Basics for Mathematical Economics

Lagrange multipliers

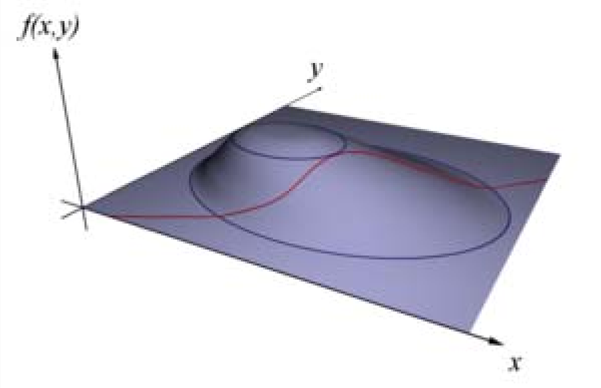


Figure 1: Find x and y to maximize $f(x,y)$ subject to a constraint (shown in red) $g(x,y) = c$.

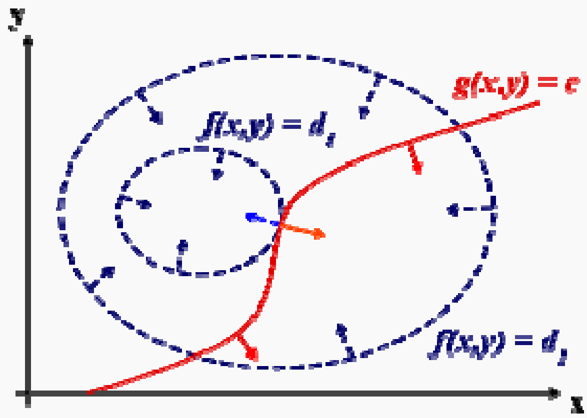


Figure 2: Contour map of Figure 1. The red line shows the constraint $g(x,y) = c$. The blue lines are contours of $f(x,y)$. The intersection of red and blue lines is our solution.

In mathematical optimization, the method of **Lagrange multipliers** (named after Joseph Louis Lagrange) is a method for finding the maximum/minimum of a function subject to constraints.

For example (see Figure 1 on the right) if we want to solve:

maximize $f(x, y)$
subject to $g(x, y) = c$

We introduce a new variable (λ) called a Lagrange multiplier to rewrite the problem as:

maximize $f(x, y) - \lambda (g(x, y) - c)$

Solving this new equation for x, y, and λ will give us the solution (x, y) for our original equation.

Introduction

Consider a two-dimensional case. Suppose we have a function *f*(*x*,*y*) we wish to maximize or minimize subject to the constraint

$$g(x, y) = c,$$

where *c* is a constant. We can visualize contours of *f* given by

$$f(x, y) = d_n$$

for various values of d_n, and the contour of *g* given by *g*(*x*,*y*) = *c*.

Suppose we walk along the contour line with *g* = *c*. In general the contour lines of *f* and *g* may be distinct, so traversing the contour line for *g* = *c* could intersect with or cross the contour lines of *f*. This is equivalent to saying that while moving along the contour line for *g* = *c* the value of *f* can vary. Only when the contour line for *g* = *c* touches contour lines of *f* tangentially, we do not increase or decrease the value of *f* - that is, when the contour lines touch but do not cross.

This occurs exactly when the tangential component of the total derivative vanishes: $df_{\parallel} = 0$, which is at the constrained stationary points of *f* (which include the constrained local extrema, assuming *f* is differentiable). Computationally, this is when the gradient of *f* is normal to the constraint(s): when $\nabla f = \lambda \nabla g$ for some scalar λ (where ∇ is the gradient). Note that the constant λ is required because, even though the directions of both gradient vectors are equal, the magnitudes of the gradient vectors are generally not equal.

A familiar example can be obtained from weather maps, with their contour lines for temperature and pressure: the constrained extrema will occur where the superposed maps show touching lines (isopleths).

Geometrically we translate the tangency condition to saying that the gradients of *f* and *g* are parallel vectors at the maximum, since the gradients are always normal to the contour lines. Thus we want points (*x*,*y*) where *g*(*x*,*y*) = *c* and

$$\nabla_{x,y} f = \lambda \nabla_{x,y} g,$$

where

$$\nabla_{x,y} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right).$$

To incorporate these conditions into one equation, we introduce an auxiliary function

$$F(x,y,\lambda) = f(x,y) + \lambda(g(x,y) - c),$$

and solve

$$\nabla_{x,y,\lambda} F(x,y,\lambda) = 0.$$

Justification

As discussed above, we are looking for stationary points of *f* seen while travelling on the level set $g(x,y) = c$. This occurs just when the gradient of *f* has no component tangential to the level sets of *g*. This condition is equivalent to $\nabla_{x,y} f(x,y) = \lambda \nabla_{x,y} g(x,y)$ for some λ. Stationary points (x,y,λ) of *F* also satisfy $g(x,y) = c$ as can be seen by considering the derivative with respect to λ. In other words, taking the derivative of the auxillary function with respect to λ and setting it equal to zero is the same thing as taking the constraint equation into account.

Caveat: extrema versus stationary points

Be aware that the solutions are the *stationary points* of the Lagrangian *F*, and are saddle points: they are not necessarily *extrema* of *F*. *F* is unbounded: given a point (x,y) that doesn't lie on the constraint, letting $\lambda \to \pm\infty$ makes *F* arbitrarily large or small. However, under certain stronger assumptions, as we shall see below, the **strong Lagrangian principle** holds, which states that the maxima of *f* maximize the Lagrangian globally.

A more general formulation: The weak Lagrangian principle

Denote the objective function by $f(\mathbf{x})$ and let the constraints be given by $h_k(\mathbf{x}) = c_k$, perhaps by moving constants to the left, as in $h_k(\mathbf{x}) - c_k = g_k(\mathbf{x})$. The domain of *f* should be an open set containing all points satisfying the constraints. Furthermore, *f* and the g_k must have continuous first partial derivatives and the gradients of the g_k must not be zero on the domain.[1] Now, define the Lagrangian, Λ, as

$$\Lambda(\mathbf{x},\boldsymbol{\lambda}) = f + \sum_k \lambda_k g_k.$$

k is an index for variables and functions associated with a particular constraint, *k*.
$\boldsymbol{\lambda}$ without a subscript indicates the vector with elements λ_k, which are taken to be independent variables.

Observe that both the optimization criteria and constraints $g_k(x)$ are compactly encoded as stationary points of the Lagrangian:

$\nabla_{\mathbf{x}}\Lambda = \mathbf{0}$ if and only if $$\nabla_{\mathbf{x}} f = -\sum_k \lambda_k \nabla_{\mathbf{x}} g_k,$$
$\nabla_{\mathbf{x}}$ means to take the gradient only with respect to each element in the vector $\mathbf{x}$, instead of all variables.

and

$\nabla_{\lambda}\Lambda = \mathbf{0}$ implies $g_k = 0$.

Collectively, the stationary points of the Lagrangian,

$$\nabla\Lambda = \mathbf{0},$$

give a number of unique equations totaling the length of $\mathbf{x}$ plus the length of λ.

Interpretation of λ_i

Often the Lagrange multipliers have an interpretation as some salient quantity of interest. To see why this might be the case, observe that:

$$\frac{\partial\Lambda}{\partial g_k} = \lambda_k.$$

So, λ_k is the rate of change of the quantity being optimized as a function of the constraint variable. As examples, in Lagrangian mechanics the equations of motion are derived by finding stationary points of the action, the time integral of the difference between kinetic and potential energy. Thus, the force on a particle due to a scalar potential, $F = -\nabla V$, can be interpreted as a Lagrange multiplier determining the change in action (transfer of potential to kinetic energy) following a variation in the particle's constrained trajectory. In economics, the optimal profit to a player is calculated subject to a constrained space of actions, where a Lagrange multiplier is the value of relaxing a given constraint (e.g. through bribery or other means).

The method of Lagrange multipliers is generalized by the Karush-Kuhn-Tucker conditions.

Examples

Very simple example

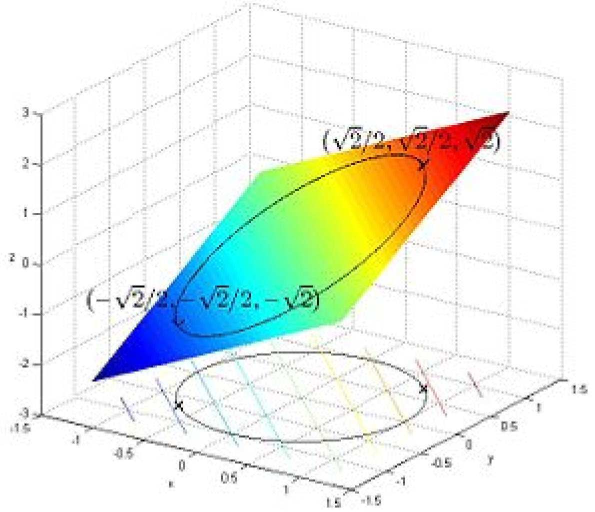


Fig. 3. Illustration of the constrained optimization problem.

Suppose you wish to maximize $f(x,y) = x + y$ subject to the constraint $x^2 + y^2 = 1$. The constraint is the unit circle, and the level sets of f are diagonal lines (with slope -1), so one can see graphically that the maximum occurs at $(\sqrt{2}/2, \sqrt{2}/2)$ (and the minimum occurs at $(-\sqrt{2}/2, -\sqrt{2}/2)$

Formally, set $g(x,y) = x^2 + y^2 - 1$, and

$$\Lambda(x,y,\lambda) = f(x,y) + \lambda g(x,y) = x + y + \lambda(x^2 + y^2 - 1)$$

Set the derivative $d\Lambda = 0$, which yields the system of equations:

$$\begin{aligned} \frac{\partial \Lambda}{\partial x} &= 1 + 2\lambda x &&= 0, \qquad \text{(i)} \\ \frac{\partial \Lambda}{\partial y} &= 1 + 2\lambda y &&= 0, \qquad \text{(ii)} \\ \frac{\partial \Lambda}{\partial \lambda} &= x^2 + y^2 - 1 &&= 0, \qquad \text{(iii)} \end{aligned}$$

As always, the $\partial \lambda$ equation is the original constraint.

Combining the first two equations yields $x = y$ (explicitly, $x \neq 0$,otherwise (i) yields 1 = 0), so one can solve for λ, yielding $\lambda = -1/(2x)$, which one can substitute into (ii)).

Substituting into (iii) yields $2x^2 = 1$, so $x = \pm\sqrt{2}/2$ and the stationary points are $(\sqrt{2}/2, \sqrt{2}/2)$ and $(-\sqrt{2}/2, -\sqrt{2}/2)$. Evaluating the objective function f on these yields

$$f(\sqrt{2}/2, \sqrt{2}/2) = \sqrt{2} \text{ and } f(-\sqrt{2}/2, -\sqrt{2}/2) = -\sqrt{2},$$

thus the maximum is $\sqrt{2}$, which is attained at $(\sqrt{2}/2, \sqrt{2}/2)$ and the minimum is $-\sqrt{2}$, which is attained at $(-\sqrt{2}/2, -\sqrt{2}/2)$.

Simple example

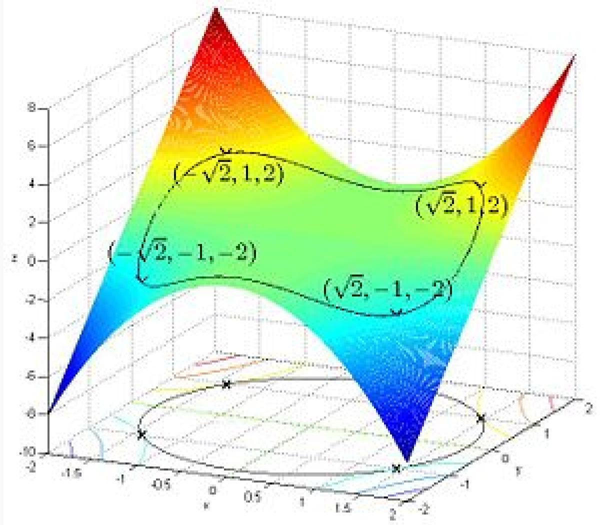


Fig. 4. Illustration of the constrained optimization problem.

Suppose you want to find the maximum values for

$$f(x, y) = x^2 y$$

with the condition that the *x* and *y* coordinates lie on the circle around the origin with radius √3, that is,

$$x^2 + y^2 = 3.$$

As there is just a single condition, we will use only one multiplier, say λ.

Use the constraint to define a function $g(x, y)$:

$$g(x, y) = x^2 + y^2 - 3.$$

The function g is identically zero on the circle of radius √3. So any multiple of $g(x, y)$ may be added to $f(x, y)$ leaving $f(x, y)$ unchanged in the region of interest (above the circle where our original constraint is satisfied). Let

$$\Lambda(x, y, \lambda) = f(x, y) + \lambda g(x, y) = x^2 y + \lambda(x^2 + y^2 - 3).$$

The critical values of Λ occur when its gradient is zero. The partial derivatives are

$$\frac{\partial \Lambda}{\partial x} = 2xy + 2\lambda x \quad = 0, \qquad \text{(i)}$$
$$\frac{\partial \Lambda}{\partial y} = x^2 + 2\lambda y \quad = 0, \qquad \text{(ii)}$$
$$\frac{\partial \Lambda}{\partial \lambda} = x^2 + y^2 - 3 \quad = 0. \qquad \text{(iii)}$$

Equation (iii) is just the original constraint. Equation (i) implies $x = 0$ *or* $\lambda = -y$. In the first case, if $x = 0$ then we must have $y = \pm\sqrt{3}$ by (iii) and then by (ii) $\lambda=0$. In the second case, if $\lambda = -y$ and substituting into equation (ii) we have that,

$$x^2 - 2y^2 = 0.$$

Then $x^2 = 2y^2$. Substituting into equation (iii) and solving for y gives this value of y:

$$y = \pm 1.$$

Thus there are six critical points:

$$(\sqrt{2}, 1); \quad (-\sqrt{2}, 1); \quad (\sqrt{2}, -1); \quad (-\sqrt{2}, -1); \quad (0, \sqrt{3}); \quad (0, -\sqrt{3}).$$

Evaluating the objective at these points, we find

$$f(\pm\sqrt{2}, 1) = 2; \quad f(\pm\sqrt{2}, -1) = -2; \quad f(0, \pm\sqrt{3}) = 0.$$

Therefore, the objective function attains a global maximum (with respect to the constraints) at $(\pm\sqrt{2}, 1)$ and a global minimum at $(\pm\sqrt{2}, -1)$. The point $(0, \sqrt{3})$ is a local minimum and $(0, -\sqrt{3})$ is a local maximum.

Example: entropy

Suppose we wish to find the discrete probability distribution with maximal information entropy. Then

$$f(p_1, p_2, \ldots, p_n) = -\sum_{k=1}^{n} p_k \log_2 p_k.$$

Of course, the sum of these probabilities equals 1, so our constraint is $g(\mathbf{p}) = 1$ with

$$g(p_1, p_2, \ldots, p_n) = \sum_{k=1}^{n} p_k.$$

We can use Lagrange multipliers to find the point of maximum entropy (depending on the probabilities). For all *k* from 1 to *n*, we require that

$$\frac{\partial}{\partial p_k}(f + \lambda(g - 1)) = 0,$$

which gives

$$\frac{\partial}{\partial p_k}\left(-\sum_{k=1}^{n} p_k \log_2 p_k + \lambda(\sum_{k=1}^{n} p_k - 1)\right) = 0.$$

Carrying out the differentiation of these *n* equations, we get

$$-\left(\frac{1}{\ln 2} + \log_2 p_k\right) + \lambda = 0.$$

This shows that all p_i are equal (because they depend on λ only). By using the constraint $\sum_k p_k = 1$, we find

$$p_k = \frac{1}{n}.$$

Hence, the uniform distribution is the distribution with the greatest entropy.

Economics

Constrained optimization plays a central role in economics. For example, the choice problem for a consumer is represented as one of maximizing a utility function subject to a budget constraint. The Lagrange multiplier has an economic interpretation as the shadow price associated with the constraint, in this case the marginal utility of income.

The strong Lagrangian principle: Lagrange duality

Given a convex optimization problem in standard form

$$\begin{aligned} \text{minimize } & f_0(x) \\ \text{subject to } & f_i(x) \leq 0,\ i \in \{1, \ldots, m\} \\ & h_i(x) = 0,\ i \in \{1, \ldots, p\} \end{aligned}$$

with the domain having non-empty interior, the **Lagrangian function** $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is defined as

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x).$$

The vectors λ and ν are called the *dual variables* or *Lagrange multiplier vectors* associated with the problem. The **Lagrange dual function** $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is defined as

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right).$$

The dual function *g* is concave, even when the initial problem is not convex. The dual function yields lower bounds on the optimal value p^* of the initial problem; for any $\lambda \geq 0$ and any ν we have $g(\lambda, \nu) \leq p^*$. If a constraint qualification such as Slater's condition holds and the original problem is convex, then we have strong duality, i.e. $d^* = \max_{\lambda \geq 0, \nu} g(\lambda, \nu) = \inf f_0 = p^*$.

See also

- Karush-Kuhn-Tucker conditions: generalization of the method of Lagrange multipliers.
- Lagrange multipliers on Banach spaces: another generalization of the method of Lagrange multipliers.

References

1. ^ Gluss, David and Weisstein, Eric W., *Lagrange Multiplier* at MathWorld.

External links

Exposition

- Conceptual introduction(plus a brief discussion of Lagrange multipliers in the calculus of variations as used in physics)
- Lagrange Multipliers without Permanent Scarring(tutorial by Dan Klein)

For additional text and interactive applets

- Simple explanation with an example of governments using taxes as Lagrange multipliers

- Applet
- Tutorial and applet
- Good Video Lecture of Lagrange Multipliers

Karush–Kuhn–Tucker conditions

In mathematics, the **Karush–Kuhn–Tucker** conditions (also known as the Kuhn-Tucker or the **KKT** conditions) are necessary for a solution in nonlinear programming to be optimal, provided some regularity conditions are satisfied. It is a generalization of the method of Lagrange multipliers to inequality constraints.

Let us consider the following nonlinear optimization problem:

$$\begin{array}{l} \min\limits_{x} \quad f(x) \\ \text{subject to:} \\ g_i(x) \le 0, h_j(x) = 0 \end{array}$$

where $f(x)$ is the function to be minimized, $g_i(x)\ (i = 1, \ldots, m)$ are the inequality constraints and $h_j(x)\ (j = 1, \ldots, l)$ are the equality constraints, and m and l are the number of inequality and equality constraints, respectively.

The necessary conditions for this general equality-inequality constrained problem were first published in the Masters thesis of William Karush[1], although they only became renowned after a seminal conference paper by Harold W. Kuhn and Albert W. Tucker.[2]

Necessary conditions

Suppose that the objective function, i.e., the function to be minimized, is $f : \mathbb{R}^n \to \mathbb{R}$ and the constraint functions are $g_i : \mathbb{R}^n \to \mathbb{R}$ and $h_j : \mathbb{R}^n \to \mathbb{R}$. Further, suppose they are continuously differentiable at a point x^*. If x^* is a local minimum that satisfies some regularity conditions, then there exist constants $\mu_i\ (i = 1, \ldots, m)$ and $\nu_j\ (j = 1, \ldots, l)$ such that [3]

Stationarity

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{l} \nu_j \nabla h_j(x^*) = 0,$$

Primal feasibility

$$g_i(x^*) \le 0, \text{ for all } i = 1, \ldots, m$$
$$h_j(x^*) = 0, \text{ for all } j = 1, \ldots, l$$

Dual feasibility

$$\mu_i \ge 0\ (i = 1, \ldots, m)$$

Complementary slackness

$$\mu_i g_i(x^*) = 0 \text{ for all } i = 1, \ldots, m.$$

Regularity conditions (or constraint qualifications)

In order for a minimum point x^* be KKT, it should satisfy some regularity condition, the most used ones are listed below:

- Linear independence constraint qualification (LICQ): the gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at x^*.
- Mangasarian-Fromowitz constraint qualification (MFCQ): the gradients of the active inequality constraints and the gradients of the equality constraints are positive-linearly independent at x^*.
- Constant rank constraint qualification (CRCQ): for each subset of the gradients of the active inequality constraints and the gradients of the equality constraints the rank at a vicinity of x^* is constant.
- Constant positive linear dependence constraint qualification (CPLD): for each subset of the gradients of the active inequality constraints and the gradients of the equality constraints, if it is positive-linear dependent at x^* then it is positive-linear dependent at a vicinity of x^*. ($v_1, \ldots, v_n$ is positive-linear dependent if there exists $a_1 \geq 0, \ldots, a_n \geq 0$ not all zero such that $a_1 v_1 + \ldots + a_n v_n = 0$)
- Quasi-normality constraint qualification (QNCQ): if the gradientes of the active inequality constraints and the gradientes of the equalify constraints are positive-lineary independent at x^* with associated multipliers λ_i for equalities and μ_j for inequalities than it doesn't exist a sequence $x_k \to x^*$ such that: $\lambda i \neq 0 \Rightarrow \lambda i$ hi(xk)>0 and $\mu j \neq 0 \Rightarrow \mu j$ gj(x_k)>0.
- Slater condition: for a convex problem, there exists a point x such that $h(x) = 0$ and $g_i(x) < 0$ for all i active in x^*.
- Linearity constraints: If f and g are affine functions, then no other condition is needed to assure that the minimum point is KKT.

It can be shown that LICQ⇒MFCQ⇒CPLD⇒QNCQ, LICQ⇒CRCQ⇒CPLD⇒QNCQ (and the converses are not true), although MFCQ is not equivalent to CRCQ. In practice weaker constraint qualifications are preferred since they provide stronger optimality conditions.

Sufficient conditions

The most common sufficient conditions are stated as follows. Let the objective function $f : \mathbb{R}^n \to \mathbb{R}$ be **convex**, the constraint functions $g_i : \mathbb{R}^n \to \mathbb{R}$ be **convex** functions and $h_j : \mathbb{R}^n \to \mathbb{R}$ be**affine** functions, and let x^* be a point in $\mathbb{R}^n$. If there exist constants μ_i $(i = 1, \ldots, m)$ and ν_j $(j = 1, \ldots, l)$ such that

Stationarity

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) + \sum_{j=1}^{l} \nu_j \nabla h_j(x^*) = 0$$

Primal feasibility

$$g_i(x^*) \le 0 \text{ for all } i = 1, \ldots, m$$
$$h_j(x^*) = 0 \text{ for all } j = 1, \ldots, l$$

Dual feasibility

$$\mu_i \ge 0 \ (i = 1, \ldots, m)$$

Complementary slackness

$$\mu_i g_i(x^*) = 0 \text{ for all } i = 1, \ldots, m,$$

then the point x^* is a global minimum.

It was shown by Martin in 1985 that the broader class of functions in which KKT conditions guarantees global optimality are the so called **invex** functions. So if equality constraints are affine functions, inequality constraints and the objective function are invex functions then KKT conditions are sufficient for global optimality.

Value function

If we reconsider the optimization problem as a maximization problem with constant inequality constraints,

$$\max_x \ f(x)$$
$$\text{subject to:}$$
$$g_i(x) \le a_i, h_j(x) = 0.$$

The value function is defined as:

$$V(a_1, \ldots a_n) = \sup_x \ f(x)$$
$$\text{subject to:}$$
$$g_i(x) \le a_i, h_j(x) = 0$$
$$j \in \{1, \ldots l\}, i \in 1, \ldots, m.$$

(So the domain of *V* is $\{a \in \mathbb{R}^m | \text{for some } x \in X g_i(x) \le a_i, i \in \{1, \ldots, m\}.$)

Given this definition, each coefficient, λ_i, is the rate at which the value function increases as a_i increases. Thus if each a_i is interpreted as a resource constraint, the coefficients tell you how much increasing a resource will increase the optimum value of our function *f*. This interpretation is especially important in economics and is used, for instance, in utility maximization problems.

References

1. ^ W. Karush (1939). "*Minima of Functions of Several Variables with Inequalities as Side Constraints*". M.Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago, Chicago, Illinois.. Available from http://wwwlib.umi.com/dxweb/details?doc_no=7371591 (for a fee)
2. ^ Kuhn, H. W.; Tucker, A. W. (1951). "Nonlinear programming". *Proceedings of 2nd Berkeley Symposium*: 481-492, Berkeley: University of California Press.
3. ^ "The Karush-Kuhn-Tucker Theorem". Retrieved on 2008-03-11.

Further reading

- J. Nocedal, S. J. Wright, *Numerical Optimization.* Springer Publishing. ISBN 978-0-387-30303-1.
- Avriel, Mordecai (2003). *Nonlinear Programming: Analysis and Methods.* Dover Publishing. ISBN 0-486-43227-0.
- R. Andreani, J. M. Martínez, M. L. Schuverdt, On the relation between constant positive linear dependence condition and quasinormality constraint qualification. *Journal of Optimization Theory and Applications, vol. 125, no2*, pp. 473-485 (2005).
- Jalaluddin Abdullah, *Optimization by the Fixed-Point Method, Version 1.97*. [1].
- Martin, D.H, The essence of invexity, *J. Optim. Theory Appl. 47*, (1985) 65-76.

Envelope theorem

The **envelope theorem** is a basic theorem used to solve maximization problems in microeconomics. It may be used to prove Hotelling's lemma, Shephard's lemma, and Roy's identity. The statement of the theorem is:

Consider an arbitrary maximization problem where the objective function (f) depends on some parameter (a):

$$M(a) = \max_x f(x, a)$$

where the function $M(a)$ gives the maximized value of the objective function (f) as a function of the parameter (a). Now let $x(a)$ be the (arg max) value of x that solves the maximization problem in terms of the parameter (a), i.e. so that $M(a) = f(x(a),a)$. The envelope theorem tells us how $M(a)$ changes as the parameter (a) changes, namely:

$$\frac{dM(a)}{da} = \left.\frac{\partial f(x^*, a)}{\partial a}\right|_{x^* = x(a)}.$$

That is, the derivative of M with respect to a is given by the partial derivative of $f(x,a)$ with respect to a, holding x fixed, and then evaluating at the optimal choice (x^*). The vertical bar to the right of the partial derivative denotes that we are to make this evaluation at $x^* = x(a)$.

Envelope theorem in generalized calculus

In the calculus of variations, the envelope theorem relates evolutes to single paths. This was first proved by Jean Gaston Darboux and Ernst Zermelo (1894) and Adolf Kneser (1898). The theorem can be stated as follows:

"*When a single-parameter family of external paths from a fixed point* O *has an envelope, the integral from the fixed point to any point* A *on the envelope equals the integral from the fixed point to any second point* B *on the envelope plus the integral along the envelope to the first point on the envelope,* $J_{OA} = J_{OB} + J_{BA}$." [1]

See also

- Optimization problem
- Random optimization
- Simplex algorithm

- Topkis's Theorem
- Variational calculus

References

1. ^ Kimball, W. S., *Calculus of Variations by Parallel Displacement*. London: Butterworth, p. 292, 1952.

Hotelling's lemma

Hotelling's lemma is a result in microeconomics that relates the supply of a good to the profit of the good's producer. It was first shown by Harold Hotelling, and is widely used in the theory of the firm. The lemma is very simple, and can be stated:

Let y(p) be a firm's net supply function in terms of a certain good's price (p). Then:

$$y(p) = \frac{\partial \pi(p)}{\partial p}$$

for π the profit function of the firm in terms of the good's price, assuming that p > 0 and that derivative exists.

The proof of the theorem stems from the fact that for a profit-maximizing firm, the maximum of the firm's profit at some output $y^*(p)$ is given by the minimum of $\pi(p^*) - p^* y^*(p)$ at some price, p^*, namely where $\partial\pi(p) - y^* = 0$ holds. Thus, $y(p) = \frac{\partial \pi(p)}{\partial p}$, and we are done.

The proof is also a simple corollary of the envelope theorem.

See also

- Hotelling's law
- Hotelling's rule
- Supply and demand
- Shephard's lemma

References

- Hotelling, H. (1932). Edgeworth's taxation paradox and the nature of demand and supply functions. Journal of Political Economy, 40, 577-616.

Shephard's lemma

Shephard's lemma is a major result in microeconomics having applications in consumer choice and the theory of the firm. The lemma states that if indifference curves of the expenditure or cost function are convex, then the cost minimizing point of a given good (*i*) with price p_i is unique. The idea is that a consumer will buy a unique ideal amount of each item to minimize the price for obtaining a certain level of utility given the price of goods in the market. It was named after Ronald Shephard who gave a proof using the distance formula in a paper published in 1953, although it was already used by John Hicks (1939) and Paul Samuelson (1947).

Definition

The lemma give a precise formulation for the demand of each good in the market with respect to that level of utility and those prices: the derivative of the expenditure function ($e(p,u)$) with respect to that price:

$$h_i(u,p) = \frac{\partial e(p,u)}{\partial p_i}$$

where $h_i(u,p)$ is the Hicksian demand for good i, $e(p,u)$ is the expenditure function, and both functions are in terms of prices (a vector p) and utility u.

Although Shephard's original proof used the distance formula, modern proofs of the Shephard's lemma use the envelope theorem.

Application

Shephard's lemma gives a relationship between expenditure (or cost) functions and Hicksian demand. The lemma can be re-expressed as Roy's identity, which gives a relationship between anindirect utility function and a corresponding Marshallian demand function.

Roy's identity

Roy's identity (named for French economist Rene Roy) is a major result in microeconomics having applications in consumer choice and the theory of the firm. The lemma relates the ordinary demand function to the derivatives of the indirect utility function.

Derivation of Roy's identity

Roy's identity reformulates Shephard's lemma in order to get a Marshallian demand function for an individual and a good (*i*) from some indirect utility function.

The first step is to consider the trivial identity obtained by substituting the expenditure function for wealth or income (*m*)in the indirect utility function ($\psi\ (m,p)$, at a utility of *u*):

$$\psi\ (e(p,u),p) = u$$

This says that the indirect utility function evaluated in such a way that minimizes the cost for achieving a certain utility given a set of prices (a vector *p*) is equal to that utility when evaluated at those prices.

Taking the derivative of both sides of this equation with respect to the price of a single good p_i (with the utility level held constant) gives:

$$\frac{\partial\psi\ [e(u,p),p]}{\partial m}\frac{\partial e(u,p)}{\partial p_i}+\frac{\partial\psi\ [e(u,p),p]}{\partial p_i}=0.$$

Rearranging gives the desired result:

$$\frac{\partial e(u,p)}{\partial p_i}=-\frac{\frac{\partial\psi\ [e(u,p),p]}{\partial p_i}}{\frac{\partial\psi\ [e(u,p),p]}{\partial m}}=x_i(m,p)$$

Application

This gives a method of deriving the Marshallian demand function of a good for some consumer from the indirect utility function of that consumer. It is also fundamental in deriving the Slutsky equation.

References

- Roy, René (1947). "La Distribution du Revenu Entre Les Divers Biens," Econometrica, 15, 205-225.

Fixed Point Theorems

Contents

Fixed-point theorems are one of the major tools economists use for proving existence, etc. One of the oldest fixed-point theorems - Brouwer's - was developed in 1910 and already by 1928, John von Neumann was using it to prove the existence of a "minimax" solution to two-agent games (which translates itself mathematically into the existence of a saddlepoint). von Neumann (1937) used a generalization of Brouwer's theorem to prove existence again for a saddlepoint - this time for a balanced growth equilibrium for his expanding economy. This generalization was later simplified by Kakutani (1941). Working on the theory of games, John Nash (1950) was among the first to use Kakutani's Fixed Point Theorem. Gerard Debreu (1952), generalizing upon Nash, came across this. The existence proofs of Arrow and Debreu (1954), McKenzie (1954) and others gave Kakutani's Fixed Point Theorem a central role. Brouwer's Theorem made a reapparence in Lionel McKenzie (1959), Hirofumi Uzawa (1962) and, later, in the computational work of Herbert Scarf (1973).

(A) Intermediate Value Theorem (IVT)

Theorem: (**Bolzano, IVT**) Let $f : [a, b] \to R$ be a continuous function, where [a, b] is a non-empty, compact, convex subset of R and $f(a) \cdot f(b) < 0$, then there exists a $x^* \in [a, b]$ such that $f(x^*) = 0$.

Proof: (i) Suppose $f(a) > 0$, then this implies $f(b) < 0$. Define $M^+ = \{x \in [a, b] \mid f(x) \geq 0\}$ and $M^- = \{x \in [a, b] \mid f(x) \leq 0\}$. By continuity of f on a connected set $[a, b] \subset R$, then M^+ and M^- are closed and $M^+ \cap M^- \neq \varnothing$. Suppose not. Suppose $M^+ \cap M^- = \varnothing$ so $x \in M^+ \Rightarrow x \notin M^-$ and $x \in M^- \Rightarrow x \notin M^+$. But $M^+ \cup M^- = [a, b]$, which implies that $M^- = (M^+)^c$. But as M^+ is closed, then M^- is open. A contradiction. Thus, $M^+ \cap M^- \neq \varnothing$, i.e. there is an $x^* \in [a, b]$ such that $x^* \in M^+ \cap M^-$, i.e. there is an x^* such that $f(x^*) \leq 0$ and $f(x^*) \geq 0$. Thus, there is an $x^* \in [a, b]$ such that $f(x^*) = 0$.♣

We can follow up on this with the following demonstration:

Corollary: Let $f : [0, 1] \to [0, 1]$ be a continuous function. Then, there exists a fixed point, i.e. $\exists\, x^* \in [0, 1]$ such that $f(x^*) = x^*$.

Proof: there are two essential possibilities: (i) if $f(0) = 0$ or if $f(1) = 1$, then we are done.

(ii) If $f(0) \neq 0$ and $f(1) \neq 1$, then define $F(x) = f(x) - x$. In this case:

$F(0) = f(0) - 0 = f(0) > 0$

$F(1) = f(1) - 1 < 0$

So $F: [0, 1] \to R$, where $F(0) \cdot F(1) < 0$. As $f(.)$ is continuous, then $F(.)$ is also continuous. Then by using the Intermediate Value Theorem (IVT), $\exists\, x^* \in [0, 1]$ such that $F(x^*) = 0$. By the definition of $F(.)$, then $F(x^*) = f(x^*) - x^* = 0$, thus $f(x^*) = x^*$.♣

(B) <u>Brouwer's Fixed Point Theorem</u>

Brouwer's fixed point theorem (Th. 1.10.1 in Debreu, 1959) is a generalization of the corollary to the IVT set out above. Specifically:

Theorem: (**Brouwer**) Let $f: S \to S$ be a continuous function from a non-empty, compact, convex set $S \subset R^n$ into itself, then there is a $x^* \in S$ such that $x^* = f(x^*)$ (i.e. x^* is a fixed point of function f).

Proof: Omitted. See Nikaido (1968: p. 63), Scarf (1973: p. 52) or Border (1985: p.28).

Thus, the previous corollary is simply a special case (where $S = [0, 1]$) of Brouwer's fixed point theorem. The intuition can be gathered from Figure 1 where we have a function f mapping from [0, 1] to [0, 1]. At point d, obviously x' maps to $x'' = f(x')$ and $x' \neq x''$, thus point d is not a fixed point. f intersects the 45° line at three points - a, b and c - all of which represent different fixed points as, for instance, at point a, $x^* = f(x^*)$.

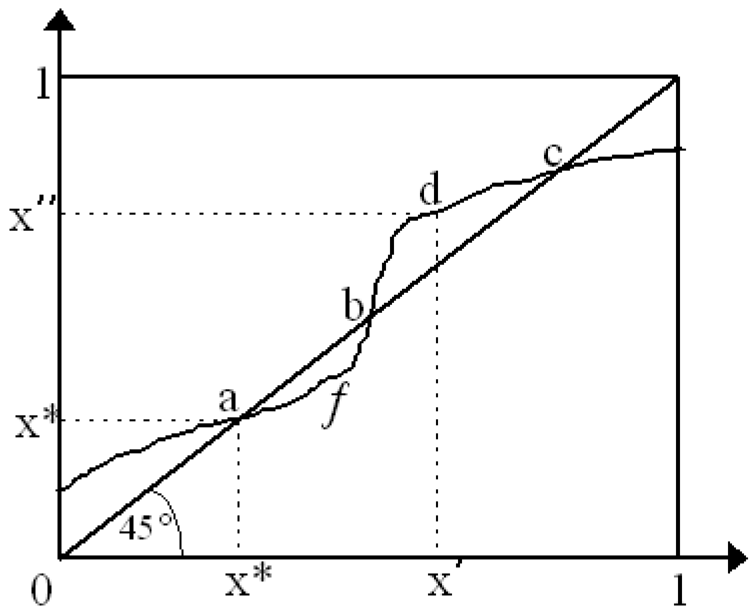


Figure 1 - Brouwer's Fixed Point Theorem

(C) Kakutani's Fixed Point Theorem

The following, Kakutani's fixed-point theorem for correspondences (Th. 1.10.2 in Debreu, 1959), can be derived from Brouwer's Fixed Point Theorem via a continuous selection argument.

Theorem: (**Kakutani**) Let $\varphi : S \rightrightarrows S$ be an upper semi-continuous correspondence from a non-empty, compact, convex set $S \subset R^n$ into itself such that for all $x \in S$, the set $\varphi(x)$ is convex and non-empty, then $\varphi(.)$ has a fixed point, i.e. there is an x^* where $x^* \in \varphi(x^*)$.

Proof: Omitted. See Nikaido (1968: p.67) or Border (1985: p.72).

We can see this in Figure 2 below where S = [0, 1] and the correspondence φ is obviously upper semicontinuous and convex-valued. Obviously, we have a fixed-point at point the intersection of the correspondence with the 45 ° line at point (a) where $x^* \in \varphi(x^*)$.

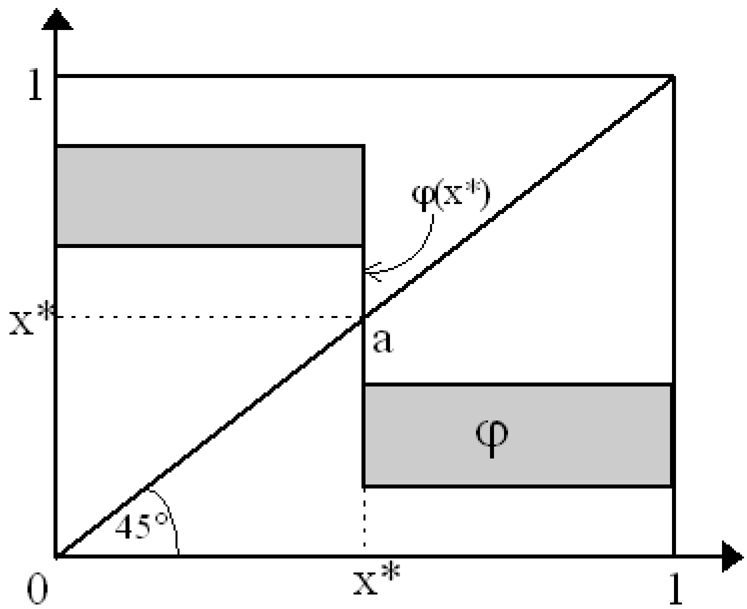


Figure 8 - Kakutani's Fixed Point Theorem

Note the importance of convex-valuedness for this result in Figure 2: if the upper and lower portions of the correspondence φ were *not* connected by a line at φ (x*) (e.g. if φ (x*) was merely the end of the upper "box" and the end of the lower "box" only and thus not a convex set), then the correspondence would *still* be upper semicontinuous (albeit not convex-valued) but it would not intersect the 45° line (thus

$x^* \notin \varphi(x^*)$) and thus there would be no fixed point, i.e. no x^* such that $x^* \in \varphi(x^*)$. Thus, convex-valuedness is instrumental.

Selected References

Kim C. Border (1985) *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge, UK: Cambridge University Press.

Gerard Debreu (1959) *Theory of Value: An axiomatic analysis of economic equilibrium*. New York: Wiley.

J. von Neumann (1937) "A Model of General Economic Equilibrium", in K. Menger, editor, *Ergebnisse eines mathematischen Kolloquiums, 1935-36*. 1945 Translation, *Review of Economic Studies*, Vol. 13 (1), p.1-9.

Hukukane Nikaido (1968) *Convex Structures and Economic Theory*. New York: Academic Press.

Herbert Scarf (1973) *The Computation of Economic Equilibria*. (with the collaboration of Terje Hansen), New Haven: Yale University Press.

Complementarity theory

A **complementarity problem** is a problem where one of the constraints is that the inner product of two elements of vector space must equal zero. i.e. <X,Y>=0[1]. In particular for finite-dimensional vector spaces this means that with vectors X and Y (with nonnegative components) then for each pair of components x_i and y_i one of the pair must be zero, hence the name*complementarity*. e.g. X=(1,0) and Y=(0,2) are complementary, but X=(1,1) and Y=(2,0) are not. A **complementarity problem** is a special case of a variational inequality.

History

Complementarity problems were originally studied because the Karush–Kuhn–Tucker conditions in linear programming and quadratic programming constitute a Linear complementarity problem(LCP) or a Mixed complementarity problem (MCP). In 1963 Lemke and Howson showed that, for two person games, computing a Nash equilibrium point is equivalent to an LCP. In 1968 Cottleand Danzig unified linear and quadratic programming and bimatrix games. Since then the study of **complementarity problems** and **variational inequalities** has expanded enormously.

Areas of mathematics and science that contributed to the development of **complementarity theory** include: optimization, equilibrium problems, variational inequality theory, fixed point theory,topological degree theory and nonlinear analysis.

See also

- Mathematical programming with equilibrium constraints

References

1. ^ Billups, Stephen; Murty, Katta (1999), *Complementarity Problems* http://www-personal.umich.edu/~murty/LCPart.ps

Further reading

- Richard Cottle, F. Giannessi, Jacques Louis Lions (1980). *Variational Inequalities and Complementarity Problems: Theory and Applications*. John Wiley & Sons. ISBN 978-0471276104.
- George Isac (1992). *Complementarity Problems*. Springer. ISBN 978-3540562511.

- (1997) in Michael C. Ferms, Jong-Shi Pang: *Complementarity and Variational Problems: State of the Art*. SIAM. ISBN 978-0898713916.
- George Isac (2000). *Topological Methods in Complementarity Theory*. Springer. ISBN 978-0792362746.
- Francisco Facchinei, Jong-Shi Pang (2003). *Finite-Dimensional Variational Inequalities and Complementarity Problems: v.1 and v.2*. Springer. ISBN 978-0387955803.

External links

- CPNET:Complementarity Problem Net

Mixed complementarity problem

Mixed Complementarity Problem (**MCP**) is a problem formulation in mathematical programming. Many well-known problem types are special cases of, or may be reduced to MCP. It is a generalization of Nonlinear Complementarity Problem (NCP).

Definition

The mixed complementarity problem is defined by a mapping $F(x) : \mathbb{R}^n \to \mathbb{R}^n$, lower values $\ell_i \in \mathbb{R} \cup \{-\infty\}$ and upper values $u_i \in \mathbb{R} \cup \{\infty\}$.

The **solution** of the MCP is a vector $x \in \mathbb{R}^n$ such that for each index $i \in \{1, \ldots, n\}$ one of the following alternatives holds:

- $x_i = \ell_i, \; F_i(x) \ge 0$;
- $\ell_i < x_i < u_i, \; F_i(x) = 0$;
- $x_i = u_i, \; F_i(x) \le 0$.

Another definition for MCP is: it is a variational inequality on the parallelepiped $[\ell, u]$.

See also

- Complementarity theory

References

- Stephen C. Billups (1995). "*Algorithms for complementarity problems and generalized equations*" (PS). Retrieved on 2006-08-14.

Variational inequality

Variational inequality is a mathematical theory intended for the study of equilibrium problems. Guido Stampacchia put forth the theory in 1964 to study partial differential equations. The applicability of the theory has since been expanded to include problems from economics, finance, optimization and game theory.

The problem is commonly restricted to $\mathbf{R}^n$. Given a subset **K** of $\mathbf{R}^n$ and a mapping $F : \mathbf{K} \to \mathbf{R}^n$, the finite-dimensional variational inequality problem associated with **K** is

$$\text{finding } x \in \mathbf{K} \text{ so that } \langle F(x), y - x \rangle \geq 0 \text{ for all } y \in \mathbf{K}$$

where $<\cdot,\cdot>$ is the standard inner product on $\mathbf{R}^n$.

In general, the variational inequality problem can be formulated on any finite- or infinite-dimensional Banach space. Given a Banach space **E**, a subset **K** of **E**, and a mapping $F : \mathbf{K} \to \mathbf{E}^*$, the variational inequality problem is the same as above where $<\cdot,\cdot> : \mathbf{E}^* \times \mathbf{E} \to \mathbf{R}$ is the duality pairing.

Examples

Consider the problem of finding the minimal value of a continuous differentiable function f over a closed interval $I = [a,b]$. Let x be a point in I where the minimum occurs. Three cases can occur:

1) if $a < x < b$ then $f'(x)=0$;
2) if $x = a$ then $f'(x) \geq 0$;
3) if $x = b$ then $f'(x) \leq 0$.

These necessary conditions can be summarized as the problem of

$$\text{finding } x \in I \text{ so that } f'(x)(y - x) \geq 0 \text{ for all } y \in I.$$

References

- Kinderlehrer, David; Stampacchia, Guido (1980), *An Introduction to Variational Inequalities and Their Applications*, New York: Academic Press, ISBN 0-89871-466-4
- G. Stampacchia. *Formes Bilineaires Coercitives sur les Ensembles Convexes*, Comptes Rendus de l'Academie des Sciences, Paris, 258, (1964), 4413–4416.

See also

- projected dynamical system
- differential variational inequality
- Complementarity theory
- Mathematical programming with equilibrium constraints

Linear complementarity problem

From Wikipedia, the free encyclopedia

In mathematical optimization theory, the **linear complementarity problem**, or **LCP**, is a special case of quadratic programming which arises frequently in computational mechanics. Given a real matrix **M** and vector **b**, the linear complementarity problem seeks a vector **x** which satisfies the following two constraints:

- $\mathbf{Mx} + \mathbf{b} \geq \mathbf{0}$ and $\mathbf{x} \geq \mathbf{0}$; that is, each component of these two vectors is non-negative, and
- $\mathbf{x}^{\mathrm{T}}(\mathbf{Mx} + \mathbf{b}) = 0$, the **complementarity condition**.

A sufficient condition for existence and uniqueness of a solution to this problem is that **M** be symmetric positive-definite.

Relation to Quadratic Programming

Finding a solution to the linear complementarity problem is equivalent to minimizing the quadratic function

$$f(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}(\mathbf{Mx} + \mathbf{b})$$

subject to the constraints

$$\mathbf{Mx} + \mathbf{b} \geq \mathbf{0} \text{ and } \mathbf{x} \geq \mathbf{0}.$$

Indeed, these constraints ensure that *f* is always non-negative, so that it attains a minimum of 0 at **x** if and only if **x** solves the linear complementarity problem.

If **M** is positive definite, any algorithm for solving (convex) QPs can of course be used to solve an LCP. However, there also exist more efficient, specialized algorithms, such as Lemke's algorithmand Dantzig's algorithm.

See also

- Complementarity theory

Further reading

- Cottle, Richard W. et al. (1992) *The linear complementarity problem*. Boston, Mass. : Academic Press

- R. W. Cottle and G. B. Dantzig. Complementary pivot theory of mathematical programming. *Linear Algebra and its Applications*, 1:103-125, 1968.

Nash equilibrium

In game theory, **Nash equilibrium** (named after John Forbes Nash, who proposed it) is a solution concept of a game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing only his or her own strategy (i.e., by changing unilaterally). If each player has chosen a strategy and no player can benefit by changing his or her strategy while the other players keep theirs unchanged, then the current set of strategy choices and the corresponding payoffs constitute a Nash equilibrium. In other words, to be a Nash equilibrium, each player must answer negatively to the question: "Knowing the strategies of the other players, and treating the strategies of the other players as set in stone, can I benefit by changing my strategy?"

Stated simply, Amy and Bill are in Nash equilibrium if Amy is making the best decision she can, taking into account Bill's decision, and Bill is making the best decision he can, taking into account Amy's decision. Likewise, many players are in Nash equilibrium if each one is making the best decision that they can, taking into account the decisions of the others. However, Nash equilibrium does not necessarily mean the best cumulative payoff for all the players involved; in many cases all the players might improve their payoffs if they could somehow agree on strategies different from the Nash equilibrium (e.g. competing businessmen forming a cartel in order to increase their profits).

History

The concept of the Nash equilibrium (NE) in pure strategies was first developed by Antoine Augustin Cournot in his theory of oligopoly (1838). Firms choose a quantity of output to maximize their own profit. However, the best output for one firm depends on the outputs of others. A Cournot equilibrium occurs when each firm's output maximizes its profits given the output of the other firms, which is a pure-strategy NE. However, the modern game-theoretic concept of NE is defined in terms of mixed-strategies, where players choose a probability distribution over possible actions. The concept of the mixed strategy NE was introduced by John von Neumann and Oskar Morgenstern in their 1944 book *The Theory of Games and Economic Behavior*. However, their analysis was restricted to the very special case of zero-sum games. They showed that a mixed-strategy NE will exist for any zero-sum game with a finite set of actions. The contribution of John Forbes Nash in his 1951 article *Non-Cooperative Games* was to define a mixed strategy NE for any game with a finite set of actions and prove that at least one (mixed strategy) NE must exist.

Definitions

Informal definition

Informally, a set of strategies is a Nash equilibrium if no player can do better by unilaterally changing his or her strategy. As a heuristic, one can imagine that each player is told the strategies of the other players. If any player would want to do something different after being informed about the others' strategies, then that set of strategies is not a Nash equilibrium. If, however, the player does not want to switch (or is indifferent between switching and not) then the set of strategies is a Nash equilibrium.

The Nash equilibrium may sometimes appear non-rational in a third-person perspective. This is because it may happen that a Nash equilibrium is not pareto optimal.

The Nash-equilibrium may also have non-rational consequences in sequential games because players may "threat"-en each other with non-rational moves. For such games the Subgame perfect Nash equilibrium may be more meaningful as a tool of analysis.

Formal definition

Let *(S, f)* be a game, where S_i is the strategy set for player *i*, $S=S_1 X S_2 \ldots X S_n$ is the set of strategy profiles and $f=(f_1(x), \ldots, f_n(x))$ is the payoff function. Let x_{-i} be a strategy profile of all players except for player *i*. When each player $i \in \{1, \ldots, n\}$ chooses strategy x_i resulting in strategy profile $x = (x_1, \ldots, x_n)$ then player *i* obtains payoff $f_i(x)$. Note that the payoff depends on the strategy profile chosen, i.e. on the strategy chosen by player *i* as well as the strategies chosen by all the other players. A strategy profile $x^* \in S$ is a Nash equilibrium (NE) if no unilateral deviation in strategy by any single player is profitable for that player, that is

$$\forall i, x_i \in S_i, x_i \neq x_i^* : f_i(x_i^*, x_{-i}^*) \geq f_i(x_i, x_{-i}^*).$$

A game can have a pure strategy NE or an NE in its mixed extension (that of choosing a pure strategy stochastically with a fixed frequency). Nash proved that, if we allow *mixed strategies* (players choose strategies randomly according to pre-assigned probabilities), then every n-player game in which every player can choose from finitely many strategies admits at least one Nash equilibrium.

When the inequality above holds strictly (with > instead of $\geq$) for all players and all feasible alternative strategies, then the equilibrium is classified as a **strict Nash equilibrium**. If instead, for some player, there is exact equality between x_i^* and some other strategy in the set *S*, then the equilibrium is classified as a **weak Nash equilibrium**.

Examples

Competition game

	Player 2 chooses '0'	Player 2 chooses '1'	Player 2 chooses '2'	Player 2 chooses '3'
Player 1 chooses '0'	*0, 0*	*2, -2*	*2, -2*	*2, -2*
Player 1 chooses '1'	*-2, 2*	*1, 1*	*3, -1*	*3, -1*
Player 1 chooses '2'	*-2, 2*	*-1, 3*	*2, 2*	*4, 0*
Player 1 chooses '3'	*-2, 2*	*-1, 3*	*0, 4*	*3, 3*

A competition game

This can be illustrated by a two-player game in which both players simultaneously choose a whole number from 0 to 3 and they both win the smaller of the two numbers in points. In addition, if one player chooses a larger number than the other, then he/she has to give up two points to the other. This game has a unique pure-strategy Nash equilibrium: both players choosing 0 (highlighted in light red). Any other choice of strategies can be improved if one of the players lowers his number to one less than the other player's number. In the table to the left, for example, when starting at the green square it is in player 1's interest to move to the purple square by choosing a smaller number, and it is in player 2's interest to move to the blue square by choosing a smaller number. If the game is modified so that the two players win the named amount if they both choose the same number, and otherwise win nothing, then there are 4 Nash equilibria (0,0...1,1...2,2...and 3,3).

Coordination game

	Player 2 adopts strategy 1	Player 2 adopts strategy 2
Player 1 adopts strategy 1	*A*, *A*	*B*, *C*
Player 1 adopts strategy 2	*C*, *B*	*D*, *D*

A coordination game

The *coordination game* is a classic (symmetric) two player, two strategy game, with the payoff matrix shown to the right, where the payoffs satisfy $A>C$ and $D>B$. The players should thus coordinate, either on *A* or on *D*, to receive a high payoff. If the players' choices do not coincide, a lower payoff is rewarded. An example of a coordination game is the setting where two technologies are available to two firms with compatible products, and they have to elect a strategy to become the market standard. If both firms agree on the chosen technology, high sales are expected for both firms. If the firms do not agree on the standard technology, few sales result. Both strategies are Nash equilibria of the game.

Driving on a road, and having to choose either to drive on the left or to drive on the right of the road, is also a coordination game. For example, with payoffs 100 meaning no crash and 0 meaning a crash, the coordination game can be defined with the following payoff matrix:

	Drive on the Left	Drive on the Right
Drive on the Left	100, 100	0, 0
Drive on the Right	0, 0	100, 100

The driving game

In this case there are two pure strategy Nash equilibria, when both choose to either drive on the left or on the right. If we admit mixed strategies (where a pure strategy is chosen at random, subject to some fixed probability), then there are three Nash equilibria for the same case: two we have seen from the pure-strategy form, where the probabilities are (0%,100%) for player one, (0%, 100%) for player two; and (100%, 0%) for player one, (100%, 0%) for player two respectively. We add another where the probabilities for each player is (50%, 50%).

Prisoner's dilemma

(note differences in the orientation of the payoff matrix)

The Prisoner's Dilemma has the same payoff matrix as depicted for the Coordination Game, but now C > A > D > B. Because C > A and D > B, each player improves his situation by switching from strategy #1 to strategy #2, no matter what the other player decides. The Prisoner's Dilemma thus has a single Nash Equilibrium: both players choosing strategy #2 ("betraying"). What has long made this an interesting case to study is the fact that D < A (ie., "both betray" is globally inferior to "both remain loyal"). The globally optimal strategy is unstable; it is not an equilibrium.

Nash equilibria in a payoff matrix

There is an easy numerical way to identify Nash Equilibria on a Payoff Matrix. It is especially helpful in two-person games where players have more than two strategies. In this case formal analysis may become too long. This rule does not apply to the case where mixed (stochastic) strategies are of interest. The rule goes as follows: if the first payoff number, in the duplet of the cell, is the maximum of the column of the cell and if the second number is the maximum of the row of the cell - then the cell represents a Nash equilibrium.

We can apply this rule to a 3x3 matrix:

	Option A	Option B	Option C
Option A	0, 0	**25, 40**	5, 10
Option B	**40, 25**	0, 0	5, 15

Option C	10, 5	15, 5	**10, 10**

A Payoff Matrix

Using the rule, we can very quickly (much faster than with formal analysis) see that the Nash Equlibria cells are (B,A), (A,B), and (C,C). Indeed, for cell (B,A) 40 is the maximum of the first column and 25 is the maximum of the second row. For (A,B) 25 is the maximum of the second column and 40 is the maximum of the first row. Same for cell (C,C). For other cells, either one or both of the duplet members are not the maximum of the corresponding rows and columns.

This said, the actual mechanics of finding equilibrium cells is obvious: find the maximum of a column and check if the second member of the pair is the maximum of the row. If these conditions are met, the cell represents a Nash Equilibrium. Check all columns this way to find all NE cells. An NxN matrix may have between 0 and NxN pure strategy Nash equilibria.

Stability

The concept of stability, useful in the analysis of many kinds of equilibrium, can also be applied to Nash equilibria.

A Nash equilibrium for a mixed strategy game is stable if a small change (specifically, an infinitesimal change) in probabilities for one player leads to a situation where two conditions hold:

1. the player who did not change has no better strategy in the new circumstance
2. the player who did change is now playing with a strictly worse strategy

If these cases are both met, then a player with the small change in his mixed-strategy will return immediately to the Nash equilibrium. The equilibrium is said to be stable. If condition one does not hold then the equilibrium is unstable. If only condition one holds then there are likely to be an infinite number of optimal strategies for the player who changed. John Nash showed that the latter situation could not arise in a range of well-defined games.

In the "driving game" example above there are both stable and unstable equilibria. The equilibria involving mixed-strategies with 100% probabilities are stable. If either player changes his probabilities slightly, they will be both at a disadvantage, and his opponent will have no reason to change his strategy in turn. The (50%,50%) equilibrium is unstable. If either player changes his probabilities, then the other player immediately has a better strategy at either (0%, 100%) or (100%, 0%).

Stability is crucial in practical applications of Nash equilibria, since the mixed-strategy of each player is not perfectly known, but has to be inferred from statistical distribution of his actions in the game. In this case unstable equilibria are very unlikely to arise in practice, since any minute change in the proportions of each strategy seen will lead to a change in strategy and the breakdown of the equilibrium.

A Coalition-Proof Nash Equilibrium (CPNE) (similar to a Strong Nash Equilibrium) occurs when players cannot do better even if they are allowed to communicate and collaborate before the game. Every correlated strategy supported by iterated strict dominance and on the Pareto frontier is a CPNE[1]. Further, it is possible for a game to have a Nash equilibrium that is resilient against coalitions less than a specified size, k. CPNE is related to the theory of the core.

Occurrence

If a game has a unique Nash equilibrium and is played among players under certain conditions, then the NE strategy set will be adopted. Sufficient conditions to guarantee that the Nash equilibrium is played are:

1. The players all will do their utmost to maximize their expected payoff as described by the game.
2. The players are flawless in execution.
3. The players have sufficient intelligence to deduce the solution.
4. The players know the planned equilibrium strategy of all of the other players.
5. The players believe that a deviation in their own strategy will not cause deviations by any other players.
6. There is common knowledge that all players meet these conditions, including this one. So, not only must each player know the other players meet the conditions, but also they must know that they all know that they meet them, and know that they know that they know that they meet them, and so on.

Where the conditions are not met

Examples of game theory problems in which these conditions are not met:

1. The first condition is not met if the game does not correctly describe the quantities a player wishes to maximize. In this case there is no particular reason for that player to adopt an equilibrium strategy. For instance, the prisoner's dilemma is not a dilemma if either player is happy to be jailed indefinitely.
2. Intentional or accidental imperfection in execution. For example, a computer capable of flawless logical play facing a second flawless

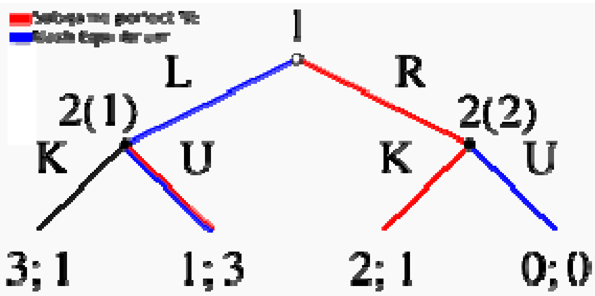
computer will result in equilibrium. Introduction of imperfection will lead to its disruption either through loss to the player who makes the mistake, or through negation of the 4th 'common knowledge' criterion leading to possible victory for the player. (An example would be a player suddenly putting the car into reverse in the game of 'chicken', ensuring a no-loss no-win scenario).

3. In many cases, the third condition is not met because, even though the equilibrium must exist, it is unknown due to the complexity of the game, for instance in Chinese chess[2]. Or, if known, it may not be known to all players, as when playing tic-tac-toe with a small child who desperately wants to win (meeting the other criteria).
4. The fourth criterion of common knowledge may not be met even if all players do, in fact, meet all the other criteria. Players wrongly distrusting each other's rationality may adopt counter-strategies to expected irrational play on their opponents' behalf. This is a major consideration in "Chicken" or an arms race, for example.

Where the conditions are met

Due to the limited conditions in which NE can actually be observed, they are rarely treated as a guide to day-to-day behaviour, or observed in practice in human negotiations. However, as a theoretical concept in economics, and evolutionary biology the NE has explanatory power. The payoff in economics is money, and in evolutionary biology gene transmission, both are the fundamental bottom line of survival. Researchers who apply games theory in these fields claim that agents failing to maximize these for whatever reason will be competed out of the market or environment, which are ascribed the ability to test all strategies. This conclusion is drawn from the "stability" theory above. In these situations the assumption that the strategy observed is actually a NE has often been borne out by research.

NE and non-credible threats



	(K,K)	(K,U)	(U,U)	(U,K)
L	3,1	3,1	1,3	1,3
R	2,1	0,0	0,0	2,1

Extensive and Normal form illustrations that show the difference between SPNE and other NE. The blue equilibrium is not subgame perfect because player two makes a non-credible threat at 2(2) to be unkind (U).

The nash equilibrium is a superset of the subgame perfect nash equilibrium. The subgame perfect equilibrium in addition to the Nash Equilibrium requires that the strategy also is a Nash equilibrium in every subgame of that game. This eliminates all non-credible threats, that is, strategies that contain non-rational moves in order to make the counter-player change his strategy.

The image to the right shows a simple sequential game that illustrates the issue with subgame imperfect Nash equilibria. In this game player one chooses left(L) or right(R), which is followed by player two being called upon to be kind (K) or unkind (U) to player one, However, player two only stands to gain from being unkind if player one goes left. If player one goes right the rational player two would de facto be kind to him in that subgame. However, The non-credible threat of being unkind at 2(2) is still part of the blue (L, (U,U)) nash equilibrium. Therefore, if rational behavior can be expected by both parties the subgame perfect Nash equilibrium may be a more meaningful solution concept when such dynamic inconsistencies arise.

Proof of existence

As above, let σ_{-i} be a mixed strategy profile of all players except for player *i*. We can define a best response correspondence for player *i*, b_i. b_i is a relation from the set of all probability distributions over opponent player profiles to a set of player *i*'s strategies, such that each element of

$$b_i(\sigma_{-i})$$

is a best response to σ_{-i}. Define

$$b(\sigma) = b_1(\sigma_{-1}) \times b_2(\sigma_{-2}) \times \cdots \times b_n(\sigma_{-n}).$$

One can use the Kakutani fixed point theorem to prove that *b* has a fixed point. That is, there is a σ^* such that $\sigma^* \in b(\sigma^*)$. Since $b(\sigma^*)$ represents the best response for all players to σ^*, the existence of the fixed point proves that there is some strategy set which is a best response to itself. No player could do any better by deviating, and it is therefore a Nash equilibrium.

When Nash made this point to John von Neumann in 1949, von Neumann famously dismissed it with the words, "That's trivial, you know. That's just a fixed point theorem." (See Nasar, 1998, p. 94.)

Alternate proof using the Brouwer fixed point theorem

We have a game $G = (N,A,u)$ where N is the number of players and $A = A_1 \times \ldots \times A_N$ is the action set for the players. All of the actions sets A_i are finite. Let $\Delta = \Delta_1 \times \ldots \times \Delta_N$ denote the set of mixed strategies for the players. The finiteness of the A_is insures the compactness of Δ.

We can now define the gain functions. For a mixed strategy $\sigma \in \Delta$, we let the gain for player i on action $a \in A_i$ be

$$Gain_i(\sigma,a) = \max\{0, u_i(a_i,\sigma_{-i}) - u_i(\sigma_i,\sigma_{-i})\}$$

The gain function represents the benefit a player gets by unilaterally changing his strategy. We now define $g = (g_1, \ldots, g_N)$ where

$$g_i(\sigma)(a) = \sigma_i(a) + Gain_i(\sigma,a)$$

for $\sigma \in \Delta, a \in A_i$. We see that

$$\sum_{a \in A_i} g_i(\sigma)(a) = \sum_{a \in A_i} \sigma_i(a) + Gain_i(\sigma,a) = 1 + \sum_{a \in A_i} Gain_i(\sigma,a) > 0$$

We now use g to define $f : \Delta \to \Delta$ as follows. Let

$$f_i(\sigma)(a) = \frac{g_i(\sigma)(a)}{\sum_{b \in A_i} g_i(\sigma)(b)}$$

for $a \in A_i$. It is easy to see that each f_i is a valid mixed strategy in Δ_i. It is also easy to check that each f_i is a continuous function of σ, and hence f is a continuous function. Now Δ is the cross product of a finite number of compact convex sets, and so we get that Δ is also compact and convex. Therefore we may apply the Brouwer fixed point theorem to f. So f has a fixed point in Δ, call itσ^*.

I claim that σ^* is a Nash Equilibrium in G. For this purpose, it suffices to show that

$$\forall 1 \le i \le N,\ \forall a \in A_i,\ Gain_i(\sigma^*,a) = 0.$$

This simply states the each player gains no benefit by unilaterally changing his strategy which is exactly the necessary condition for being a Nash Equilibrium.

Now assume that the gains are not all zero. Therefore, $\exists i$, $1 \le i \le N$, and $a \in A_i$ such that $Gain_i(\sigma^*,a) > 0$. Note then that

$$\sum_{a \in A_i} g_i(\sigma^*,a) = 1 + \sum_{a \in A_i} Gain_i(\sigma^*,a) > 1$$

So let $C = \sum_{a \in A_i} g_i(\sigma^*, a)$. Also we shall denote $Gain(i, \cdot)$ as the gain vector indexed by actions in A_i. Since $f(\sigma^*) = \sigma^*$ we clearly have that $f_i(\sigma^*) = \sigma_i^*$. Therefore we see that

$$\sigma_i^* = \frac{g_i(\sigma^*)}{\sum_{a \in A_i} g_i(\sigma^*)(a)} \Rightarrow \sigma_i^* = \frac{\sigma_i^* + Gain_i(\sigma^*, \cdot)}{C} \Rightarrow C\sigma_i^* = \sigma_i^* + Gain_i(\sigma^*, \cdot)$$

$$(C-1)\,\sigma_i^* = Gain_i(\sigma^*, \cdot) \Rightarrow \sigma_i^* = \left(\frac{1}{C-1}\right) Gain_i(\sigma^*, \cdot)$$

Since $C > 1$ we have that σ_i^* is some positive scaling of the vector $Gain_i(\sigma^*, \cdot)$. Now I claim that

$$\sigma_i^*(a)(u_i(a_i, \sigma_{-i}^*) - u_i(\sigma_i^*, \sigma_{-i}^*)) = \sigma_i^*(a) Gain_i(\sigma^*, a)$$

$\forall a \in A_i$. To see this, we first note that if $Gain_i(\sigma^*, a) > 0$ then this is true by definition of the gain function. Now assume that $Gain_i(\sigma^*, a) = 0$. By our previous statements we have that

$$\sigma_i^*(a) = \left(\frac{1}{C-1}\right) Gain_i(\sigma^*, a) = 0$$

and so the left term is zero, giving us that the entire expression is 0 as needed.

So we finally have that

$$\begin{aligned}
0 &= u_i(\sigma_i^*, \sigma_{-i}^*) - u_i(\sigma_i^*, \sigma_{-i}^*) \\
&= \left(\sum_{a \in A_i} \sigma_i^*(a) u_i(a_i, \sigma_{-i}^*)\right) - u_i(\sigma_i^*, \sigma_{-i}^*) \\
&= \sum_{a \in A_i} \sigma_i^*(a)(u_i(a_i, \sigma_{-i}^*) - u_i(\sigma_i^*, \sigma_{-i}^*)) \\
&= \sum_{a \in A_i} \sigma_i^*(a) Gain_i(\sigma^*, a) \qquad \text{by the previous statements} \\
&= \sum_{a \in A_i} (C-1)\,\sigma_i^*(a)^2 > 0
\end{aligned}$$

where the last inequality follows since σ_i^* is a non-zero vector. But this is a clear contradiction, so all the gains must indeed be zero. Therefore σ^* is a Nash Equilibrium for *G* as needed.

See also

- Adjusted Winner procedure

- Best response
- Conflict resolution research
- Evolutionarily stable strategy
- Game theory
- Glossary of game theory
- Hotelling's law
- Mexican Standoff
- Minimax theorem
- Optimum contract and par contract
- Prisoner's dilemma
- Relations between equilibrium concepts
- Solution concept
- Equilibrium selection
- Stackelberg competition
- Subgame perfect Nash equilibrium
- Wardrop's principle
- Complementarity theory

References

Game Theory textbooks

- Dutta, Prajit K. (1999), *Strategies and games: theory and practice*, MIT Press, ISBN 978-0-262-04169-0. Suitable for undergraduate and business students.

- Fudenberg, Drew and Jean Tirole (1991) *Game Theory* MIT Press.

- Morgenstern, Oskar and John von Neumann (1947) *The Theory of Games and Economic Behavior* Princeton University Press

- Myerson, Roger B. (1997), *Game theory: analysis of conflict*, Harvard University Press, ISBN 978-0-674-34116-6

- Rubinstein, Ariel; Osborne, Martin J. (1994), *A course in game theory*, MIT Press, ISBN 978-0-262-65040-3. A modern introduction at the graduate level.

Original Papers

- Nash, John (1950) "Equilibrium points in n-person games" *Proceedings of the National Academy of Sciences* 36(1):48-49.
- Nash, John (1951) "Non-Cooperative Games" *The Annals of Mathematics* 54(2):286-295.

Other References

- Mehlmann, A. *The Game's Afoot! Game Theory in Myth and Paradox,* American Mathematical Society (2000).

- Nasar, Sylvia (1998), "A Beautiful Mind", Simon and Schuster, Inc.

Notes

1. ^ D. Moreno, J. Wooders (1996). "Coalition-Proof Equilibrium". *Games and Economic Behavior* **17**: 80–112. doi:10.1006/game.1996.0095.
2. ^ Nash proved that a perfect NE exists for this type of finite extensive form game – it can be represented as a strategy complying with his original conditions for a game with a NE. Such games may not have unique NE, but at least one of the many equilibrium strategies would be played by hypothetical players having perfect knowledge of all 10^{150} game trees.

External links

- Complete Proof of Existence of Nash Equilibria