problem 5. Find an equation of the plane that passes through the point $(1, 2, 3)$ and cuts off the smallest volume in the first octant.

Problem 6. Use the method of Lagrange multipliers to complete the following.

(1) Maximize $f(x, y) = \sqrt{6 - x^2 - y^2}$ subject to the constraint $x + y - 2 = 0$.

(2) Minimize $f(x, y) = 3x^2 - y^2$ subject to the constraint $2x - 2y + 5 = 0$.

(3) Minimize $f(x, y) = x^2 + y^2$ subject to the constraint $xy^2 = 54$.

(4) Maximize $f(x, y, z) = e^{xyz}$ subject to the constraint $2x^2 + y^2 + z^2 = 24$.

(5) Maximize $f(x, y, z) = \ln(x^2+1)+\ln(y^2+1)+\ln(z^2+1)$ subject to the constraint $x^2+y^2+z^2 = 12$.

(6) Maximize $f(x, y, z) = x + y + z$ subject to the constraint $x^2 + y^2 + z^2 = 1$.

(7) Maximize $f(x, y, z, t) = x + y + z + t$ subject to the constraint $x^2 + y^2 + z^2 + t^2 = 1$.

(8) Minimize $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraints $x + 2z = 6$ and $x + y = 12$.

(9) Maximize $f(x, y, z) = z$ subject to the constraints $x^2 + y^2 + z^2 = 36$ and $2x + y - z = 2$.

(10) Maximize $f(x, y, z) = yz + xy$ subject to the constraint $xy = 1$ and $y^2 + z^2 = 1$.

Problem 7. Use the method of Lagrange multipliers to find the extreme values of the function $f(x_1, x_2, \cdots, x_n) = x_1 + x_2 + \cdots + x_n$ subject to the constraint $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$.

Proof. Let $g(x_1, \cdots, x_n) = x_1^2+x_2^2+\cdots+x_n^2-1$. Suppose that f, under the constraint $g = 0$, attains its extrema at $(a_1, \cdots, a_n)$. Since $(\nabla g)(x_1, \cdots, x_n) = (x_1, \cdots, x_n)$ so that $(\nabla g)(x_1, \cdots, x_n) \neq 0$ if $g(x_1, \cdots, x_n) = 0$, by the Lagrange Multiplier Theorem there exists $\lambda \in \mathbb{R}$ such that

$$(1, \cdots, 1) = (\nabla f)(a_1, \cdots, a_n) = 2\lambda(\nabla g)(a_1, \cdots, a_n) = 2\lambda(a_1, \cdots, a_n)\,;$$

that is, $2\lambda a_j = 1$ for all $1 \leqslant j \leqslant n$. Then $\lambda \neq 0$; thus $a_j = \frac{1}{2\lambda}$ for all $1 \leqslant j \leqslant n$. Since $g(a_1, \cdots, a_n) = 0$, we find that

$$\frac{n}{4\lambda^2} = \sum_{i=1}^{n} a_i^2 = 1$$

which shows that $\lambda = \pm\frac{\sqrt{n}}{2}$ so that $a_j = \pm\frac{1}{\sqrt{n}}$ for all $1 \leqslant j \leqslant n$. At these two points, $f(\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}) = \sqrt{n}$ and $f(-\frac{1}{\sqrt{n}}, \cdots, -\frac{1}{\sqrt{n}}) = -\sqrt{n}$. Since the constraint $g = 0$ defines a closed and bounded set, f attains its maximum and minimum of the level set $g = 0$. Therefore, f attains its maximum and minimum at $(\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}})$ and $-(\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}})$, respectively, and the maximum is $\sqrt{n}$ and the minimum is $-\sqrt{n}$. □

Problem 8. Find the extreme value of $f(x, y, z) = z$ subject to the constraints $x^4 + y^4 - z^3 = 0$ and $y = z$.

Proof. Let $g(x, y, z) = x^4 + y^4 - z^3$ and $h(x, y, z) = y - z$. Then

$$(\nabla g)(x, y, z) = (4x^3, 4y^3, -3z^2) \qquad \text{and} \qquad (\nabla h)(x, y, z) = (0, 1, -1)$$

which implies that

$$(\nabla g)(x, y, z) \times (\nabla h)(x, y, z) = (3z^2 - 4y^3, 4x^3, 4x^3)\,.$$

Suppose the extreme value of f, under the constraints $g = h = 0$, occurs at (x_0, y_0, z_0).

1. If $(\nabla g)(x_0, y_0, z_0) \times (\nabla h)(x_0, y_0, z_0) = 0$, then $(x_0, y_0, z_0) = (0, 0, 0)$ and $f(0, 0, 0) = 0$.

2. If $(\nabla g)(x_0, y_0, z_0) \times (\nabla h)(x_0, y_0, z_0) \neq 0$, then the Lagrange Multiplier Theorem implies that there exist $\lambda, \mu \in \mathbb{R}$ such that

$$(\nabla f)(x_0, y_0, z_0) = \lambda(\nabla g)(x_0, y_0, z_0) + \mu(\nabla h)(x_0, y_0, z_0)\,.$$

Therefore, (x_0, y_0, z_0) satisfies that

$$4\lambda x_0^3 = 0\,, \tag{0.1a}$$
$$4\lambda y_0^3 + \mu = 0\,, \tag{0.1b}$$
$$-3\lambda z_0^2 - \mu = 1\,, \tag{0.1c}$$
$$x_0^4 + y_0^4 - z_0^3 = 0\,, \tag{0.1d}$$
$$y_0 - z_0 = 0\,. \tag{0.1e}$$

Then (0.1a) implies that $\lambda = 0$ or $x_0 = 0$.

(a) If $\lambda = 0$, then (0.1b) shows $\mu = 0$; thus using (0.1c), we obtain a contradiction $0 = -1$. Therefore, $\lambda \neq 0$.

(b) If $x_0 = 0$ (and $\lambda \neq 0$), then (0.1d) implies that $y_0^4 - z_0^3 = 0$. Together with (0.1e), we find that $y_0 = 0$ or $y_0 = 1$. However, if $y_0 = 0$, then (0.1b) shows that $\mu = 0$ which again implies a contradiction $0 = 1$ from (0.1c). Therefore, $y_0 = z_0 = 1$ (and there are λ, μ satisfying (0.1b,c) for $y_0 = z_0 = 1$ but the values of λ and μ are not important).

Therefore, the Lagrange Multiplier Theorem only provides one possible $(x_0, y_0, z_0) = (0, 1, 1)$ where f attains its extreme value.

Since the intersection of the level surface $g = 0$ and $h = 0$ is closed and bounded, f must attains its maximum and minimum subject to the constraints $g = h = 0$. Since $(0,0,0)$ and $(0,1,1)$ are the only possible points where f attains its extrema, the maximum and minimum of f, subject to the constraint $g = h = 0$, is $f(0,1,1) = 1$ and $f(0,0,0) = 0$, respectively. ▫

Problem 9. Let A be a full rank $m \times n$ real matrix, where $m < n$. and A have full rank. For a given $b \in \mathbb{R}^m$, show that the function f given by

$$f(x_1, \cdots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$$

under the constraint $Ax = b$, where $x = [x_1, \cdots, x_n]^{\mathrm{T}}$, attains its minimum at the point $A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b$.

Solution. For $1 \leqslant i \leqslant m$, let a_i denote the i-th column of A^{T} and b_i denote the i-th component of b. Then $Ax = b$ if and only if $a_i^{\mathrm{T}} x = b_i$ for all $1 \leqslant i \leqslant m$.

Let $g_i(x) = a_i^{\mathrm{T}} x - b_i$. Suppose that the function $f/2$, under the constraint $g = 0$, attains its extrema at $x_* = [x_1^*, \cdots, x_n^*]^{\mathrm{T}}$. Then by the fact that A has full rank, $\{\nabla g_i(x_*)\}_{i=1}^m = \{a_i\}_{i=1}^m$ is a linearly independent set. Therefore, there exist $\lambda_1, \cdots, \lambda_m \in \mathbb{R}$ such that

$$x_* = \frac{1}{2}(\nabla f)(x_*) = \sum_{i=1}^m \lambda_i a_i = A^{\mathrm{T}} \lambda_* \,,$$

where $\lambda_* = [\lambda_1, \cdots, \lambda_m]^{\mathrm{T}}$. Since A has full rank, the $m \times m$ matrix AA^{T} is non-singular; thus

$$x_* = A^{\mathrm{T}} \lambda_* \quad \Rightarrow \quad (AA^{\mathrm{T}})\lambda_* = Ax_* = b \quad \Rightarrow \quad \lambda_* = (AA^{\mathrm{T}})^{-1} b \,.$$

Therefore, $x_* = A^{\mathrm{T}}\lambda_* = A^{\mathrm{T}}(AA^{\mathrm{T}})^{-1}b$. Such x_* must be the point at which f attains its minimum since the maximum of f, subject to $Ax = b$, is ∞ since there are points far far away from the origin but satisfying $Ax = b$. ▫

Problem 10. (1) Use the method of Lagrange multipliers to show that the product of three positive numbers x, y, and z, whose sum has the constant value S, is a maximum when the three numbers are equal. Use this result to show that

$$\frac{x + y + z}{3} \geqslant \sqrt[3]{xyz} \qquad \forall\, x, y, z > 0 \,.$$

(2) Generalize the result of part (1) to prove that the product $x_1 x_2 x_3 \cdots x_n$ is maximized, under the constraint that $\sum\limits_{i=1}^n x_i = S$ and $x_i \geqslant 0$ for all $1 \leqslant i \leqslant n$, when

$$x_1 = x_2 = x_3 = \cdots = x_n \,.$$

Then prove that

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leqslant \frac{x_1 + x_2 + \cdots + x_n}{n} \qquad \forall\, x_1, x_2, \cdots, x_n \geqslant 0 \,.$$

Problem 11. (1) Maximize $\sum\limits_{i=1}^n x_i y_i$ subject to the constraints $\sum\limits_{i=1}^n x_i^2 = 1$ and $\sum\limits_{i=1}^n y_i^2 = 1$.

(2) Put $x_i = \dfrac{a_i}{\sqrt{\sum\limits_{j=1}^{n} a_j^2}}$ and $y_i = \dfrac{b_i}{\sqrt{\sum\limits_{j=1}^{n} b_j^2}}$ to show that

$$\sum_{i=1}^{n} a_i b_i \leqslant \sqrt{\sum_{j=1}^{n} a_j^2}\sqrt{\sum_{j=1}^{n} b_j^2}$$

for any numbers $a_1, a_2, \cdots, a_n$, $b_1, b_2, \cdots, b_n$. This inequality is known as the Cauchy-Schwarz Inequality.

Problem 12. Find the points on the curve $x^2 + xy + y^2 = 1$ in the xy-plane that are nearest to and farthest from the origin.

Problem 13. If the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ is to enclose the circle $x^2 + y^2 = 2y$, what values of a and b minimize the area of the ellipse?

Problem 14. (1) Use the method of Lagrange multipliers to prove that the rectangle with maximum area that has a given perimeter p is a square.

(2) Use the method of Lagrange multipliers to prove that the triangle with maximum area that has a given perimeter p is equilateral.

Hint: Use Heron's formula for the area:

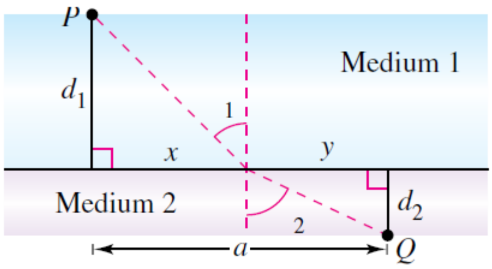
$$A = \sqrt{s(s-x)(s-y)(s-z)}\,,$$

where $s = \dfrac{p}{2}$ and x, y, z are the lengths of the sides.

Problem 15. When light waves traveling in a transparent medium strike the surface of a second transparent medium, they tend to "bend" in order to follow the path of minimum time. This tendency is called refraction and is described by Snell' s Law of Refraction,

$$\frac{\sin\theta_1}{\mathrm{v}_1} = \frac{\sin\theta_2}{\mathrm{v}_2}\,,$$

where θ_1 and θ_2 are the magnitudes of the angles shown in the figure, and v_1 and v_2 are the velocities of light in the two media. Use the method of Lagrange multipliers to derive this law using $x + y = a$.



Problem 16. A set $C \subseteq \mathbb{R}^n$ is said to be convex if

$$t\boldsymbol{x} + (1-t)\boldsymbol{y} \in C \qquad \forall\, \boldsymbol{x}, \boldsymbol{y} \in C \text{ and } t \in [0,1]\,.$$

(一個 $\mathbb{R}^n$ 中的集合 C 被稱為凸集合如果 C 中任兩點 $\boldsymbol{x}$ 與 $\boldsymbol{y}$ 之連線所形成的線段也在 C 中)。

Suppose that $C \subseteq \mathbb{R}^n$ is a convex set, and $f : C \to \mathbb{R}$ be a differentiable real-valued function. Show that if f on C attains its minimum at a point $\boldsymbol{x}^*$, then

$$(\nabla f)(\boldsymbol{x}^*) \cdot (\boldsymbol{x} - \boldsymbol{x}^*) \geqslant 0 \qquad \forall\, \boldsymbol{x} \in C\,. \tag{$\star$}$$

Hint: Recall that $(\nabla f)(\boldsymbol{x}^*) \cdot (\boldsymbol{x} - \boldsymbol{x}^*)$, when f is differentiable at x^*, is the directional derivative of f at $\boldsymbol{x}^*$ in the "direction" $(\boldsymbol{x} - \boldsymbol{x}^*)$.
Remark: A point $\boldsymbol{x}^*$ satisfying $(\star)$ is sometimes called a ***stationary point*** of f in C.

Problem 17. Let B be the unit closed ball centered at the origin given by

$$B = \left\{ \boldsymbol{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \,\middle|\, \|\boldsymbol{x}\|^2 = x_1^2 + x_2^2 + \cdots + x_n^2 \leqslant 1 \right\},$$

and $f : B \to \mathbb{R}$ be a differentiable real-valued function. Consider the minimization problem $\min\limits_{\boldsymbol{x} \in B} f(\boldsymbol{x})$.

(1) Show that if f attains its minimum at $\boldsymbol{x}^* \in B$, then there exists $\lambda \leqslant 0$ such that

$$(\nabla f)(\boldsymbol{x}^*) = \lambda \boldsymbol{x}^*\,.$$

(2) Find the minimum of the function $f(x, y) = x^2 + 2y^2 - x$ on the unit closed disk centered at the origin $\left\{(x, y) \,\middle|\, x^2 + y^2 \leqslant 1\right\}$ using (1).

Problem 18. Let $\boldsymbol{a} \in \mathbb{R}^3$ be a vector, $b \in \mathbb{R}$, and C be a half plane given by

$$C = \left\{ \boldsymbol{x} = (x_1, x_2, x_3) \in \mathbb{R}^3 \,\middle|\, \boldsymbol{a} \cdot \boldsymbol{x} \leqslant b \right\},$$

and $f : C \to \mathbb{R}$ be a differentiable real-valued function. Consider the minimization problem $\min\limits_{\boldsymbol{x} \in C} f(\boldsymbol{x})$. Show that if f attains its minimum at $\boldsymbol{x}^* \in C$, then there exists $\lambda \leqslant 0$ such that

$$(\nabla f)(\boldsymbol{x}^*) = \lambda \boldsymbol{a}\,.$$