7. Optimization

Mwaso Mnensa 08-16-2022

Unconstrained optimization

- This is the simplest type of optimization
- The solution requires finding the critical points of a function
- For a single variable function f, the solution is given by $f' = 0$
- For a multivariate function, the solution requires that $\nabla F(\boldsymbol{x}) = \boldsymbol{0}$
- That is for every n, $\frac{\partial F}{\partial x_n} = 0$

Question

- Is the requirement to be a critical point a necessary or a sufficient condition?

Second Order (sufficient) conditions

Remember:

$Df(x^*)$	$D^2f(x^*)$	Max/Min
$= 0$	Negative semidefinite	Local max
$= 0$	Positive semidefinite	Local min
$= 0$	Neither	Saddle point or inflexion

Recap on +ve/-ve definiteness

- The $n \times n$ matrix A is positive definite if and only if its n principal minors are all greater than 0: $detA_1 > 0, detA_2 > 0,..., detA_n > 0.$
- The $n\times n$ matrix A is negative definite if and only if its n principal minors alternate in sign with the odd order ones being negative and the even order ones being positive: $det\ A_1 < 0, det\ A_2 > 0, detA_3 < 0, detA_4 > 0,....$
- The $n \times n$ matrix A is positive semidefinite if and only if its principal minors are all greater than or equal to 0: $det\ A_1 \geq 0\ and\ det\ A_2 \geq 0,..., A_n \geq 0.$
- The $n \times n$ matrix A is negative semidefinite if and only if its n principal minors alternate in sign with the odd order ones being less than or equal to zero and the even order ones being greater than or equal to 0:

  $det\ A_1 \leq 0\ and\ det\ A_2 \geq 0,\ . . . ..$

Example

$$f(x, y) = x^4 + x^2 - 6xy + 3y^2$$

Let's find the critical points and classify them as local max, local min or saddle points.

First find the Jacobian = $[\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y}]$

- $J = [4x^3 + 2x - 6y \qquad - 6x + 6y] = [0 \;\; 0]$
- Therefore x=y
- And $4x^3 + 2x - 6(x) = 0 \; or \; 4x^3 - 4x = 0$
- $x = 0 \; or \; 1 \; or \; -1$

Now Check the Hessian for 2^{nd} order conditions

- $H = \begin{bmatrix} 12x^2 + 2 & -6 \\ -6 & 6 \end{bmatrix}$

- Check all possible critical point (1,1) , (-1,-1) and (0,0)

- $(1,1)$: $A_1 = 12 + 2 = 14 \ \ and \ A_2 = 14 * 6 - (-6) * (-6) = 48$
- Both are +ve, therefore it is a local minimum
- (-1,-1): $A_1 = 14 \ \ and \ A_2 = 48. \quad (local\ minimum)$
- (0,0) : $A_1 = \ 2 * 6 - (-6 * -6) = \ -24 \ \ and \ A_2 = 6$
  $(saddle\ point)$

Exercise

- Find the critical points for this function:

$$f(x,\ y,\ z) = e^{x-y} + e^{y-x} + e^{x^2} + z^2.$$

- Are these points minimum, maximum or saddle points

Application: Profit maximization

Suppose a firm uses n inputs to produce a single product. $\mathbf{x} \in \mathbb{R}^n$ represents an input bundle. $y = Q(\mathbf{x})$ is the production function. p is the selling price of the product and $\mathbf{w}$ is the cost of inputs. The firm's profit function is

$$\pi(\mathbf{x}) = pQ(\mathbf{x}) - \mathbf{w}\mathbf{x}$$

First order conditions

$$\frac{\partial \pi}{\partial x_i}(\mathbf{x}^*) = 0$$

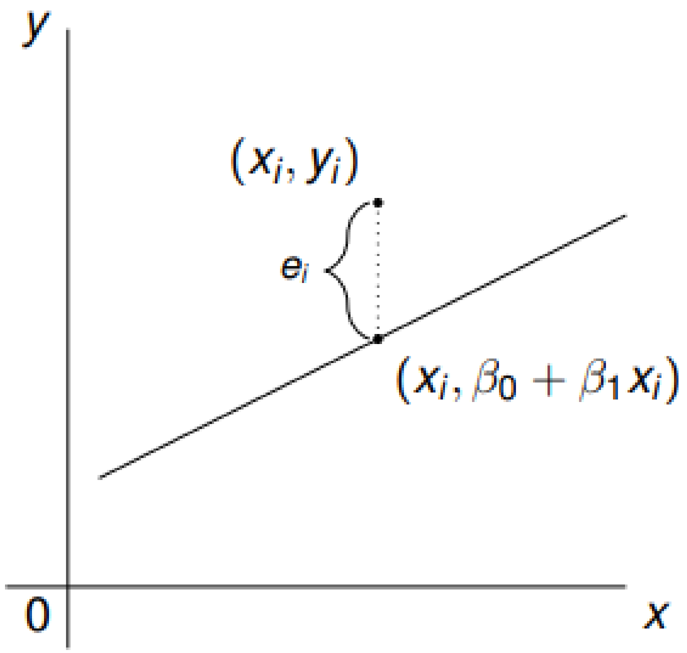
What does this imply? What is the second order necessary conditions? What does it imply?

Application: OLS

Suppose we want to estimate the following single variable linear model with N observations

$$y = \beta_0 + \beta_1 x + e$$

Our goal is to minimize the sum of the squared estimation error. Derive the estimator of β_0 and β_1.



$$\mathrm{SSE}(\beta) = \sum_{i=1}^{n} (Y_i - X_i\beta)^2 = \left(\sum_{i=1}^{n} Y_i^2\right) - 2\beta\left(\sum_{i=1}^{n} X_i Y_i\right) + \beta^2\left(\sum_{i=1}^{n} X_i^2\right).$$

Constrained optimization

- Consumers make decisions on what to buy constrained by the fact that their choice must be affordable.
- Firms make production decisions to maximize their profits subject to the constraint that they have limited production capacity.
- Households make decisions on how much to work/play with the constraint that there are only so many hours in the day.
- Firms minimize costs subject to the constraint that they have orders to fulfill.

Form of constrained optimization problem

- Maximize/Minimize

$$f(x_1, x_2, ..., x_n)$$

subject to:

$$g_1(x_1, x_2, ..., x_n) \leq b_1$$
$$\vdots$$
$$g_k(x_1, x_2, ..., x_n) \leq b_k$$

$$h_1(x_1, x_2, ..., x_n) = c_1$$
$$\vdots$$
$$h_m(x_1, x_2, ..., x_n) = c_m$$

Constrained optimization

- $f(x_1, \ldots, x_n)$ is called the objective function and
- $g_1, \ldots., g_k, h_1, \ldots, h_k$ are constraint functions

Typical consumer problem

$$u(x_1, x_2, ...x_n)$$

subject to:

$$p_1x_1 + p_2x_2 + \ ... \ + p_nx_n \leq I$$

and $x_i \geq 0 \ \forall i = 1, 2, 3, ..., n$

Simple optimization problem

$$f(x_1, x_2)$$

subject to:

$$h(x_1, x_2) = c$$

We want to choose the values of $(x_1,\ x_2)$ that maximize f and that are in the set $\{(x_1,\ x_2) :\ h(x_1,\ x_2)\ =\ c\}$.

The Lagrangian

Lagrange's Theorem: Let $f(x)$ and $g_j(x)$, $j = 1, \ldots, m$ be a continuously differentiable real-valued functions over some domain $D \to R^n$. Let x^* be an interior point of D and suppose that x^* is an optimum (maximum or minimum) of f subject to the constraints, $g_j(x) = 0$. If the gradient vectors $\nabla g_j(x^*)$, $j = 1, \ldots, m$, are linearly independent, then there exists m unique numbers λ_j^*, $j = 1, \ldots, m$, such that:

$$\frac{\partial L(x^*, \lambda^*)}{\partial x_i} = \frac{f(x^*)}{\partial x_i} + \sum_{j=1}^{m} \lambda_j^* \frac{\partial g_j(x^*)}{\partial x_i} = 0$$

For $i = 1, 2, \ldots, n$. $L(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x^*)$

In other words

Let $f, h_1, ..., h_m$ be C^1 functions of n variables. Consider the problem of maximizing (or minimizing) $f(\mathbf{x})$ on the constraint set

$$C_{\mathbf{h}} = \{\mathbf{x} = (x_1, ..., x_n) : h_1(\mathbf{x} = a_1, ..., h_m(\mathbf{x} = a_m)\}$$

Suppose that $\mathbf{x}^* \in C_{\mathbf{h}}$ and it is a (local) max or min of f on $C_{\mathbf{h}}$. Suppose further that $\mathbf{x}^*$ is not the critical point of $\mathbf{h} = (h_1, ..., h_m)$ (i.e.the rank of $D\mathbf{h}(\mathbf{x}^*)$ is $< m$). Then there exists real numbers $\mu_1^*, ...\mu_m^*$ such that $(x_1^*, ..., x_n^*, \mu_1^*, ...\mu_m^*)$ is a critical point of the Lagrangian function

$$L(\mathbf{x}^*, \boldsymbol{\mu}^*) \equiv f(\mathbf{x}) - \mu_1[h(\mathbf{x}) - a_1] - \cdots - \mu_m[h(\mathbf{x}) - a_m]$$

In other words, at (x_1^*, x_2^*, μ^*)

$$\frac{\partial L}{\partial x_1}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0, ..., \quad \frac{\partial L}{\partial x_n}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0$$

$$\frac{\partial L}{\partial \mu_1}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0, ..., \quad \frac{\partial L}{\partial \mu_m}(\mathbf{x}^*, \boldsymbol{\mu}^*) = 0$$

Second order conditions

To know that we have a maximum, all we really need is that the second differential of the objective function at the critical point is decreasing **along the constraint**.

By the implicit function theorem,

$$\frac{dx_2}{dx_1} = -\frac{\partial h/\partial x_1}{\partial h/\partial x_2}$$

Let $y = f(x_1, x_2(x_1))$ be the value of objective function subject to the constraint. By the chain rule,

$$\frac{dy}{dx_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dx_1} = \frac{\partial f}{\partial x_1} - \frac{\partial f}{\partial x_2}\frac{\partial h/\partial x_1}{\partial h/\partial x_2}$$

The second order sufficient condition requires that

$$\frac{d^2y}{dx_1^2} < 0$$

It can be shown that

$$\frac{d^2y}{dx_1^2} = \frac{-1}{(\partial h/\partial x_2)^2}\bar{D}$$

where $\bar{D}$ is the determinant of a **boarded Hessian** of L

$$\begin{pmatrix} 0 & \frac{\partial h}{\partial x_1} & \frac{\partial h}{\partial x_2} \\ \frac{\partial h}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \frac{\partial^2 L}{\partial x_1 \partial x_2} \\ \frac{\partial h}{\partial x_2} & \frac{\partial^2 L}{\partial x_2 \partial x_1} & \frac{\partial^2 L}{\partial x_2^2} \end{pmatrix}$$

Exercise

(Simon & Blume exercise 18.7)
Maximize $f(x, y, z) = yz + xz$ subject to $y^2 + z^2 = 1$ and $xz = 3$.

Exercise

$$- ax_1^2 - bx_2^2$$

subject to:

$$x_1 + x_2 = 1$$

Exercise

$$x_1^2 x_2$$

subject to

$$2x_1^2 + x_2^2 = 3$$