Mathematics for Microeconomics, Part II

Ming-Ching Luoh

2020.9.25.

Second-Order Conditions and Curvature

Functions of one variable

- Consider the case of

$$y = f(x)$$

A necessary condition for a maximum is

$$\frac{dy}{dx} = f'(x) = 0$$

- For a maximum, y must be decreasing for movements away from it. Change in y is

$$dy = f'(x)dx$$

To be at a maximum, dy must be decreasing for small increases in x.

Change of dy is the second derivative of y

$$d(dy) = d^2y = \frac{d[f'(x)dx]}{dx} = f''(x)dx \cdot dx = f''(x)dx^2$$

- $d^2y < 0$ implies $f''(x)dx^2 < 0$. Since dx^2 (square of dx) must be positive, $f''(x) < 0$
- This means that the function f must have a concave shape at the critical point. This is the curvature condition for a maximum,

Functions of two variables

- Next, consider y as a function of two independent variables, $y = f(x_1, x_2)$.
- First order conditions for a maximum are

$$\frac{\partial y}{\partial x_1} = f_1 = 0$$
$$\frac{\partial y}{\partial x_2} = f_2 = 0$$

- For a local maximum, f_1 and f_2 must be diminishing at the critical point.
- Conditions must also be placed on the cross-partial derivative ($f_{12} = f_{21}$) to ensure that dy is decreasing for movements through the critical point in any direction.

- The total differential of y is

$$dy = f_1 dx_1 + f_2 dx_2$$

and the change in dy is

$$\begin{aligned} d^2y &= (f_{11}dx_1 + f_{12}dx_2)\,dx_1 + (f_{21}dx_1 + f_{22}dx_2)\,dx_2 \\ &= f_{11}dx_1^2 + f_{12}dx_2dx_1 + f_{21}dx_1dx_2 + f_{22}dx_2^2 \end{aligned}$$

- By Young's theorem, $f_{12} = f_{21}$, then

$$d^2y = f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2$$

- For d^2y to be unambiguously negative for any change in the x's, it is necessary that $f_{11} < 0$, $f_{22} < 0$.

- For example, if $dx_2 = 0$, then $d^2y = f_{11}dx_1^2$ and $d^2y < 0$ implies $f_{11} < 0$.
- An identical argument can be made for $dx_1 = 0$, and $f_{22} < 0$.
- If neither dx_1 nor dx_2 is zero, we must consider the cross-partial, f_{12}, in deciding whether d^2y is unambiguously negative.

$$
\begin{aligned}
d^2y &= f_{11}dx_1^2 + 2f_{12}dx_2dx_1 + f_{22}dx_2^2 \\
&= f_{11}dx_1^2 + 2f_{12}dx_2dx_1 + \frac{(f_{12}dx_2)^2}{f_{11}} - \frac{(f_{12}dx_2)^2}{f_{11}} + f_{22}dx_2^2 \\
&= \frac{1}{f_{11}}(f_{11}dx_1 + f_{12}dx_2)^2 + \frac{1}{f_{11}}(f_{11}f_{22} - f_{12}^2)dx_2^2
\end{aligned}
$$

d_y^2 to be unambiguously negative only if $f_{11}f_{22} - f_{12}^2 > 0$ since $f_{11} < 0$.

- See Extensions to this chapter for the general case.

Concave functions

- $f_{11} f_{22} - f_{12}^2 > 0$ requires that the own second partial derivatives (f_{11} and f_{22}) be sufficiently negative so that their product will outweigh any possible perverse effects from the cross-partial derivatives ($f_{12} = f_{21}$).
- Functions that obey such a condition is called **concave functions**.
- Concave functions have the property that they always lie below any plane that is tangent to them.
- The plane defined by the maximum value of the function is simply a special case of this property.

Example 2.10 Second-Order Conditions:Health Status for the Last Time

Health status function from Example 2.6, where y is the health status (0 to 10), x_1, x_2 are daily dosages of two health-enhancing drugs.

$$y = f(x_1, x_2) = -x_1^2 + 2x_1 - x_2^2 + 4x_2 + 5$$

First-order conditions are

$$\begin{aligned} f_1 &= -2x_1 + 2 = 0 \\ f_2 &= -2x_2 + 4 = 0 \\ \text{or} \quad x_1^* &= 1, x_2^* = 2 \end{aligned}$$

Second-order partial derivatives

$$f_{11} = -2, f_{22} = -2, f_{12} = 0, \text{ and } f_{11}f_{22} - f_{12}^2 > 0$$

Both necessary and sufficient conditions for are satisfied.

Constrained maximization

- As another example, consider the problem of choosing x_1 and x_2 to maximize

$$y = f(x_1, x_2)$$

subject to the linear constraint

$$c - b_1 x_1 - b_2 x_2 = 0$$

where c, b_1, b_2 are constant parameters.

- Lagrangian expression and first-order conditions are

$$\mathcal{L} = f(x_1, x_2) + \lambda(c - b_1 x_1 - b_2 x_2)$$

and

$$\begin{aligned} f_1 - \lambda b_1 &= 0 \\ f_2 - \lambda b_2 &= 0 \\ c - b_1 x_1 - b_2 x_2 &= 0 \end{aligned}$$

- Use the "second" total differential to ensure a local maximum.

$$d^2y = f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2$$

  Only those values of x_1 and x_2 that satisfy the constraint can be considered valid alternatives to the critical point.

- Total differential of the constraint $c - b_1x_1 - b_2x_2 = 0$ is

$$\begin{aligned} -b_1dx_1 - b_2dx_2 &= 0, \\ dx_2 &= -\frac{b_1}{b_2}dx_1 \end{aligned}$$

  This shows the allowable relative changes in x_1 and x_2.

- The first-order conditions imply

$$\frac{f_1}{f_2} = \frac{b_1}{b_2},$$

therefore

$$dx_2 = -\frac{f_1}{f_2}dx_1$$

and thus

$$\begin{aligned} d^2y &= f_{11}dx_1^2 + 2f_{12}dx_1dx_2 + f_{22}dx_2^2 \\ &= f_{11}dx_1^2 - 2f_{12}\frac{f_1}{f_2}dx_1^2 + f_{22}\frac{f_1^2}{f_2^2}dx_1^2 \\ &= \left(f_{11}f_2^2 - 2f_{12}f_1f_2 + f_{22}f_1^2\right)\frac{dx_1^2}{f_2^2} \end{aligned}$$

- Therefore, for $d_y^2 < 0$, it must be true that

$$f_{11}f_2^2 - 2f_{12}f_1f_2 + f_{22}f_1^2 < 0$$

- This equation characterizes a set of functions termed *quasi-concave functions.*

Quasi-concave functions

- Quasi-concave functions have the property that the set of all points for which such a function takes on a value greater than any specific constant is a convex set.
- A set of points is said to be *convex* if any two points in the set can be joined by a straight line that is contained completely within the set.
- Problems 2.9 and 2.10 examine two specific quasi-concave functions that we will frequently encounter in this book.

Example 2.11 Concave and Quasi-Concave Functions

- The differences between concave and quasi-concave functions can be illustrated with the function

$$y = f(x_1, x_2) = (x_1 \cdot x_2)^k$$

where $x_1 > 0, x_2 > 0$, and $k > 0$.

- No matter what value k takes, this function is quasi-concave. To show this, look at the "level curves" of the function at a specific value c.

$$y = c = (x_1 x_2)^k, \text{ or } x_1 x_2 = c^{1/k} = c'$$

- This is the equation of a standard rectangular hrperbola. Clearly the set of points for which y takes on values larger than c is convex because it is bounded by this hyperbola.

- If every point on the line segment joining any two points lies on the set, then it is called a convex set.
- To show the quasi-concavity directly

$$
\begin{aligned}
f_1 &= kx_1^{k-1}x_2^k \\
f_2 &= kx_1^k x_2^{k-1} \\
f_{11} &= k(k-1)x_1^{k-2}x_2^k \\
f_{22} &= k(k-1)x_1^k x_2^{k-2} \\
f_{12} &= k^2 x_1^{k-1}x_2^{k-1}
\end{aligned}
$$

$$
\begin{aligned}
& f_{11}f_2^2 - 2f_{12}f_1f_2 + f_{22}f_1^2 \\
= \; & k^3(k-1)x_1^{3k-2}x_2^{3k-2} - 2k^4x_1^{3k-2}x_2^{3k-2} + k^3(k-1)x_1^{3k-2}x_2^{3k-2} \\
= \; & (-2)k^3x_1^{3k-2}x_2^{3k-2} < 0
\end{aligned}
$$

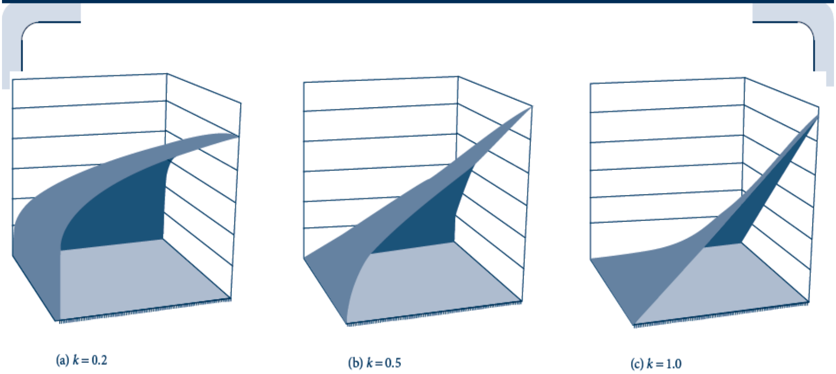
No matter what value k takes, this function is quasi-concave.

- For concavity,

$$\begin{aligned} f_{11}f_{22} - f_{12}^2 &= k^2(k-1)^2 x_1^{2k-2} x_2^{2k-2} - k^4 x_1^{2k-2} x_2^{2k-2} \\ &= x_1^{2k-2} x_2^{2k-2}[k^2(k-1)^2 - k^4] \\ &= x_1^{2k-2} x_2^{2k-2}[k^2(-2k+1)] \end{aligned}$$

- Whether or not the function is concave depends on the value of k.
- If $k < 0.5$, the function is concave since $f_{11}f_{22} - f_{12}^2 > 0$.
- If $k > 0.5$, the function is convex since $f_{11}f_{22} - f_{12}^2 < 0$.
- Intuitively, for points where $x_1 = x_2$, $y = (x_1^2)^k = x_1^{2k}$.

FIGURE 2.4 Concave and Quasi-Concave Functions



In all three cases these functions are quasi-concave. For a fixed y, their level curves are convex. But only for k = 0.2 is the function strictly concave. The case k = 1.0 clearly shows nonconcavity because the function is not below its tangent plane.

Homogeneous Functions

- A function $f(x_1, x_2, \cdots, x_n)$ is said to be homogeneous of degree k if

$$f(tx_1, tx_2, \cdots, tx_n) = t^k f(x_1, x_2, \cdots, x_n).$$

- When $k = 1$, a doubling of all of its arguments doubles the value of the function itself.
- When $k = 0$, a doubling of all of its arguments leaves the value of the function unchanged.

- If a function is homogeneous of degree k, the partial derivatives of the function will be homogeneous of degree $k-1$.

From definition,

$$\begin{aligned}
f(tx_1, tx_2, \cdots, tx_n) &= t^k f(x_1, x_2, \cdots, x_n) \\
\frac{\partial f(tx_1, \cdots, tx_n)}{\partial x_1} &= t^k \frac{\partial f(x_1, \cdots, x_n)}{\partial x_1} \\
\text{and } \frac{\partial f(tx_1, \cdots, tx_n)}{\partial x_1} &= \frac{\partial f(tx_1, \cdots, tx_n)}{\partial tx_1} \cdot \frac{\partial tx_1}{\partial x_1} \\
&= f_1(tx_1, \cdots, tx_n) \cdot t
\end{aligned}$$

Therefore,

$$f_1(tx_1, \cdots, tx_n) = t^{k-1} f_1(x_1, \cdots, x_n)$$

Euler's theorem

- Differentiate the definition for homogeneity with respect to the proportionality factor t yields

$$kt^{k-1}f(x_1,\cdots,x_n) \quad = \quad x_1 f_1(tx_1,\cdots,tx_n)+\cdots+x_n f_n(tx_1,\cdots,tx_n)$$

- For $t=1$:

$$kf(x_1,\cdots,x_n)=x_1 f_1(x_1,\cdots,x_n)+\cdots+x_n f_n(x_1,\cdots,x_n)$$

  This is termed *Euler's theorem*.

- For a homogeneous function, there is a definite relationship between the value of the function and the values of its partial derivatives.

Homothetic functions

- A homothetic function is one that is formed by taking a monotonic transformation of a homogeneous function.
- Monotonic transformations, by definition, preserve the order of the relationship between the arguments of a function and the value of that function.
- They generally do not possess the homogeneity properties of their underlying functions.
- Homothetic functions, however, do preserve the implicit trade-offs among the variables in the function, which depends only on the ratios of those variables, not on their absolute values.

- For example, consider a two-variable function of the form $y = f(x_1, x_2)$ the implicit trade-off between x_1 and x_2 is

$$\frac{dx_2}{dx_1} = -\frac{f_1}{f_2}$$

- If we assume that f is homogeneous of degree k then its partial derivatives will be homogeneous of degree $k - 1$. The implicit trade-off between x_1 and x_2 is

$$\frac{dx_2}{dx_1} = -\frac{t^{k-1} f_1(x_1, x_2)}{t^{k-1} f_2(x_1, x_2)} = -\frac{f_1(tx_1, tx_2)}{f_2(tx_1, tx_2)}$$

Let $t = \frac{1}{x_2}$, then

$$\frac{dx_2}{dx_1} = -\frac{f_1(x_1/x_2, 1)}{f_2(x_1/x_2, 1)}$$

- If we apply any monotonic transformation F(with $F' > 0$) to the original homogeneous function f, the trade-off implied by the new homothetic function $F([f(x_1, x_2)]$ are unchanged

$$\frac{dx_2}{dx_1} = -\frac{F' f_1(x_1/x_2, 1)}{F' f_2(x_1/x_2, 1)} = -\frac{f_1(x_1/x_2, 1)}{f_2(x_1/x_2, 1)}$$

Example 2.12 Cardinal (Numerical) and Ordinal Properties

- Consider various values of the parameter k for the function

$$f(x_1, x_2) = (x_1 x_2)^k$$

- Quasi-concavity is preserved for all values of k.
- It is concave (a cardinal property) only for a narrow range of values of k, many monotonic transformations destroy the concavity of f.
- A proportional increase in the two arguments would yield

$$f(tx_1, tx_2) = t^{2k}(x_1 x_2)^k = t^{2k} f(x_1, x_2)$$

  The degree of homogeneity depends on k.
- The function is homothetic because

$$\frac{dx_2}{dx_1} = -\frac{f_1}{f_2} = -\frac{k x_1^{k-1} x_2^k}{k x_1^k x_2^{k-1}} = -\frac{x_2}{x_1}$$

Integration

Antiderivatives

- Integration is the inverse of differentiation.
- Let $F(x)$ be the integral of $f(x)$, then $f(x)$ is the derivative of $F(x)$.

$$\frac{dF(x)}{dx} = F'(x) = f(x)$$

then

$$F(x) = \int f(x)dx$$

- If $f(x) = x$ then

$$F(x) = \int f(x)dx = \int x dx = \frac{x^2}{2} + C$$

where C is an arbitrary "constant of integration."

Calculation of antiderivatives

Three methods.

1. Creative guesswork. What function will yield $f(x)$ as its derivative? Then use differentiation to check your answer.
   - $F(x) = \int x^2 dx = \frac{x^3}{3} + C$
   - $F(x) = \int x^n dx = \frac{x^{n+1}}{n+1} + C$
   - $F(x) = \int (ax^2 + bx + c) dx = \frac{ax^3}{3} + \frac{bx^2}{2} + cx + C$
   - $F(x) = \int e^x dx = e^x + C$
   - $F(x) = \int a^x dx = \frac{a^x}{\ln a} + C$
   - $F(x) = \int \left(\frac{1}{x}\right) dx = \ln(|x|) + C$
   - $F(x) = \int (\ln x) dx = x \ln x - x + C$

2. Change of variable. Redefine variables to make the function easier to integrate.

- Let $y = 1 + x^2$, then $dy = 2x dx$ and

$$\int \frac{2x}{1+x^2} dx = \int \frac{1}{y} dy = \ln(|y|) = \ln(|1+x^2|)$$

3. Integration by parts. $duv = udv + vdu$

- For any two functions u and v

$$\begin{aligned} \int duv &= uv = \int udv + \int vdu \\ \int udv &= uv - \int vdu \end{aligned}$$

- What the integral of xe^x is? Let $u = x$ (thus, $du = dx$) and $dv = e^x dx$ (thus, $v = e^x$)

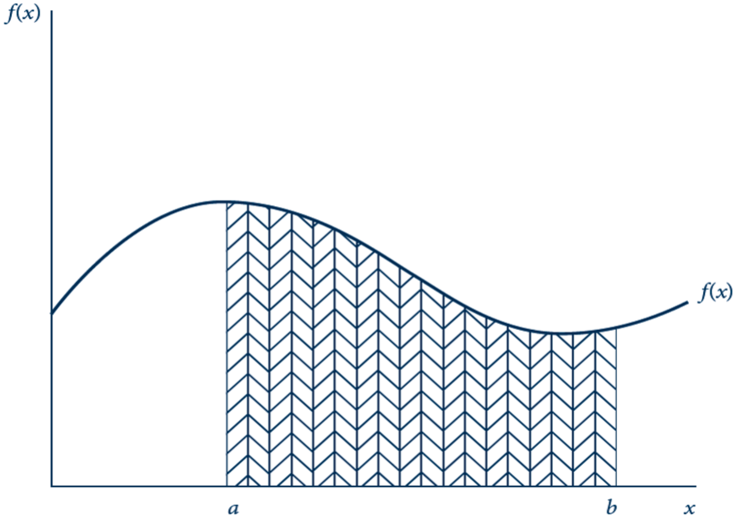
$$\begin{aligned} \int xe^x dx &= \int udv = uv - \int vdu \\ &= xe^x - \int e^x dx = (x-1)e^x + C \end{aligned}$$

Definite integrals

- To sum up the area under a graph of a function over some defined interval.

- Area under $f(x)$ from $x = a$ to $x = b$
- area under $f(x) \approx \sum_i f(x_i)\Delta x_i$
- area under $f(x) \approx \int_a^b f(x_i)dx_i$

Figure 2.5 Definite Integrals Show the Areas Under the Graph of a Function



Fundamental theorem of calculus

- The *fundamental theorem of calculus* directly ties together the two principal tools of calculus: derivatives and integrals.
- It can be used to illustrate the distinction between "stocks" and "flows."

  area under $f(x) = \int_a^b f(x)dx = F(b) - F(a)$

Example 2.13 Stocks and Flows

- Suppose that net population increase for a country can be approximated by the function

$$f(t) = 1,000e^{0.02t}$$

- Net population is growing ("flow" concept) at the rate of 2 percent per year.
- How much in total the population ("stock" concept) will increase over a 50 year period?

$$\int_{t=0}^{t=50} f(t)dt = \int_{t=0}^{t=50} 1,000e^{0.02t}dt = F(t)\Big|_0^{50}$$

$$= \frac{1,000e^{0.02t}}{0.02}\Big|_0^{50} = \frac{1,000e^{0.02\cdot 50}}{0.02} - 50,000 = 85,914$$

Another example.

- Suppose that total costs for a particular firm are given by

$$C(q) = 0.1q^2 + 500$$

- q – output during some period
- Variable costs: $0.1q^2$
- Fixed costs: 500
- Marginal costs $MC = dC(q)/dq = 0.2q$
- Total costs for $q = 100$ is Fixed cost (500) + Variable cost where variable cost is

$$\int_{q=0}^{q=100} 0.2q\, dq = 0.1q^2 \Big|_0^{100} = 1,000 - 0 = 1,000$$

Differentiating a definite integral

1. Differentiation with respect to the variable of integration.
   - A definite integral has a constant value, hence its derivative is zero

$$\frac{d \int_a^b f(x)dx}{dx} = 0$$

2. Differentiation with respect to the upper bound of integration.

- Changing the upper bound of integration will change the value of a definite integral

$$\frac{d\int_a^x f(t)dt}{dx} = \frac{d[F(x) - F(a)]}{dx} = f(x) - 0 = f(x)$$

- If the upper bound of integration is a function of x,

$$\frac{d\int_a^{g(x)} f(t)dt}{dx} = \frac{d[F(g(x)) - F(a)]}{dx}$$
$$= \frac{d[F(g(x))]}{dx} = f\frac{dg(x)}{dx} = f(g(x))g'(x)$$

- • If the lower bound of integration is a function of x,

$$\frac{d\int_{g(x)}^{b} f(t)dt}{dx} = \frac{d[F(b) - F(g(x))]}{dx}$$

$$= -\frac{d[F(g(x))]}{dx} = -f(g(x))g'(x)$$

3. Differentiation with respect to another relevant variable
Suppose we want to integrate $f(x,y)$ with respect to x. How will this be affected by changes in y?

$$\frac{d\int_a^b f(x,y)dx}{dy} = \int_a^b f_y(x,y)dx$$

Dynamic Optimization

Some optimization problems involve multiple periods.

- Need to find the optimal time path for a variable that succeeds in optimizing some goal.
- Decisions made in one period affect outcomes in later periods.

The optimal control problem

- Find the optimal path for $x(t)$ over a specified time interval $[t_0, t_1]$.
- Changes in x are governed by

$$\frac{dx(t)}{dt} = g\left[x(t), c(t), t\right]$$

where $c(t)$ is used to “control" the change in $x(t)$.

- In each period, the decision-maker derive value from x and c according to $f[x(t), c(t), t]$.
- To optimize

$$\int_{t_0}^{t_1} f[x(t), c(t), t]dt$$

- There may also be endpoint constraints:

$$x(t_0) = x_0, \quad x(t_1) = x_1$$

- This problem is “dynamic" since any decision about how much to change x this period will affect not only the future value of x, but it will also affect future values of the outcome function f.

The maximum principle

- At a single point in time, the decision maker must be concerned with both the current value of the objective function $f[x(t), c(t), t]$ and with the implied change in the value of $x(t)$.
- The current value of $x(t)$ is given by $\lambda(t)x(t)$, the instantaneous rate of change of this value is given by

$$\frac{d[\lambda(t)x(t)]}{dt} = \lambda(t)\frac{dx(t)}{dt} + x(t)\frac{d\lambda(t)}{dt}$$

- At any time t, a comprehensive measure of the value of concern to the decision maker is:

$$
\begin{aligned}
H &= f[x(t), c(t), t] + \frac{d[\lambda(t)x(t)]}{dt} \\
&= f[x(t), c(t), t] + \lambda(t)g[x(t), c(t), t] + x(t)\frac{d\lambda(t)}{dt}
\end{aligned}
$$

- The comprehensive value represents both the current benefits being received and the instantaneous change in the value of x.

- What conditions muse hold for $x(t)$ and $c(t)$ to optimize this Hamiltonian expression?

- The two optimality conditions, referred to as the *maximum principle*.

$$
\begin{aligned}
\frac{\partial H}{\partial c} &= f_c + \lambda g_c, \ \text{ or } f_c = -\lambda g_c \\
\frac{\partial H}{\partial x} &= f_x + \lambda g_x + \frac{d\lambda(t)}{dt} = 0, \ \text{ or } f_x + \lambda g_x = -\frac{d\lambda(t)}{dt}
\end{aligned}
$$

- The first condition suggests that, at the margin, the gain from increasing c in terms of the function f must be balanced against future costs.
- The second condition suggests that the net current gain from more x must be weighed against the declining future value of x.

Example 2.14 Allocating a Fixed Supply

- Assume that someone has inherited 1,000 bottles of wine from a rich uncle. He or she intends to drink these bottles over the next 20 years.
- Suppose this person's utility function for wine is given by

$$u[c(t)] = \ln c(t),$$

which exhibits diminishing marginal utility: $u' > 0, u'' < 0$

- This person's goal is to maximize

$$\int_0^{20} u[c(t)]dt = \int_0^{20} \ln c(t)dt$$

- Let $x(t)$ be the number of bottles of wine remaining at time t. This series is constrained by $x(0) = 1,000$ and $x(20) = 0$.
- The differential equation determining the evolution of $x(t)$is

$$\frac{dx(t)}{dt} = -c(t)$$

  That is, each instant's consumption reduces the stock of bottles by the amount consumed.
- The current value Hamiltonian expression is

$$H = \ln c(t) + \lambda[-c(t)] + x(t)\frac{d\lambda}{dt}$$

  and the first-order conditions are

$$\begin{aligned} \frac{\partial H}{\partial c} &= \frac{1}{c} - \lambda = 0 \\ \frac{\partial H}{\partial x} &= \frac{d\lambda}{dt} = 0 \end{aligned}$$

- With λ being constant over time, $c(t)$ is also constant over time. If $c(t) = k$, the number of bottles remaining at any time will be

$$x(t) = 1000 - kt$$

- Since $x(0) = 1000$ and $x(20) = 0$, we have $k = 50$.
- The optimum plan is to drink the wine at the rate of 50 bottles per year for 20 years.

- A more complicated utility function:

$$u[c(t)] = \begin{cases} c(t)^{\gamma}/\gamma, & \text{if } \gamma \neq 0, \gamma < 1, \\ \ln c(t) & \text{if } \gamma = 0 \end{cases}$$

- Assume that the consumer discounts future consumption at the rate δ. Hence this person's goal is to maximize

$$\int_0^{20} u[c(t)]dt = \int_0^{20} e^{-\delta t} \frac{c(t)^{\gamma}}{\gamma} dt$$

subject to the constraints:

$$\frac{dx(t)}{dt} = -c(t), x(0) = 1,000, x(20) = 0$$

.

- The current value Hamiltonian expression is

$$H = e^{-\delta t}\frac{c(t)^{\gamma}}{\gamma} + \lambda(-c) + x(t)\frac{d\lambda(t)}{dt}$$

- The maximum principle requires that

$$\begin{aligned}\frac{\partial H}{\partial c} &= e^{-\delta t}[c(t)]^{\gamma-1} - \lambda = 0\\ \frac{\partial H}{\partial x} &= 0 + 0 + \frac{d\lambda}{dt} = 0\end{aligned}$$

- The value of the wine stock should be constant over time ($\lambda = k$, a constant). and that

$$e^{-\delta t}[c(t)]^{\gamma-1} = k, \text{ or } , c(t) = k^{1/(\gamma-1)}e^{\delta t/(\gamma-1)}$$

- Optimal wine consumption should fall over time since $\gamma - 1 < 0$.

- For example, let $\delta = 0.1$ and $\gamma = -1$, then

$$c(t) = k^{-0.5}e^{-0.05t}$$

- next, we need to choose k to satisfy the endpoint constraints.

$$\begin{aligned}\int_0^{20} c(t)dt &= \int_0^{20} k^{-0.5}e^{-0.05t}dt = -20k^{-0.5}e^{-0.05t}\Big|_0^{20} \\ &= -20k^{-0.5}(e^{-1}-1) = 12.64k^{-0.5} = 1,000\end{aligned}$$

- Finally, the optimal consumption plan is:

$$c(t) \approx 79e^{-0.05t}$$

Mathematical Statistics

- For issued raised by uncertainty and imperfect information, we need a good background in mathematical statistics.

Random variables and probability density functions

- A *random variable* describes the outcomes from an experiment that is subject to chance.

e.g., flipping a coin

$$x = \begin{cases} 1, & \text{if coin is heads} \\ 0 & \text{if coin is tails} \end{cases}$$

Discrete and continuous random variables

- For discrete random variables, the outcomes from a random experiment are a finite number of possibilities.

  e.g.: recording the number that comes up on a single die (random variable with six outcomes)

- For continuous random variable, the outcomes from a random experiment are a continuum of possibilities.

  e.g.: outdoors temperature tomorrow

Probability density function (PDF)

- For any random variable, the *probability density function* (PDF) shows the probability that each outcome will occur.
- The probabilities specified by the PDF must sum to 1.

  Discrete case:

$$\sum_{i=1}^{n} f(x_i) = 1$$

  Continuous case:

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

A few important PDFs

Figure 2.6 Four Common Probability Density Functions

(a) Binomial distribution

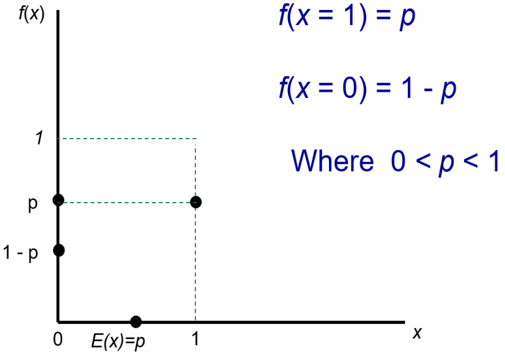


Figure 2.6 Four Common Probability Density Functions
(b) Uniform distribution

$$f(x) = \frac{1}{b-a} \text{ for } a \leq x \leq b$$

$$f(x) = 0 \text{ for } x < a \text{ or } x > b$$

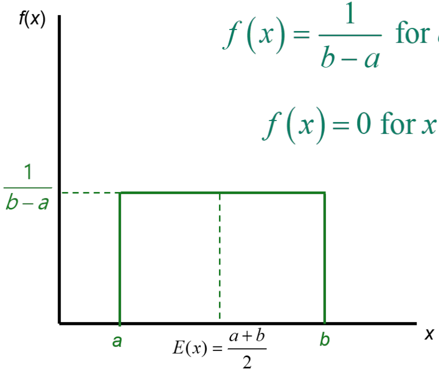


Figure 2.6 Four Common Probability Density Functions
(c) Exponential distribution

$$f(x) = \lambda e^{-\lambda x} \text{ if } x > 0$$
$$f(x) = 0 \text{ if } x \leq 0$$

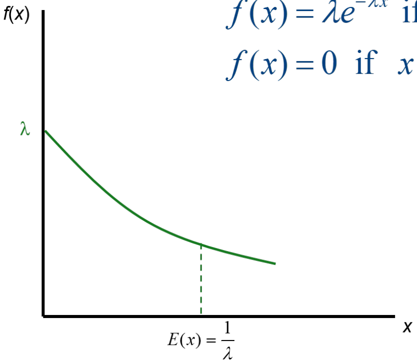
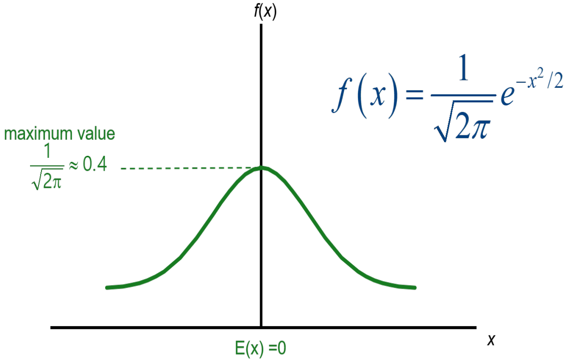


Figure 2.6 Four Common Probability Density Functions (d) Normal distribution



Expected value

- The *expected value* of a random variable is the numerical value that the random variable might be expected to have, on average.
- It is the "center of gravity" of the PDF.
- Discrete case:

$$E(x) = \sum_{i=1}^{n} x_i f(x_i)$$

- Continuous case:

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx$$

- The concept of expected value can be generalized to include the expected value of any function of a random variable [say, $g(x)$].

$$E(g(x)) = \int_{-\infty}^{+\infty} g(x)f(x)dx$$

- As a special case, consider a linear function $y = ax + b$. Then

$$\mathrm{E}(y) = E(ax + b) = \int_{-\infty}^{+\infty} (ax + b)f(x)dx = aE(x) + b$$

- Expected value of a random variable can be phrased in terms of the cumulative distribution function (CDF) $F(x)$,

$$F(x) = \int_{-\infty}^{+\infty} f(t)dt$$

- $F(x)$ represents the probability that the random variable t is less than or equal to x. The expected value of x can be written as Expected value of x:

$$E(x) = \int_{-\infty}^{+\infty} x d{\color{red}F(x)}$$

Example 2.15 Expected Values of a Few Random Variables

1. Binomial:

$$E(x) = 1 \cdot f(x=1) + 0 \cdot f(x=0) = 1 \cdot p + 0 \cdot (1-p) = p$$

2. Uniform:

$$E(x) = \int_a^b \frac{x}{b-a} dx = \frac{b+a}{2}$$

3. Exponential:

$$\begin{aligned} E(x) &= \int_0^\infty x\lambda e^{-\lambda x} dx \\ &= -xe^{-\lambda x}\Big|_0^\infty - \int_0^\infty -e^{-\lambda x} dx \\ &= -\frac{1}{\lambda} e^{-\lambda x}\Big|_0^\infty = \frac{1}{\lambda} \end{aligned}$$

Integration by parts, let $u = x$, $dv = \lambda e^{-\lambda x} dx$, $v = -e^{-\lambda x}$.

4. Normal:

$$\begin{aligned} E(x) &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \left[-e^{-x^2/2} \right] \Big|_{-\infty}^{+\infty} \\ &= \frac{1}{\sqrt{2\pi}} [0 - 0] = 0 \end{aligned}$$

Variance and standard deviation

- The variance of a random variable is A measure of dispersion.
- The variance is defined as the "expected squared deviation' of a random variable from its expected value.

$$\begin{aligned} Var(x) &= \sigma_x^2 = E\left[(x - E(x))^2\right] \\ &= \int_{-\infty}^{+\infty} (x - E(x))^2 f(x) dx \end{aligned}$$

- The square root of the variance is called the *standard deviation* and is denoted as σ_x.

$$\sigma_x = \sqrt{Var(x)} = \sqrt{\sigma_x^2}$$

Example 2.16 Variances and Standard Deviations for Simple Random Variables

1. Binomial:

$$\begin{aligned} \sigma_x^2 &= \sum_{i=1}^{n} (x - E(x))^2 f(x_i) \\ \sigma_x^2 &= (1-p)^2 \cdot p + (0-p)^2 \cdot (1-p) = p \cdot (1-p) \\ \sigma_x &= \sqrt{p \cdot (1-p)} \end{aligned}$$

2. Uniform: $\sigma_x^2 = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \frac{1}{b-a} dx = \frac{(b-a)^2}{12}$

3. Exponential: $\sigma_x^2 = \frac{1}{\lambda^2}$ and $\sigma_x = \frac{1}{\lambda}$

4. Normal: $\sigma_x^2 = \sigma_x = 1$

Standardizing the Normal

- If the random variable x has a standard Normal PDF, it will have an expected value of 0, a standard deviation of 1.
- Linear transformation $y = \sigma x + \mu$ can be used to give this random variable any desired expected value (μ) and standard deviation (σ)

$$\begin{aligned} \mathrm{E}(y) &= \sigma E(x) + \mu \\ \mathrm{Var}(y) &= \sigma_y^2 = \sigma^2 Var(x) = \sigma^2 \end{aligned}$$

Covariance

- The *covariance* between x and y seeks to measure the direction of association between the variables. It is defined as

$$\text{Cov}(x,y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} [x - \text{E}(x)][y - E(y)] f(x,y) dx dy$$

- Two random variables are *independent* if the probability of any particular value of one is not affected by the particular value of the other than may occur'
- This means that the PDF must have the property that $f(x, y) = g(x) \cdot h(y)$.
- If x and y are independent, their covariance will be zero.

$$\text{Cov}(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} [x - \text{E}(x)][y - E(y)] f(x, y) dx dy = 0$$

- However, a zero covariance does not necessarily imply statistical independent.

Extension: Second-Order Conditions and Matrix Algebra

Matrix Algebra background

- An $n \times k$ matrix, A, is a rectangular array of terms with $i = 1, \cdots, n$ and $j = 1, \cdots, k$

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nk} \end{bmatrix}$$

- If $n = k$, then A is a square matrix, A square matrix is symmetric if $a_{ij} = a_{ji}$.
- The *identity matrix*, I_n, is a $n \times n$ square matrix where $a_{ij} = 1$ if $i = j$ and $a_{ij} = 0$ if $i \neq j$
- The **determinant** of a square matrix, (denoted by $|A|$) is a scalar found by suitably multiplying together all the terms in the matrix. If A is 2×2,

$$|A| = a_{11}a_{22} - a_{21}a_{12}.$$

- The *inverse* of an $n \times n$ matrix, A, is another $n \times n$ matrix, A^{-1}, such that $A \times A^{-1} = I_n$

- A necessary and sufficient condition for the existence of A^{-1} is $|A| \neq 0$
- The *leading principal minors* of an $n \times n$ square matrix A are the series of determinants of the first p rows and columns of A, where $p = 1, \cdots, n$.

  If A is 2×2, then the first leading principal minor is a_{11} and the second is $a_{11}a_{22} - a_{21}a_{12}$.
- An $n \times n$ square matrix, A, is positive definite if all its leading principal minors are positive.

  The matrix is negative definite if its principal minors alternate in sign starting with a minus.

- The Hessian matrix is formed by all the second-order partial derivatives of a function.

  If f is a continuous and twice differentiable function of n variables, then its Hessian is given by

$$H(f) = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

E2.1 Concave and Convex Functions

- A concave function is one that is always below (or on) any tangent to it. A convex function is one that is always above (or on) any tangent to it.
- The concavity or convexity of any function is determined by its second derivative(s).
- For a function of a single variable $f(x)$, the Taylor approximation at any point (x_0)

$$\begin{aligned} f(x_0+dx) &= f(x_0)+f'(x_0)dx+f''(x_0)\frac{dx^2}{2} \\ &+ \text{higher} - \text{order terms.} \end{aligned}$$

Assuming that the higher-order terms are 0, we have

$$\begin{aligned} f(x_0+dx) &\leq f(x_0)+f'(x_0)dx \text{ if } f''(x_0)\leq 0 \\ f(x_0+dx) &\geq f(x_0)+f'(x_0)dx \text{ if } f''(x_0)\geq 0 \end{aligned}$$

where $f(x_0)+f'(x_0)dx$ is the equation tangent to the function at x_0.

- For functions with many variables, concavity requires that the Hessian matrix be negative definite, whereas convexity requires that this matrix be positive definite.
- If $f(x_1, x_2)$ is a function of two variables, the Hessian is given by

$$H = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$$

This is negative definite if

$$f_{11} < 0 \text{ and } f_{11}f_{22} - f_{21}f_{12} > 0.$$

Example 1

From Example 2.6: Suppose that y is a function of x_1 and x_2

$$y = -x_1^2 + 2x_1 - x_2^2 + 4x_2 + 5$$

First-order conditions

$$\frac{\partial y}{\partial x_1} = -2x_1 + 2 = 0, \frac{\partial y}{\partial x_2} = -2x_2 + 4 = 0$$

The Hessian is given by

$$H = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

and the first and second leading principal monors are

$$H_1 = -2 < 0$$
$$H_2 = (-2)(-2) - 0 = 4 > 0$$

The Hessian matrix is negative definite, hence the function is concave.

Example 2

The Cobb-Douglas production or utility function $x^a y^b$, where $a, b \in (0,1)$. The first- and second-order derivatives of the function are

$$\begin{aligned} f_x &= ax^{a-1}y^b, \\ f_y &= bx^a y^{b-1}, \\ f_{xx} &= a(a-1)x^{a-2}y^b, \\ f_{yy} &= b(b-1)x^a y^{b-2}. \\ f_{xy} = f_{yx} &= abx^{a-1}y^{b-1} \end{aligned}$$

Hence the Hessian for this function is

$$H = \begin{bmatrix} a(a-1)x^{a-2}y^b & abx^{a-1}y^{b-1} \\ abx^{a-1}y^{b-1} & b(b-1)x^a y^{b-2} \end{bmatrix}$$

The first leading principal minor of this Hessian is

$$H_1 = a(a-1)x^{a-2}y^b < 0,$$

the second leading principal minor is

$$\begin{aligned} H_2 &= a(a-1)b(b-1)x^{2a-2}y^{2b-2} - a^2b^2x^{2a-2}y^{2b-2} \\ &= ab(1-a-b)x^{2a-2}y^{2b-2} \end{aligned}$$

Hence $H_2 > 0$ and thus this function is concave if $a + b < 1$.

E2.2 Maximization

- The first-order conditions for an unconstrained maximum of a function of many variables requires finding a point at which the partial derivatives are zero.
- If the function is concave it will be below its tangent plane at this point; therefore, the point will be a true maximum.

E2.3 Constrained maxima

We wish to maximize

$$f(x_1, \cdots, x_n)$$

subject to the constraint

$$g(x_1, \cdots, x_n) = 0$$

- First-order conditions for a maximum:

$$f_i + \lambda g_i = 0$$

where λ is the Lagrange multiplier.

- Second-order conditions for a maximum:

  Augmented ("bordered") Hessian, H_b

$$H_b = \begin{bmatrix} 0 & g_1 & g_2 & \cdots & g_n \\ g_1 & f_{11} & f_{12} & \cdots & f_{1n} \\ g_2 & f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_n & f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

- For a maximum, $(-1)H_b$ must be negative definite. That is, the leading principal minor of H_b must follow the pattern $- + - + -$ and so forth, starting with the second such minor.

Example: In the optimal fence problem (Example 2.8), the first order conditions are

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial \lambda} &= P - 2x - 2y = 0\\ \frac{\partial \mathcal{L}}{\partial x} &= y - 2\lambda = 0\\ \frac{\partial \mathcal{L}}{\partial y} &= x - 2\lambda = 0,\end{aligned}$$

the bordered Hessian is

$$H = \begin{bmatrix} 0 & -2 & -2 \\ -2 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$$

and $H_{b2} = -4$, $H_{b3} = 8$, this the leading principal minors have the sign pattern required for a maximum.

E2.4 Quasi-concavity

- If the constraint, g, is linear, then the second-order conditions can be related solely to the shape of the function to be optimized.

- The constraint can be written as

$$g(x_1, \cdots, x_n) = c - b_1 x_1 - b_2 x_2 - \cdots - b_n x_n = 0$$

and the first-order conditions for a maximum are

$$f_i = \lambda b_i, i = 1, \cdots, n.$$

- It is clear that The bordered Hessian H_b and the matrix H' have the same leading principal minors except for a (positive) constant of proportionality.

$$H' = \begin{bmatrix} 0 & f_1 & f_2 & \cdots & f_n \\ f_1 & f_{11} & f_{12} & \cdots & f_{1n} \\ f_2 & f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ f_n & f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

- The conditions for a maximum of f subject to a linear constraint will be satisfied provided H' follows the same sign conventions as H_b. That is, $(-1)H'$ must be negative definite.
- A function f for which H' does follow this pattern is called *quasi-concave*.

Example

For the fences problem, $f(x, y) = xy$ and H' is given by

$$H' = \begin{bmatrix} 0 & y & x \\ y & 0 & 1 \\ x & 1 & 0 \end{bmatrix}$$

Thus,

$$\begin{aligned} H'_2 &= -y^2 < 0 \\ H'_3 &= 2xy > 0 \end{aligned}$$

and the function is quasi-concave.

Example

More generally, if f is a function of only two variable,

$$H' = \begin{bmatrix} 0 & f_1 & f_2 \\ f_1 & f_{11} & f_{12} \\ f_2 & f_{21} & f_{22} \end{bmatrix}$$

then quasi-concavity requires that

$$\begin{aligned} H'_2 &= -(f_1)^2 < 0 \text{ and} \\ H'_3 &= 2f_1 f_2 f_{12} - f_{11} f_2^2 - f_{22} f_1^2 > 0 \end{aligned}$$

E2.5 Comparative Statics with two endogenous variables

- Two endogenous variables (x_1 and x_2) and a single exogenous parameter, a.
- It takes two equations (e.g. demand and supply) to determine the equilibrium values of these two endogenous variables, and the values taken by these variables will depend on a. In implicit form as

$$f^1[x_1(a), x_2(a), a] = 0$$
$$f^2[x_1(a), x_2(a), a] = 0$$

- Differentiation of these equilibrium equations with respect to a

$$f_1^1 \frac{dx_1^*}{da} + f_2^1 \frac{dx_2^*}{da} + f_a^1 = 0$$
$$f_1^2 \frac{dx_1^*}{da} + f_2^2 \frac{dx_2^*}{da} + f_a^2 = 0$$

- Solve these simultaneous equations for the comparative static values of the derivatives ($\frac{\partial x_1^*}{\partial a}$ and $\frac{\partial x_2^*}{\partial a}$) that show how the equilibrium values change when a changes.

- We can write simultaneous equations in matrix notation:

$$\begin{bmatrix} f_1^1 & f_2^1 \\ f_1^2 & f_2^2 \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial x_1^*}{\partial a} \\ \frac{\partial x_2^*}{\partial a} \end{bmatrix} = \begin{bmatrix} -f_a^1 \\ -f_a^2 \end{bmatrix}$$

- This can be solved as

$$\begin{bmatrix} \frac{\partial x_1^*}{\partial a} \\ \frac{\partial x_2^*}{\partial a} \end{bmatrix} = \begin{bmatrix} f_1^1 & f_2^1 \\ f_1^2 & f_2^2 \end{bmatrix}^{-1} \cdot \begin{bmatrix} -f_a^1 \\ -f_a^2 \end{bmatrix}$$

Cramer's rule

- Cramer'srule shows that each of the comparative static derivatives can be solved as the ratio of two determinants.

$$\frac{dx_1^*}{da} = \frac{\begin{vmatrix} -f_a^1 & f_2^1 \\ -f_a^2 & f_2^2 \end{vmatrix}}{\begin{vmatrix} f_1^1 & f_2^1 \\ f_1^2 & f_2^2 \end{vmatrix}}, \frac{dx_2^*}{da} = \frac{\begin{vmatrix} f_1^1 & -f_a^1 \\ f_1^2 & -f_a^2 \end{vmatrix}}{\begin{vmatrix} f_1^1 & f_2^1 \\ f_1^2 & f_2^2 \end{vmatrix}}$$

- Suppose that the demand and supply functions for a product are given by:

$$\begin{aligned} q &= cp + a \text{ or } q - cp - a = 0 \quad (\text{demand}, \ c < 0) \\ q &= dp \text{ or } q - dp = 0 \quad (\text{supply}, \ d > 0) \end{aligned}$$

- Differentiate these two equations with respect to a yields:

$$\frac{dq^*}{da} - c\frac{dp^*}{da} - 1 = 0$$
$$\frac{dq^*}{da} - d\frac{dp^*}{da} = 0$$

In matrix form:

$$\begin{bmatrix} 1 & -c \\ 1 & -d \end{bmatrix} \cdot \begin{bmatrix} \frac{dq^*}{da} \\ \frac{dp^*}{da} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -c \\ 1 & -d \end{bmatrix} \cdot \begin{bmatrix} \frac{dq^*}{da} \\ \frac{dp^*}{da} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Therefore,

$$\frac{dq^*}{da} = \frac{\begin{vmatrix} 1 & -c \\ 0 & -d \end{vmatrix}}{\begin{vmatrix} 1 & -c \\ 1 & -d \end{vmatrix}} = \frac{-d}{c-d} = \frac{d}{d-c} > 0$$

$$\frac{dp^*}{da} = \frac{\begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}}{\begin{vmatrix} 1 & -c \\ 1 & -d \end{vmatrix}} = \frac{-1}{c-d} = \frac{1}{d-c} > 0$$