MAXIMUM – MINIMUM PROBLEMS OF N-VARIABLE FUNCTIONS

PRELIMINARIES IN MATRIX THEORY
POSITIVE DEFINITE MATRICES

Let *P*,*S* be real symmetric matrices. If

$$y^T Py > 0 \qquad \forall y \neq 0$$

then *P* is called positive-definite matrix. If

$$y^T Sy \geq 0 \qquad \forall y \neq 0$$

then *S* is called positive semi-definite matrix.

Notes.

- If *P* is positive definite matrix then *P* is invertible.
- If *P* is positive definite and *S* is positive semi-definite then P+S is positive definite.
- The real symmetric matrix *P* is positive definite (semi-definite) iff the eigenvalues of *P* are positive (non negative).

NEGATIVE DEFINITE MATRICES

Let *P*,*S* be real symmetric matrices. If

$$y^T P y < 0 \qquad \forall y \neq 0$$

then *P* is called negative-definite matrix. If

$$y^T S y \leq 0 \qquad \forall y \neq 0$$

then *S* is called negative semi-definite matrix.

Notes.

- If *P* is negative definite matrix then *P* is invertible.
- If *P* is negative definite and *S* is negative semi-definite then P+S is negative definite.
- The real symmetric matrix *P* is negative definite (semi-definite) iff the eigenvalues of *P* are negative (non positive).

The matrix P

$$P = \begin{vmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{vmatrix}$$

is positive definite iff

$$p_{11} > 0, \begin{vmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{vmatrix} > 0, \begin{vmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{vmatrix} > 0$$

and negative definite iff

$$p_{11} < 0, \begin{vmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{vmatrix} > 0, \begin{vmatrix} p_{11} & p_{12} & p_{13} \\ p_{12} & p_{22} & p_{23} \\ p_{13} & p_{23} & p_{33} \end{vmatrix} < 0 \text{ (sign changes)}$$

In case of positive semi-definite (resp. negative semi-definite matrices) the inequality operators include the equal sign.

PRELIMINARY RESULTS IN CALCULUS THEORY

GRADIENT OF FUNCTIONS

Consider the function $f\left(\underbrace{x_1, x_2, \ldots, x_n}_{x}\right)$. We define the *gradient* of $f(x)$ with respect to *x* as

$$\nabla_x f(x) = \frac{df}{dx} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_m} \end{bmatrix} \quad \mathrm{x} = \left(\mathrm{x}_1, x_2, \ldots, x_m\right)^T$$

PROPERTIES OF THE GRADIENT OF FUNCTIONS

$$\frac{\partial\left(x^T c\right)}{\partial x} = \frac{\partial\left(x_1 c_1 + x_2 c_2 + \cdots + x_m c_m\right)}{\partial x} = c$$

$$\frac{\partial\left(x^T M x\right)}{\partial x} = \frac{\partial\left(x^T \underbrace{\left[Mx\right]}_{c_1}\right)}{\partial x} + \frac{\partial\left(\left(\underbrace{\left[M^T x\right]}_{c_2}^T\right) x\right)}{\partial x} =$$

$$= \frac{\partial\left(x^T \underbrace{\left[Mx\right]}_{c_1}\right)}{\partial x} + \frac{\partial\left(x^T \underbrace{\left[M^T x\right]}_{c_2}\right)}{\partial x} = Mx + M^T x$$

If M is a real symmetric matrix then $\frac{\partial\left(x^T M x\right)}{\partial x} = 2Mx$

HASSIENT MATRIX OF FUNCTIONS

We define the *hassient matrix* of $f\left(x\right)$ as

$$f_{xx} \triangleq \frac{\partial^2 f}{\partial x^2} = \left[\frac{\partial^2 f}{\partial x \partial x}\right] = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_m \partial x_1} & \frac{\partial^2 f}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_m^2} \end{bmatrix}$$

Properties

$$\frac{\partial^2\left(x^T M x\right)}{\partial x^2} = \frac{\partial\left(\left(M + M^T\right)x\right)}{\partial x} = \frac{\partial\left(x^T\left(M^T + M\right)\right)}{\partial x} = M + M^T$$

If M is a real symmetric matrix then $\dfrac{\partial^2\left(x^T M x\right)}{\partial x^2} = 2M$.

$$f = f\big(x(t), y(t)\big) \Rightarrow df = \left[\frac{\partial f}{\partial x}\right]^T dx + \left[\frac{\partial f}{\partial y}\right]^T dy$$

$$f = f\big(x(t), y(t), t\big), y\big(t\big) = y\Big(x\big(t\big), t\Big)$$

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial y^T}{\partial x}\frac{\partial f}{\partial y}$$

$$\frac{df}{dt} = \left[\frac{\partial f}{\partial x} + \frac{\partial y^T}{\partial x}\frac{\partial f}{\partial y}\right]^T \frac{dx}{dt} + \left[\frac{\partial f}{\partial y}\right]^T \frac{\partial y}{\partial t} + \frac{\partial f}{\partial t}$$

Taylor series of a function

$$f\left(x\right)= f\left(x_0\right)+\left[\frac{\partial f}{\partial x}\right]^T\Bigg|_{x=x_0}\left(x-x_0\right)+\frac{1}{2!}\left(x-x_0\right)^T\left[\frac{\partial^2 f}{\partial x^2}\right]\Bigg|_{x=x_0}\left(x-x_0\right)+O(3)$$

$$\Delta f\left(x\right)=\left[\frac{\partial f}{\partial x}\right]^T\Bigg|_{x=x_0} dx+\frac{1}{2!}dx^T\left[\frac{\partial^2 f}{\partial x^2}\right]\Bigg|_{x=x_0} dx+O(3)$$

RELATIVE MAXIMA-MINIMA

- A function $f: D \to \mathbb{R}, D \subseteq \mathbb{R}^n$ has a **relative maximum** $f(a)$ at the point $(a, f(a))$ if $f(a) \geq f(x)$ for all x in some region containing a.
- A function $f: D \to \mathbb{R}, D \subseteq \mathbb{R}^n$ has a **relative minimum** $f(a)$ at the point $(a, f(a))$ if $f(a) \leq f(x)$ for all x in some region containing a.

Theorem 1. (**Necessary conditions**) Suppose that $f(x)$ attains a relative minimum value (resp. relative maximum) at the point a and the gradient $\nabla_x f$ exist. Then

1) $\nabla_x f\big|_{x=a} = 0 \Leftrightarrow \left.\dfrac{\partial f}{\partial x_i}\right|_{x_i=a_i} = 0$

2) $K = \left[k_{ij}\right] = \left.\left[\dfrac{\partial^2 f(x)}{\partial x^2}\right]\right|_{x=a} = f_{xx}\big|_{x=a}$ is positive semi-definite (resp. negative semi-definite)

Sufficient condition : K is positive definite (resp. negative definite).

Notes.

- If $x^T Kx$ change signs at $x=a$ then a is a *saddle point*.
- If K is positive semi-definite or negative semi-definite then we need more information in order to design if the point is minimum or maximum (terms of order 3). The point a is called singular.

MAXIMUM – MINIMUM PROBLEMS OF TWO VARIABLE FUNCTIONS

Example 1. Find the relative minimum-maximum of

$$f(x,y) = x^2 + y^2 + z^2, \quad -\infty \le x, y, z \le +\infty$$

Since the partial derivatives $f_x = \dfrac{\partial f}{\partial x} = 2x$, $f_y = \dfrac{\partial f}{\partial y} = 2y$ and $f_z = \dfrac{\partial f}{\partial z} = 2z$ both exist and

1. $f_x = f_y = f_z = 0 \Leftrightarrow (x,y,z) = (0,0,0)$

2. $K = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} & \dfrac{\partial^2 f}{\partial x \partial z} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y \partial z} \\ \dfrac{\partial^2 f}{\partial z \partial x} & \dfrac{\partial^2 f}{\partial z \partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} > 0$

Thus we have a relative minimum at $(x,y,z) = (0,0,0)$.

$$f(x,y,z) - f(0,0,0) = x^2 + y^2 + z^2 - 0 = x^2 + y^2 + z^2 \ge 0 \Rightarrow f(x,y,z) \ge f(0,0,0)$$

Example 2. Find the relative minimum-maximum of

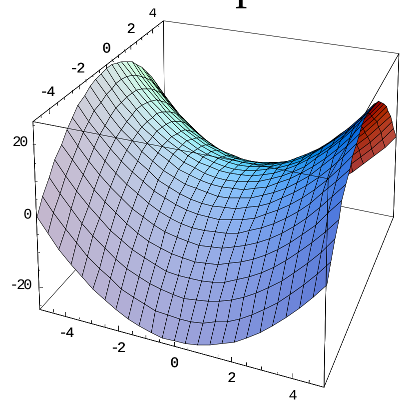
$$f(x,y) = x^2 - y^2, \qquad -\infty \le x \le +\infty$$

Since the partial derivatives $f_x = \dfrac{\partial f}{\partial x} = 2x$, $f_y = \dfrac{\partial f}{\partial y} = -2y$ both exist and

1. $f_x = f_y = 0 \Leftrightarrow (x,y) = (0,0)$

2. $K = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$

We have a saddle point at *(x,y)=(0,0)* since *K* has positive and negative eigenvalues.



CONSTRAINED OPTIMIZATION

Let $J(x,u):\mathbb{R}^n \to \mathbb{R}$ and $f:\mathbb{R}^n \to \mathbb{R}^m$, $f(x,u)=0$ where $u(t)\in\mathbb{R}^r$, $x(t)\in\mathbb{R}^n$.

Problem. Find the minimum-maximum of $J(x,u)$ under the constraints $f(x,u)=0$.

1rst solution. Solve the second equation and substitute in the function $J(x,u)$. Then apply the known criteria.

2nd solution. Lagrange multipliers

Form the function

$$L(x,u,\lambda) \triangleq J(x,u)+\lambda f(x,u), \lambda \triangleq (\lambda_1,\lambda_2,...,\lambda_m)$$

λ_k : Lagrange multipliers

$L(x,u,\lambda)$: Lagrangian

Necessary conditions

Then all the relative minimum and maximum points of $J\left(x,u\right)$ with x and y constrained to satisfy the equation $f\left(x,u\right)=0$ will be among those points $\left(x_0,u_0\right)$ for which $\left(x_0,u_0,\lambda_0\right)$ is a maximum or minimum point of $L\left(x,u,\lambda\right)$. These points $\left(x_0,u_0,\lambda_0\right)$ will be solutions of the system of simultaneous equations

$$L_x\left(x,u,\lambda\right)=0$$

$$L_u\left(x,u,\lambda\right)=0$$

$$L_\lambda\left(x,u,\lambda\right)=0 \quad \text{(this is just } f\left(x,u\right)=0\text{)}$$

Example. Find the point on the plane

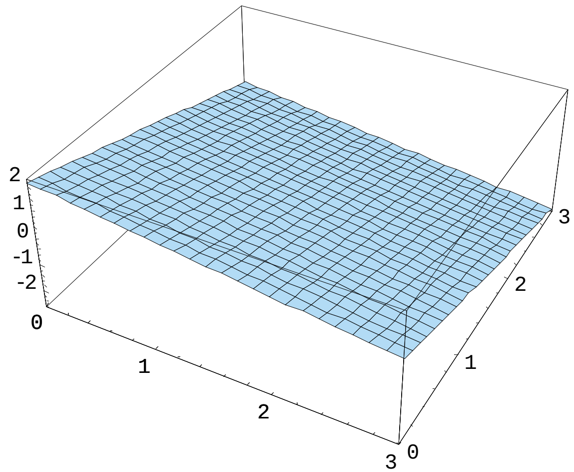
$$x + 2y + 2z = 4$$

that is closest to the origin or equivalently minimize

$$f\left(x, y, z\right) = x^2 + y^2 + z^2$$

subject to constraint

$$x + 2y + 2z = 4$$



Define the Lagrangian

$$L\left(x, y, z, \lambda\right) \triangleq x^2 + y^2 + z^2 + \lambda\left(x + 2y + 2z - 4\right)$$

The necessary conditions are

$$\left.\begin{aligned} \frac{\partial L}{\partial x} &= 2x + \lambda = 0 \\ \frac{\partial L}{\partial y} &= 2y + 2\lambda = 0 \\ \frac{\partial L}{\partial z} &= 2z + 2\lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= x + 2y + 2z - 4 = 0 \end{aligned}\right\} \Rightarrow \left\{x = \frac{4}{9}, y = \frac{8}{9}, z = \frac{8}{9}, l = -\frac{8}{9}\right\}$$

Sufficient conditions

Let

$$D_\phi(x) = \begin{pmatrix} \nabla^T_{\phi_1} \\ \nabla^T_{\phi_2} \\ \vdots \\ \nabla^T_{\phi_m} \end{pmatrix} = \left(\left(\frac{\partial f}{\partial x} \right) \quad \left(\frac{\partial f}{\partial u} \right) \right) \text{ the Jacobian of the constraints}$$

and

$$T_\phi(x) \triangleq \{ \xi : D_\phi(x)\xi = 0 \}$$

be the tangent plane at the point x on the surface defined by the constraints. Then

Theorem. Suppose $f(x), \phi_1(x), \phi_2(x), ..., \phi_m(x)$ have continuous second partial derivatives in $\mathbb{R}^n$, and let (x^*, λ^*) be a stationary point of the Lagrangian $L(x, \lambda)$. If

$$\xi^T L_{xx}(x^*, \lambda^*)\xi > 0, \forall \xi (\neq 0) \in T_\phi(x^*)$$

Then x^* is a strong local minimiser of $f(x)$ subject to constraints $\phi_1(x) = \phi_2(x) = ... = \phi_m(x) = 0$.

$$T_\phi(x) \triangleq \left\{\xi : D_\phi(x)\xi = 0\right\} = \left\{\xi : \left(\left(\frac{\partial f}{\partial x}\right) \quad \left(\frac{\partial f}{\partial u}\right)\right)\begin{pmatrix}\xi_1 \\ \xi_2\end{pmatrix} = 0\right\} = \begin{pmatrix} -\left(\frac{\partial f}{\partial x}\right)^{-1}\left(\frac{\partial f}{\partial u}\right)\xi_2 \\ \xi_2 \end{pmatrix}$$

$$\xi^T L_{xx}(x^*, \lambda^*)\xi = \xi_2^T \left(-\left(\frac{\partial f}{\partial u}\right)^T \left(\frac{\partial f}{\partial x}\right)^{-T} \quad I \right) L_{xx}(x^*, \lambda^*) \begin{pmatrix} -\left(\frac{\partial f}{\partial x}\right)^{-1}\left(\frac{\partial f}{\partial u}\right) \\ I \end{pmatrix} \xi_2 > 0, \forall \xi_2 (\neq 0)$$

Example. $$L(x,y,z,\lambda) \triangleq x^2+y^2+z^2+\lambda(x+2y+2z-4)$$

$$D_\phi(x)=\nabla^T_{\phi_1}=\begin{pmatrix}1 & 2 & 2\end{pmatrix}\ ;\ T_\phi(x)\triangleq\left\{\xi:\begin{pmatrix}1 & 2 & 2\end{pmatrix}\xi=0\right\}=\begin{pmatrix}-2\xi_2-2\xi_3\\ \xi_2\\ \xi_3\end{pmatrix}$$

$$L_{xx}=\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}>0 \text{ or}$$

$$\xi^T L_{xx}(x^*,\lambda^*)\xi=\begin{pmatrix}-2\xi_2-2\xi_3 & \xi_2 & \xi_3\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{pmatrix}\begin{pmatrix}-2\xi_2-2\xi_3\\ \xi_2\\ \xi_3\end{pmatrix}=$$

$$=2(-2\xi_2-2\xi_3)^2+2(\xi_2)^2+2(\xi_3)^2>0,\forall\xi\neq 0$$

Therefore $\left\{x=\frac{4}{9},y=\frac{8}{9},z=\frac{8}{9},l=-\frac{8}{9}\right\}$ minimize the function $f(x,y,z)=x^2+y^2+z^2$ subject to constraint $x+2y+2z=4$.

Exersizes

1. Find the minimum-maximum of the function $J(x,y)=xy$ subject to the constraint $f(x,y)=\frac{x^2}{4}+\frac{y^2}{9}-1=0$.
2. Find the minimum of the function

$$J(x,u)=\frac{1}{2}x^TQx+\frac{1}{2}u^TRu$$

$$x\in R^n, u\in R^r, f\in R^n, Q\geq 0, R>0$$

where Q and R are symmetric matrices, and x,u are subject to the constraint

$$f(x,u)=x+Bu+c=0$$