Second Order Conditions

Joseph Tao-yi Wang and Ya-Ju Tsai
2019/6/3

(Calculus 4, 19.3)

Recall the Hessian from Section 17.3

- Suppose $U \in \mathbf{R}^n$ is an open set, $F(\vec{x})$ is C^2

  Let $\vec{x}^*$ satisfies $\dfrac{\partial F}{\partial x_i}(\vec{x}) = 0, i = 1, \cdots, n$

- The Hessian at $\vec{x} = \vec{x}^*$ is

$$D^2F(\vec{x}) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(\vec{x}^*) & \cdots & \frac{\partial^2 F}{\partial x_n \partial x_1}(\vec{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_1 \partial x_n}(\vec{x}^*) & \cdots & \frac{\partial^2 F}{\partial x_n^2}(\vec{x}^*) \end{pmatrix}$$

Thm 17.2: Unconstrained Max. SOC

- Suppose $U \in \mathbf{R}^n$ is an open set, $F(\vec{x})$ is C^2

  Let $\vec{x}^*$ satisfies $\frac{\partial F}{\partial x_i}(\vec{x}) = 0, i = 1, \cdots, n$

1. If the Hessian $D^2F(\vec{x}^*)$ is negative definite, then $\vec{x}^*$ is a strict local max of F
2. If the Hessian $D^2F(\vec{x}^*)$ is positive definite, then $\vec{x}^*$ is a strict local min of F
3. If $D^2F(\vec{x}^*)$ is indefinite, then $\vec{x}^*$ is neither a strict local max nor a local min of F

Thm 19.7: One Equality Constraint SOC

- Let f, h be C^2 functions on $\mathbf{R}^2$
- To maximize f on the constraint set
$$C_h = \left\{(x, y) \middle| h(\vec{x}) = c\right\}$$
- Form the Lagrangian
$$\mathcal{L}(\vec{x}, \vec{\mu}) = f(\vec{x}) - \mu[h(\vec{x}) - c]$$
- (x^*, y^*) is a local constrained max of f on C_h if:

Thm 19.7: One Equality Constraint SOC

Exists μ^* such that (x^*, y^*, μ^*) satisfies:

1. $\frac{\partial \mathcal{L}}{\partial x} = 0, \quad \frac{\partial \mathcal{L}}{\partial y} = 0, \quad \frac{\partial \mathcal{L}}{\partial \mu} = 0 \text{ at } (x^*, y^*, \mu^*)$

2. $\det \begin{pmatrix} 0 & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{pmatrix} > 0 \text{ at } (x^*, y^*, \mu^*)$

Thm 19.6: Equality Constraints SOC

- Let f, $h_1, \ldots, h_k$ be C^2 functions on $\mathbf{R}^n$
- To maximize f on the constraint set

$$C_h = \{\vec{x} : h_1(\vec{x}) = c_1, \cdots, h_k(\vec{x}) = c_k\}$$

- Form the Lagrangian

$$\mathcal{L}(\vec{x}, \vec{\mu}) = f(\vec{x}) - \mu_1[h_1(\vec{x}) - c_1] - \cdots - \mu_k[h_k(\vec{x}) - c_k]$$

- $\vec{x}^*$ is a strict local constrained max of f on C_h if:

Thm 19.6: Equality Constraints SOC

1. $\vec{x}^* \in C_h$
2. Exists $\vec{\mu}^* = (\mu_1^*, \cdots, \mu_k^*)$ such that at $(\vec{x}^*, \vec{\mu}^*)$
   $$\frac{\partial \mathcal{L}}{\partial x_i} = 0 \; (i=1,\ldots,n), \quad \frac{\partial \mathcal{L}}{\partial \mu_j} = 0 \; (j=1,\ldots,k)$$
3. $D^2_{\vec{x}} \mathcal{L}(\vec{x}^*, \vec{\mu}^*)$, Hessian of $\mathcal{L}$ w.r.t. $\vec{x}$ at $(\vec{x}^*, \vec{\mu}^*)$ is negative definite on the linear constraint set $\{\vec{v} : D\vec{h}(\vec{x}^*)\vec{v} = \vec{0}\}$.

i.e. If $\vec{v} \neq \vec{0}$ and $D\vec{h}(\vec{x}^*)\vec{v} = \vec{0}$

$$\Rightarrow \vec{v}^T \left[D^2_{\vec{x}} \mathcal{L}(\vec{x}^*, \vec{\mu}^*) \right] \vec{v} < 0$$

- How can we check condition 3.?

Condition 3. of Thm 19.6

3. $D^2_{\vec{x}}\mathcal{L}(\vec{x}^*, \vec{\mu}^*)$, Hessian of $\mathcal{L}$ w.r.t. $\vec{x}$ at $(\vec{x}^*, \vec{\mu}^*)$ is negative definite on the linear constraint set $\{\vec{v} : D\vec{h}(\vec{x}^*)\vec{v} = \vec{0}\}$.

- Form $H = \begin{pmatrix} \vec{0} & D\vec{h}(\vec{x}^*) \\ D\vec{h}(\vec{x}^*)^T & D^2_{\vec{x}}\mathcal{L}(\vec{x}^*, \vec{\mu}^*) \end{pmatrix}$
- If the last *(n-k)* LPM of matrix H alternate in sign, and $(-1)^n \cdot \det H > 0$
- Then Condition 3. of Thm 19.6 holds.

Hessian for Thm 19.6: Equality Constraints

$$H = \begin{pmatrix} 0 & \cdots & 0 & | & \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & | & \frac{\partial h_k}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_n} \\ - & - & - & - & - & - & - \\ \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_1} & | & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n} & \cdots & \frac{\partial h_k}{\partial x_n} & | & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2} \end{pmatrix}$$

Similar to Hessian of the Lagrangian!!

$$D^2_{(\vec{\mu},\vec{x})}\mathcal{L} = \begin{pmatrix} 0 & \cdots & 0 & | & -\frac{\partial h_1}{\partial x_1} & \cdots & -\frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & | & -\frac{\partial h_k}{\partial x_1} & \cdots & -\frac{\partial h_k}{\partial x_n} \\ - & - & - & - & - & - & - \\ -\frac{\partial h_1}{\partial x_1} & \cdots & -\frac{\partial h_k}{\partial x_1} & | & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\ -\frac{\partial h_1}{\partial x_n} & \cdots & -\frac{\partial h_k}{\partial x_n} & | & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2} \end{pmatrix}$$

- Multiplying each row and each column by *(-1)* does not change the determinant and LPMs!

Thm 16.4: Definiteness of Quadratic Forms

- For $Q(\vec{x}^*) = \vec{x}^T A \vec{x}$ restricted to $B\vec{x} = \vec{0}$,
- Check the bordered matrix

$$H = \begin{pmatrix} \vec{0} & B \\ B^T & A \end{pmatrix}_{(n+m)\times(n+m)}$$

- and its Leading Principal Minors (LPM)
- The k-th order LPM is the determinant of the leading principal matrix A_k, derived by deleting last n-k rows and columns of a matrix A

Thm 16.4: Definiteness of Quadratic Forms

$$H = \begin{pmatrix} \vec{0} & B \\ B^T & A \end{pmatrix}_{(n+m)\times(n+m)}$$

1. If $(-1)^n \cdot \det H > 0$ and the last n-m LPM alternate in sign, then Q is negative definite ($\vec{x} = \vec{0}$ is a strict local max of F on $B\vec{x} = \vec{0}$)
2. If $\det H$ and its last n-m LPM all have same sign as $(-1)^m$, then Q is positive definite ($\vec{x} = \vec{0}$ is a strict local min of F on $B\vec{x} = \vec{0}$)
3. If neither, then Q is indefinite on $B\vec{x} = \vec{0}$

Thm 19.8: Add Inequality Constraints

- Let f, $g_1, \ldots, g_m$, $h_1, \ldots, h_k$ be C^2 functions on $\mathbf{R}^n$
- To maximize f on the constraint set

$$C_{g,h} = \{\vec{x} \mid g_1(\vec{x}) \le b_1, \cdots, g_m(\vec{x}) \le b_m, \\ h_1(\vec{x}) = c_1, \cdots, h_k(\vec{x}) = c_k\}$$

- Form the Lagrangian $\mathcal{L}(\vec{x}, \vec{\lambda}, \vec{\mu}) =$

$$f(\vec{x}) - \lambda_1[g_1(\vec{x}) - b_1] - \cdots - \lambda_m[g_m(\vec{x}) - b_m] \\ - \mu_1[h_1(\vec{x}) - c_1] - \cdots - \mu_k[h_k(\vec{x}) - c_k]$$

- $\vec{x}^*$ is a strict local constrained max of f on $C_{g,h}$ if:

Thm 19.8: Add Inequality Constraints

1. Exists $\vec{\lambda}^* = (\lambda_1^*, \cdots, \lambda_m^*), \vec{\mu}^* = (\mu_1^*, \cdots, \mu_k^*)$

- such that at $(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*)$, $\frac{\partial \mathcal{L}}{\partial x_i} = 0$ $(i=1,\ldots,n)$
- $\lambda_j^* \geq 0, \quad \lambda_j^* \left[g_j(\vec{x}^*) - b_1 \right] = 0$ $(j=1,\ldots,k)$

  $h_1(\vec{x}^*) = c_1, \cdots, h_k(\vec{x}^*) = c_k$

2. $\vec{g}_E = (g_1, \cdots, g_e)$ binding $(g_{e+1},\ldots, g_m$ not)

$D^2_{\vec{x}}\mathcal{L}(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*)$, Hessian of $\mathcal{L}$ wrt $\vec{x}$ at $(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*)$ is negative definite on the linear constraint set

$$\{\vec{v} : D\vec{g}_E(\vec{x}^*)\vec{v} = \vec{0} \text{ and } D\vec{h}(\vec{x}^*)\vec{v} = \vec{0}\}$$

Condition 2. of Thm 19.8

2. $\vec{g}_E = (g_1, \cdots, g_e)$ binding ($g_{e+1}, \ldots, g_m$ not)

$D^2_{\vec{x}}\mathcal{L}(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*)$, Hessian of $\mathcal{L}$ wrt $\vec{x}$ at $(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*)$

is negative definite on the linear constraint set

$$\{\vec{v} : D\vec{g}_E(\vec{x}^*)\vec{v} = \vec{0} \text{ and } D\vec{h}(\vec{x}^*)\vec{v} = \vec{0}\}$$

I.e. $\vec{v} \neq \vec{0}, D\vec{g}_E(\vec{x}^*)\vec{v} = \vec{0}, D\vec{h}(\vec{x}^*)\vec{v} = \vec{0}$

$$\Rightarrow \vec{v}^T \cdot \left[D^2_{\vec{x}}\mathcal{L}(\vec{x}^*, \vec{\lambda}^*, \vec{\mu}^*)\right] \cdot \vec{v} < 0$$

Bordered Hessian H and check if the last n-$(e+k)$ LPM of matrix H alternate in sign, and

$$(-1)^n \cdot \det H > 0$$

Hessian for Thm 19.8: + Inequality Constraint

$$
\begin{pmatrix}
0 & \cdots & 0 & 0 & \cdots & 0 & | & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & | & \frac{\partial g_e}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_n} \\
0 & \cdots & 0 & 0 & \cdots & 0 & | & \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & | & \frac{\partial h_k}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_n} \\
- & - & - & - & - & - & - & & & \\
\frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_e}{\partial x_1} & \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_k}{\partial x_1} & | & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & | & \vdots & \ddots & \vdots \\
\frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_e}{\partial x_n} & \frac{\partial h_1}{\partial x_n} & \cdots & \frac{\partial h_k}{\partial x_n} & | & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2}
\end{pmatrix}
$$