

Lecture X: Introduction to Multivariate Unconstrained and Constrained Optimality

Charles B. Moss ¹

¹University of Florida

September 13, 2012

- 1 Unconstrained Multivariate Optimality
 - First and Second Order Necessary Conditions
 - Development of the Constrained Optimum

First and Second Order Necessary Conditions

- As a first step, let us develop the first and second order necessary conditions for the multivariate function.
 - Like the univariate case, the development of the multivariate optimum is based on the Taylor series expansion.
 - The multivariate form of the Taylor series expansion is typically restricted to the first two moments by convenience, as

$$f(x) = f(x^*) + \nabla_x f(x^*) dx + \frac{1}{2} dx' \nabla_{xx}^2 f(x^*) dx. \quad (1)$$

- Again, subtracting $f(x^*)$ from each side we are left with

$$f(x) - f(x^*) = \nabla_x f(x) dx + \frac{1}{2} dx' \nabla_{xx}^2 f(x^*) dx \leq 0 \quad (2)$$

for a maximum at x^* .

- Thus, the first order necessary condition for the unconstrained maximum is that $\|\nabla_x f(x^*)\| \rightarrow 0$.
- The second order necessary condition for the unconstrained maximum is that $\nabla_{xx}^2 f(x^*)$ is negative semi-definite (or $x' \nabla_{xx}^2 f(x^*) x \leq 0$) for all x in the neighborhood around x^* .
- The later condition is if all of the eigenvalues of the Hessian matrix are less than or equal to zero.

- The first stage in our discussion will be a fairly general description of the problem at hand. We are now interested in developing the optimality conditions for the problem

$$\begin{aligned} \max_x f(x) \\ \text{s.t. } G(x) = b \end{aligned} \tag{3}$$

where $f(x)$ is a scalar valued function mapping the vector of inputs into a single value and $G(x)$ is a potentially vector valued function (i.e., it maps n inputs into m results which are potentially constrained).

- In order to solve the constrained problem we start by assuming that we have a value that solves the nonlinear constraints and form a linear approximation around that point

$$G(x) \approx G(x^*) + \nabla_x G(x^*) dx$$

$$\nabla_x G(x) = \begin{pmatrix} \frac{\partial G_1(x)}{\partial x_1} & \frac{\partial G_1(x)}{\partial x_2} & \dots & \frac{\partial G_1(x)}{\partial x_n} \\ \frac{\partial G_2(x)}{\partial x_1} & \frac{\partial G_2(x)}{\partial x_2} & \dots & \frac{\partial G_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G_m(x)}{\partial x_1} & \frac{\partial G_m(x)}{\partial x_2} & \dots & \frac{\partial G_m(x)}{\partial x_n} \end{pmatrix} \quad (4)$$

- Given the Taylor series expansion of the vector equation, we can see that starting from a feasible point $\bar{x}$ and stepping to another feasible point $\hat{x}$ involves solving the equation

$$G(\bar{x}) - G(\hat{x}) = \nabla_x G(\hat{x}) dx = 0 \quad (5)$$

thus any dx in the null space of $\nabla_x G(\hat{x})$ maintains feasibility.

- Simplifying this in the current case assume that $\nabla_x G(\hat{x}) dx$ can be written as

$$[G_1 G_2] \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} = 0 \quad (6)$$

where G_1 and G_2 are the partitioned values from the Jacobian matrix evaluated at $\hat{x}$.

- This partitioning would be such that G_1 is an m by m matrix where m is the number of equations in the vector valued function. Thus, G_2 is a m by $n - m$ matrix.

- This partitioning gives us the matrix equation

$$G_1 dx_1 + G_2 dx_2 = 0 \quad (7)$$

- Notice that this equation can be solved for dx_1 as

$$dx_1 = -G_1^{-1} G_2 dx_2 \quad (8)$$

which gives the change in x_1 associated with a change in x_2 which maintains feasibility in the constraints.

- Returning to the basic optimization problem, we develop our standard Taylor series expansion of the objective function

$$f(x) - f(x^*) = \nabla_x f(x^*) dx + \frac{1}{2} dx' \nabla_{xx}^2 f(x^*) dx \leq 0 \quad (9)$$

- Partitioning this expression in the same manner as the constraint matrix, we have

$$\begin{bmatrix} f_1 & f_2 \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} dx_1 & dx_2 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} dx_1 \\ dx_2 \end{bmatrix} \leq 0 \quad (10)$$

- Substituting for the feasibility constraint yields

$$\begin{bmatrix} f_1 & f_2 \end{bmatrix} \begin{bmatrix} -G_1^{-1} G_2 dx_2 \\ dx_2 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} -dx_2' G_2' G_1^{-1'} & dx_2 \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} -G_1^{-1} G_2 dx_2 \\ dx_2 \end{bmatrix} \leq 0 \quad (11)$$

- The first necessary condition for this problem can then be defined as

$$\begin{aligned} -f_1 G_1^{-1} G_2 dx_2 + f_2 dx_2 &= 0 \\ [f_2 - f_1 G_1^{-1} G_2] dx_2 &= 0 \\ \| f_2 - f_1 G_1^{-1} G_2 \| &= 0 \end{aligned} \quad (12)$$

- Intuitively, this condition implies that the gradient projected along a feasible path must be equal to zero at optimum. The second order necessary conditions are then defined by

$$\begin{aligned}
 & \begin{bmatrix} -dx_2' G_2' G_1^{-1'} & dx_2' \end{bmatrix} \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \begin{bmatrix} G_1^{-1} G_2 dx_2 \\ dx_2 \end{bmatrix} = \\
 & \begin{bmatrix} dx_2' G_2' G_1^{-1} F_{11} + dx_2' F_{21} & dx_2' G_2' G_1^{-1} F_{12} + dx_2' F_{22} \end{bmatrix} \begin{bmatrix} G_1^{-1} G_2 dx_2 \\ dx_2 \end{bmatrix} \\
 & \begin{bmatrix} dx_2' G_2' G_1^{-1} F_{11} G_1^{-1} G_2 dx_2 + dx_2' F_{21} G_1^{-1} G_2 dx_2 + \\ dx_2' G_2' G_1^{-1} F_{12} dx_2 + dx_2' F_{22} dx_2 \end{bmatrix} \\
 & dx_2' \left[G_2' G_1^{-1} F_{11} G_1^{-1} G_2 + F_{21} G_1^{-1} G_2 + G_2' G_1^{-1} F_{12} + F_{22} \right] dx_2 \\
 & \hspace{15em} (13)
 \end{aligned}$$