V-2-6. Maximum and minimum

V-2-6-1. Maximum and minimum in quadratic from maximum and minimum

When the quadratic form is defined by a positive real number as follow, we can draw the locus of arrow head of the vector $\boldsymbol{x}$ as shown in V-2-3. "Diagonalization and spectral decomposition of quadratic from". This means that the absolute value (length) of vector is restricted by the equation

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x} = d$$

Depending on the shape of quadratic form, the length of the vector has extreme value in several cases. In the case when the matrix is positive definite, the shape of the locus is hyperelliptic. The longest vector is the longest radius, and the maximum value of length of vector is length of longest radius. The shortest vector is shortest radius, and the minimum value of the length of vector is length of shortest radius. In the case of indefinite, shape of the locus is not simple, and we cannot discuss maximum and minimum of the length of vector without other domain of definition in many cases. When

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{1}$$

And

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{\Lambda}\boldsymbol{X}$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}$$

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$$

$$\frac{{X_1}^2}{\left(\frac{1}{\sqrt{\lambda_1}}\right)^2} + \frac{{X_2}^2}{\left(\frac{1}{\sqrt{\lambda_2}}\right)^2} + \cdots + \frac{{X_p}^2}{\left(\frac{1}{\sqrt{\lambda_p}}\right)^2} = 1$$

This is equation of hyperelliptic. Figure 58 is an example elliptic

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{1}$$

And

$$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{X}^{\boldsymbol{T}}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\boldsymbol{X}$$

$$\frac{{X_1}^2}{\left(\frac{1}{\sqrt{\lambda_1}}\right)^2} + \frac{{X_2}^2}{\left(\frac{1}{\sqrt{\lambda_2}}\right)^2} = 1$$

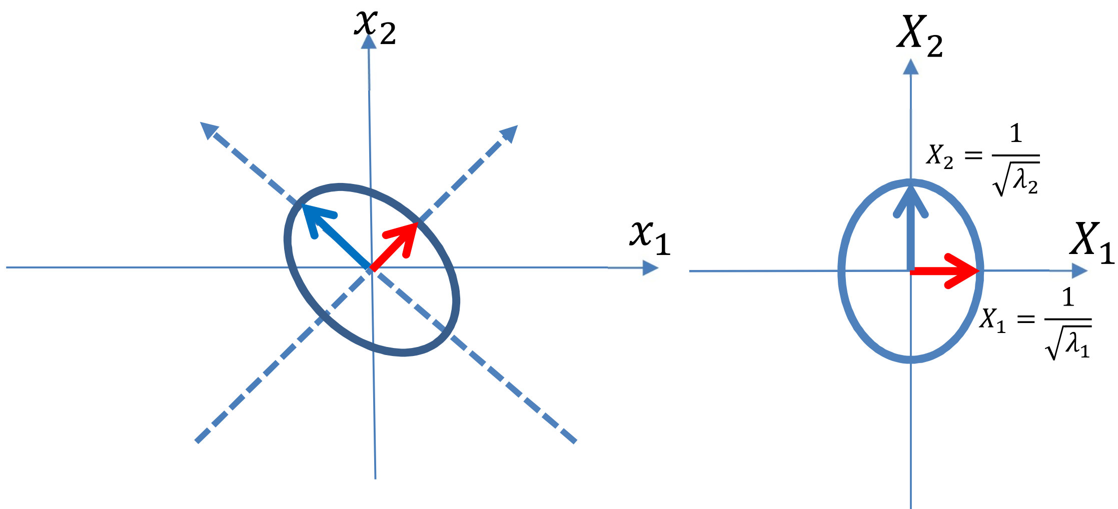

Fig. 58. An example of elliptic

From this figure, we can understand that the maximum length of vector, the absolute value of $\boldsymbol{x}$ is $\frac{1}{\sqrt{\lambda_2}}$ and the minimum length of vector $\boldsymbol{x}$ is $\frac{1}{\sqrt{\lambda_1}}$ when the matrix is positive definite.

From this, we can conclude that the minimum and maximum of following value is smallest eigen value and largest eigenvalue, when matrix $\boldsymbol{A}$ is positive definite.

$$\lambda_{smallest} \leq \frac{\boldsymbol{x^T A x}}{\boldsymbol{x^T x}} \leq \lambda_{largest} y$$

The author supposes that many readers can accept upper explanation intuitively. However, some readers cannot accept the explanation because of its too simple logic. For such readers, the author adds supplemental explanation.

Meaning of $\frac{\boldsymbol{x^T A x}}{\boldsymbol{x^T x}}$ is ratio of $\boldsymbol{x^T A x}$ and $\boldsymbol{x^T x}$. The numerator, $\boldsymbol{x^T A x}$, is magnitude of hyperelliptic, and $\boldsymbol{x^T x}$ is square of length of vector $\boldsymbol{x}$. We are discussing the maximum and minimum value of the ratio and when the ratio is maximum and minimum value. Here, $\boldsymbol{A}$ is symmetric and positive definite.

Diagonalization

$$\boldsymbol{P}^{-1}\boldsymbol{AP} = \boldsymbol{\Lambda}$$

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}$$

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$$

$$\boldsymbol{A}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^{-1}$$

、$\boldsymbol{A}$ is symmetric

$$\boldsymbol{P}^{-1} = \boldsymbol{P}^T$$

$$\boldsymbol{B}^{\frac{1}{2}} = \boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T$$

Here

$$\boldsymbol{X} = \boldsymbol{P}^T\boldsymbol{x}$$

$$\boldsymbol{x} = \boldsymbol{P}\boldsymbol{X}$$

$$\frac{\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}} = \frac{\boldsymbol{x}^T\boldsymbol{B}\boldsymbol{x}}{(\boldsymbol{P}\boldsymbol{X})^T\boldsymbol{P}\boldsymbol{X}} = \frac{\boldsymbol{x}^T\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{X}} = \frac{\boldsymbol{x}^T\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}} = \frac{\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}} = \frac{\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{P}^T\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}}$$

$$= \frac{\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T\boldsymbol{x}}{\boldsymbol{X}^T\boldsymbol{X}} = \frac{\boldsymbol{X}^T\boldsymbol{\Lambda}\boldsymbol{X}}{\boldsymbol{X}^T\boldsymbol{X}}$$

$$\frac{\boldsymbol{X}^T\boldsymbol{\Lambda}\boldsymbol{X}}{\boldsymbol{X}^T\boldsymbol{X}} = \frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} \lambda_i {X_i}^2}{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} {X_i}^2}$$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$

$$\frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} \lambda_i {y_i}^2}{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} {y_i}^2} \leq \lambda_1 \frac{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} {y_i}^2}{\sum_{\boldsymbol{i=1}}^{\boldsymbol{p}} {y_i}^2} = \lambda_1$$

Spectral decomposition

$$\boldsymbol{A} = \lambda_1\boldsymbol{e_1}\boldsymbol{e_1}^T + \lambda_2\boldsymbol{e_2}\boldsymbol{e_2}^T + \cdots + \lambda_p\boldsymbol{e_p}\boldsymbol{e_p}^T$$

$$\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} = \lambda_1\boldsymbol{x}^T\boldsymbol{e_1}\boldsymbol{e_1}^T\boldsymbol{x} + \lambda_2\boldsymbol{x}^T\boldsymbol{e_2}\boldsymbol{e_2}^T\boldsymbol{x} + \cdots + \lambda_p\boldsymbol{x}^T\boldsymbol{e_p}\boldsymbol{e_p}^T\boldsymbol{x}$$

Assigning $\boldsymbol{e}_1$ to $\boldsymbol{x}$

$$\boldsymbol{e}_1{}^T\boldsymbol{A}\boldsymbol{e} = \lambda_1\boldsymbol{e}_1{}^T\boldsymbol{e_1}\boldsymbol{e_1}^T\boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_1{}^T\boldsymbol{e_2}\boldsymbol{e_2}^T\boldsymbol{e}_1 + \cdots + \lambda_p\boldsymbol{e}_1{}^T\boldsymbol{e_p}\boldsymbol{e_p}^T\boldsymbol{e}_1$$

$\boldsymbol{e}_i$ and $\boldsymbol{e}_j$ is orthogonal each other

$$\boldsymbol{e}_i{}^T\boldsymbol{e}_j = 0 \quad i \neq j$$

$$\boldsymbol{e}_i{}^T\boldsymbol{e}_i = 1$$

$$\boldsymbol{e}_1{}^T\boldsymbol{A}\boldsymbol{e} = \lambda_1\boldsymbol{e}_1{}^T\boldsymbol{e_1}\boldsymbol{e_1}^T\boldsymbol{e}_1 = \lambda_1$$

Conclusion

When symmetric matrix $\boldsymbol{A}$ is positive definite

$$\max_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}} = \lambda_1$$

Similarly,

$$\min_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}} = \lambda_p$$

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{pmatrix}$$

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$$

Formula 63

V-2-6-2. Cauchy Schwarz' s inequation

So far, we do not know how to use this conclusion. We will use this in later paragraph of this text book. Before that, the author adds several complemental interpretations for understanding or the meaning of this inequality.

Cauchy Schwarz's inequation is often used in optimization. This inequation has various forms, though most general expression is as follow.

$$\left(a_1b_1 + a_2b_2 + \cdots + a_pb_p\right)^2 \leq \left({a_1}^2 + {a_2}^2 + \cdots + {a_p}^2\right)\left({b_1}^2 + {b_2}^2 + \cdots + {b_p}^2\right)$$

Formula 64

The most simple and elegant proof of the inequality is to change the inequation to vector form.

Denoting vector $\boldsymbol{a}$ and $\boldsymbol{b}$

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix}$$

Inner products of the vectors

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$$

$$-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$$

$$0 \leq \cos\theta \leq 1$$

Then

$$\boldsymbol{a} \cdot \boldsymbol{b} \leq |\boldsymbol{a}||\boldsymbol{b}| \qquad \text{i}$$

$$\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = \begin{pmatrix} a_1 & a_2 & \cdots & a_p \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_pb_p$$

$$|\boldsymbol{a}| = \sqrt{{a_1}^2 + {a_2}^2 + \cdots + {a_p}^2}, \qquad |\boldsymbol{b}| = \sqrt{{b_1}^2 + {b_2}^2 + \cdots + {b_p}^2}$$

Put these equations to i

$$a_1b_1 + a_2b_2 + \cdots + a_pb_p \le \sqrt{a_1{}^2 + a_2{}^2 + \cdots + a_p{}^2}\sqrt{b_1{}^2 + b_2{}^2 + \cdots + b_p{}^2}$$

Raise both side to the second power

$$\left(a_1b_1 + a_2b_2 + \cdots + a_pb_p\right)^2 \le \left(a_1{}^2 + a_2{}^2 + \cdots + a_p{}^2\right)\left(b_1{}^2 + b_2{}^2 + \cdots + b_p{}^2\right)$$

This is very sophisticated proof, as it is used only a trivial rule that inner products of the vectors are smaller than products of length of vectors.

Example of expansion of Cauchy Schwarz's inequation (1)

Letting

$$\boldsymbol{a} = \left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$\boldsymbol{b} = \left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)$$

$\boldsymbol{B}$ is symmetric matrix

Put $\boldsymbol{a}$ and $\boldsymbol{b}$ in Cauchy Schwarz's inequation.

$$(\boldsymbol{a}^T\boldsymbol{a})(\boldsymbol{b}^T\boldsymbol{b}) \ge (\boldsymbol{a}^T\boldsymbol{b})^{\boldsymbol{2}}$$

$$\text{Left side} = \left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)^T\left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)^T\left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)$$

$$= \left(\boldsymbol{\alpha}^T\boldsymbol{B}^{\frac{1}{2}^T}\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{\beta}'\boldsymbol{B}^{-\frac{1}{2}^{\boldsymbol{T}}}\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)$$

$$= (\boldsymbol{\alpha}^T\boldsymbol{B}\boldsymbol{\alpha})(\boldsymbol{\beta}^T\boldsymbol{B}^{-\boldsymbol{1}}\boldsymbol{\beta})$$

$$\text{Right side} = \left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)^T\left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)^T\left(\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)$$

$$= \left(\boldsymbol{\alpha}^T\boldsymbol{B}^{\frac{1}{2}^T}\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)\left(\boldsymbol{\alpha}^T\boldsymbol{B}^{\frac{1}{2}^T}\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}\right)$$

$$= (\boldsymbol{\alpha}^T\boldsymbol{I}\boldsymbol{\beta})(\boldsymbol{\alpha}^T\boldsymbol{I}\boldsymbol{\beta})$$

$$= (\boldsymbol{\alpha}^T\boldsymbol{\beta})^2$$

$$(\boldsymbol{\alpha}^T\boldsymbol{B}\boldsymbol{\alpha})(\boldsymbol{\beta}^T\boldsymbol{B}^{-\boldsymbol{1}}\boldsymbol{\beta}) \ge (\boldsymbol{\alpha}^{\boldsymbol{T}}\boldsymbol{\beta})^2$$

When

$$\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha} = c\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}$$

$$(\boldsymbol{\alpha}^T\boldsymbol{B}\boldsymbol{\alpha})(\boldsymbol{\beta}^T\boldsymbol{B}^{-\boldsymbol{1}}\boldsymbol{\beta}) = (\boldsymbol{\alpha}^{\boldsymbol{T}}\boldsymbol{\beta})^2$$

Formula 65

In this inequation, $\boldsymbol{\alpha}^T\boldsymbol{B}\boldsymbol{\alpha}$ is scalar and $\boldsymbol{B}$ is positive definite. We can divide both sides without change of inequality sign.

$$\frac{(\boldsymbol{\alpha}^T\boldsymbol{\beta})^2}{\boldsymbol{\alpha}^T\boldsymbol{B}\boldsymbol{\alpha}} \leq \boldsymbol{\beta}^T\boldsymbol{B}^{-1}\boldsymbol{\beta}$$

By this transformation, we can separate the inequality to two parts. Left side is function of $\boldsymbol{\alpha}$, and right side is function of $\boldsymbol{\beta}$.

$$\mathbf{F}(\boldsymbol{\alpha}) = \frac{(\boldsymbol{\alpha}^T\boldsymbol{\beta})^2}{\boldsymbol{\alpha}^T\boldsymbol{B}\boldsymbol{\alpha}}$$

$$\max_{\boldsymbol{\alpha}\neq 0}\mathbf{F}(\boldsymbol{\alpha}) = \boldsymbol{\beta}^T\boldsymbol{B}^{-1}\boldsymbol{\beta}$$

When

$$\boldsymbol{B}^{\frac{1}{2}}\boldsymbol{\alpha} = c\boldsymbol{B}^{-\frac{1}{2}}\boldsymbol{\beta}$$

$$\boldsymbol{\alpha} = c\boldsymbol{B}^{-1}\boldsymbol{\beta}$$

For clear specification of variable, we denote $\boldsymbol{\alpha}$ as $\boldsymbol{x}$

We presume that symmetric matrix $\boldsymbol{B}$ is positive definite.

$$\max_{x\neq 0}\frac{(\boldsymbol{x}'\boldsymbol{\beta})^2}{\boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}} = \boldsymbol{\beta}'\boldsymbol{B}^{-1}\boldsymbol{\beta}$$

When

$$\boldsymbol{x} = c\boldsymbol{B}^{-1}\boldsymbol{\beta}$$

Example of expansion of Cauchy Schwarz's inequation (2)

Letting

$$\boldsymbol{a} = \left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$\boldsymbol{b} = \left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$\boldsymbol{B}$ is symmetric matrix

Put $\boldsymbol{a}$ and $\boldsymbol{b}$ in Cauchy Schwarz's inequation.

$$(\boldsymbol{a}^T\boldsymbol{a})(\boldsymbol{b}^T\boldsymbol{b}) \geq (\boldsymbol{a}^T\boldsymbol{b})^{\mathbf{2}}$$

$$\text{Left side} = \left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)^T\left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{\alpha}\right)^T\left(\boldsymbol{G}^{-\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$= \left(\boldsymbol{\alpha}^T\boldsymbol{F}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\left(\boldsymbol{\alpha}^T\boldsymbol{G}^{\frac{1}{2}^{\boldsymbol{T}}}\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$= (\boldsymbol{\alpha}^T\boldsymbol{F}\boldsymbol{\alpha})(\boldsymbol{\alpha}^T\boldsymbol{G}\boldsymbol{\alpha})$$

$$\text{Right side} = \left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)^T\left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right)$$

$$= \boldsymbol{\alpha}^T\boldsymbol{G}^{\frac{1}{2}^T}\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}$$

Combining both side

$$(\boldsymbol{\alpha}^T \boldsymbol{F}\boldsymbol{\alpha})(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha}) \geq \boldsymbol{\alpha}^T \boldsymbol{G}^{\frac{1}{2}^T} \boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}$$

Dividing both side by $(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})^2$ $\quad ((\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})^2 > \boldsymbol{0})$

$$\frac{(\boldsymbol{\alpha}^T \boldsymbol{F}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})} \geq \frac{\boldsymbol{\alpha}^T \boldsymbol{g}^{\frac{1}{2}^T} \boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}}{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})^2}$$

Condition of equality

$$\boldsymbol{a} = c\boldsymbol{b}$$

$$\left(\boldsymbol{a} = \left(\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}\right), \quad \boldsymbol{b} = \left(\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}\right)\right)$$

$$\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha} = c\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}$$

When

$$\boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha} = c\boldsymbol{G}^{\frac{1}{2}}\boldsymbol{\alpha}$$

$$\frac{(\boldsymbol{\alpha}^T \boldsymbol{F}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})} \text{ is minimum value}$$

$$\frac{(\boldsymbol{\alpha}^T \boldsymbol{F}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})} = \frac{\boldsymbol{\alpha}^T \boldsymbol{g}^{\frac{1}{2}^T} \boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}}{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})^2}$$

$$\frac{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T \boldsymbol{F}\boldsymbol{\alpha})} \text{ is maximum value}$$

$$\frac{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})}{(\boldsymbol{\alpha}^T \boldsymbol{F}\boldsymbol{\alpha})} = \frac{(\boldsymbol{\alpha}^T \boldsymbol{G}\boldsymbol{\alpha})^2}{\boldsymbol{\alpha}^T \boldsymbol{g}^{\frac{1}{2}^T} \boldsymbol{F}^{\frac{1}{2}}\boldsymbol{\alpha}}$$

V-2-6-3. Method of Lagrange multiplier

The exemplified two cases of expansion of Cauchy Schwarz's inequation in upper paragraphs. Those cases can be interpreted as particular cases of method of Lagrange multiplier. Method of Lagrange multipliers is a method to calculate maximum or minimum value under constrained conditions. Figure 59 is illustration of theoretical background of method of Lagrange multipliers.

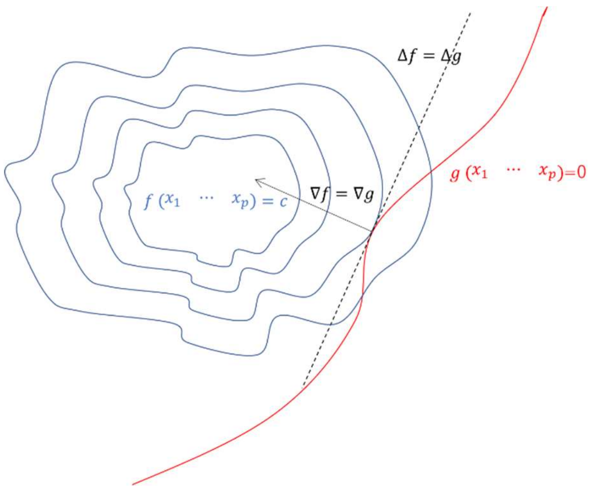

Fig. 59. Relation between minimizing function and constrained condition.

Red line is constrained condition $g(x_1 \quad \cdots \quad x_p) = 0$, and shapes written by blue line is target function of minimization $f(x_1 \quad \cdots \quad x_p) = c$. When we increase c, the area enclosed by blue line expand and the blue line will contact red line at a particular c. After that blue line will have intersections with red line. When blue line has intersections or tangent point, blue line can satisfy the constrained condition. The shape of the blue line is not simple and there will be several tangent points between blue line and red line with increase of c, however c at first tangent point is minimum c which satisfy constrained condition. In the case, when we consider maximization of function $f(x_1 \quad \cdots \quad x_p) = c$. we put $f(x_1 \quad \cdots \quad x_p) = c$ inside of $g(x_1 \quad \cdots \quad x_p) = 0$. The last point where blue line and red line has tangent point will give the maximum value of c. At tangent point blue line and red line share tangent plane ($\Delta f = \lambda \Delta g$) and normal vector($\nabla f = \lambda \nabla g$). Normal vector is expressing gradient of hyperplane and the other name of the normal vector is gradient vector of the plane.

Gradient of function $f(x_1 \quad \cdots \quad x_p)$ can be obtained by partial differentiation.

$$\nabla f = \frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial f(x_1 \quad \cdots \quad x_p)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x_1 \quad \cdots \quad x_p)}{\partial x_p} \end{pmatrix}$$

$$\lambda \nabla g = \lambda \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = \lambda \begin{pmatrix} \frac{\partial g(x_1 \quad \cdots \quad x_p)}{\partial x_1} \\ \vdots \\ \frac{\partial g(x_1 \quad \cdots \quad x_p)}{\partial x_p} \end{pmatrix}$$

For sharing normal vector

$$\nabla f = \lambda \nabla g$$

$$\nabla f - \lambda \nabla g = 0$$

$$\frac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}} - \lambda \frac{\partial g(\boldsymbol{x})}{\partial \boldsymbol{x}} = 0$$

$$\frac{\partial\big(f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})\big)}{\partial \boldsymbol{x}} = 0$$

Denoting $L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$

General description of method of Lagrange multiplier is as follow.

$$\frac{\partial L(\boldsymbol{x}, \lambda)}{\partial (\boldsymbol{x}, \lambda)} = \begin{pmatrix} \frac{\partial L(\boldsymbol{x}, \lambda)}{\partial x_1} \\ \vdots \\ \frac{\partial L(\boldsymbol{x}, \lambda)}{\partial x_p} \\ \frac{\partial L(\boldsymbol{x}, \lambda)}{\partial \lambda} \end{pmatrix}$$

Here

$$L(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$$

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x_1} \\ \vdots \\ \boldsymbol{x_p} \end{pmatrix}$$

Formula 66

Example solutions

1. Obtain extreme value of $x_1 + x_2 + x_3$, when $x_1{}^2 + x_2{}^2 + x_3{}^2 = 1$

   Extreme value of $f(x_1 \quad x_2 \quad x_3) = x_1 + x_2 + x_3$

   Subject to $g(x_1 \quad x_2 \quad x_3) = x_1{}^2 + x_2{}^2 + x_3{}^2 - 1 = 0$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$L(x_1 \quad x_2 \quad x_3 \quad \lambda) = x_1 + x_2 + x_3 - \lambda(x_1{}^2 + x_2{}^2 + x_3{}^2 - 1)$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial \boldsymbol{x}}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial x_1} = 1 - 2\lambda x_1 = 0 \qquad \text{i}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial x_2} = 1 - 2\lambda x_2 = 0 \qquad \text{ii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial x_3} = 1 - 2\lambda x_3 = 0 \qquad \text{iii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda)}{\partial \lambda} = -({x_1}^2 + {x_2}^2 + {x_3}^2 - 1) = 0 \qquad \text{iv}$$

From i, ii and iii,

$$x_1 = \frac{1}{2\lambda} \qquad \text{i}'$$

$$x_2 = \frac{1}{2\lambda} \qquad \text{ii}'$$

$$x_3 = \frac{1}{2\lambda} \qquad \text{iii}'$$

Put i', ii' and iii' in iv

$$\left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{2\lambda}\right)^2 + \left(\frac{1}{2\lambda}\right)^2 - 1 = 0$$

$$\frac{3}{4\lambda^2} = 1$$

$$\lambda^2 = \frac{3}{4}$$

$$\lambda = \pm\frac{\sqrt{3}}{2}$$

Put this in i', ii' and iii'

$$x_1 = \pm\frac{2}{2\sqrt{3}} = \pm\frac{1}{\sqrt{3}} = \pm\frac{\sqrt{3}}{3}$$

$$x_2 = \pm\frac{2}{2\sqrt{3}} = \pm\frac{1}{\sqrt{3}} = \pm\frac{\sqrt{3}}{3}$$

$$x_3 = \pm\frac{2}{2\sqrt{3}} = \pm\frac{1}{\sqrt{3}} = \pm\frac{\sqrt{3}}{3}$$

λ	$-\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$
x_1	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
x_2	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
x_3	$-\frac{\sqrt{3}}{3}$	$\frac{\sqrt{3}}{3}$
$x_1 + x_2 + x_3$	$-\sqrt{3}$	$\sqrt{3}$

2. Obtain extreme value of $x_1 + x_2 + x_3$, when $x_1{}^2 + x_2{}^2 = 1$, and $x_2{}^2 + x_3{}^2 = 1$

Extreme value of $f(x_1 \quad x_2 \quad x_3)$=$x_1 + x_2 + x_3$

Subject to $g_1(x_1 \quad x_2 \quad x_3)$=$x_1{}^2 + x_2{}^2 - 1 = 0$

$g_2(x_1 \quad x_2 \quad x_3)$=$x_2{}^2 + x_3{}^2 - 1 = 0$

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2) = x_1 + x_2 + x_3 - \lambda_1(x_1{}^2 + x_2{}^2 - 1) - \lambda_2(x_2{}^2 + x_3{}^2 - 1)$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial \boldsymbol{x}}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial x_1} = 1 - 2\lambda_1 x_1 = 0 \qquad \text{i}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial x_2} = 1 - 2(\lambda_1 + \lambda_2)x_2 = 0 \qquad \text{ii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial x_3} = 1 - 2\lambda_2 x_3 = 0 \qquad \text{iii}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial \lambda_1} = x_1{}^2 + x_2{}^2 - 1 = 0 \qquad \text{iv}$$

$$\frac{\partial L(x_1 \quad x_2 \quad x_3 \quad \lambda_1 \quad \lambda_2)}{\partial \lambda_1} = x_2{}^2 + x_3{}^2 - 1 = 0 \qquad \text{v}$$

From i, ii and iii,

$$x_1 = \frac{1}{2\lambda_1} \qquad \text{i}'$$

$$x_2 = \frac{1}{2(\lambda_1 + \lambda_2)} \qquad \text{ii}'$$

$$x_3 = \frac{1}{2\lambda_2} \qquad \text{iii}'$$

$$g_1(x_1 \quad x_2 \quad x_3) - g_2(x_1 \quad x_2 \quad x_3) = (x_1{}^2 + x_2{}^2 - 1) - (x_2{}^2 + x_3{}^2 - 1) = x_1{}^2 - x_3{}^2 = 0$$

$$x_1{}^2 - x_3{}^2 = (x_1 + x_3)(x_1 - x_3) = 0$$

$$x_1 = x_3 \text{ or } x_1 = -x_3$$

When $x_1 = -x_3$, $\lambda_1 = -\lambda_2$ and $\lambda_1 + \lambda_2$=0.

$(\lambda_1 + \lambda_2)$ is dominator of ii$'$. So, we could not accept $x_1 = -x_3$

$x_1 = x_3, \quad \lambda_1 = -\lambda_2 = \lambda$

Put this in i$'$ and ii’

$$x_1 = \frac{1}{2\lambda_1} = \frac{1}{2\lambda} \qquad \text{i}''$$

$$x_2 = \frac{1}{2(\lambda_1 + \lambda_2)} = \frac{1}{4\lambda} \qquad \text{ii}''$$

Put i″ and ii" in iv.

$$x_1{}^2 + x_2{}^2 - 1 = \frac{1}{4\lambda^2} + \frac{1}{16\lambda^2} - 1 = \frac{5}{16\lambda^2} - 1 = 0 \qquad \text{iv}$$

$$\lambda^2 = \frac{5}{16}$$

$$\lambda = \pm\frac{\sqrt{5}}{4}$$

Put this in i', ii' and iii'

$$x_1 = \pm\frac{1}{2\frac{\sqrt{5}}{4}} = \pm\frac{2}{\sqrt{5}}$$

$$x_2 = \pm\frac{1}{4\frac{\sqrt{5}}{4}} = \pm\frac{1}{\sqrt{5}}$$

$$x_3 = \pm\frac{1}{2\frac{\sqrt{5}}{4}} = \pm\frac{2}{\sqrt{5}}$$

λ	$-\frac{\sqrt{5}}{4}$	$\frac{\sqrt{5}}{4}$
x_1	$-\frac{2}{\sqrt{5}}$	$\frac{2}{\sqrt{5}}$
x_2	$-\frac{1}{\sqrt{5}}$	$\frac{1}{\sqrt{5}}$
x_3	$-\frac{2}{\sqrt{5}}$	$\frac{2}{\sqrt{5}}$
$x_1 + x_2 + x_3$	$-\sqrt{5}$	$\sqrt{5}$