Math 408A
Testing Positive Definiteness

Second-Order Sufficiency
and
Testing Positive Definiteness

January 25, 2012

More on Second-Order Sufficient Conditions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable at $\overline{x} \in \mathbb{R}^n$.

More on Second-Order Sufficient Conditions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable at $\overline{x} \in \mathbb{R}^n$. If $\nabla f(\overline{x}) = 0$ and $\nabla^2 f(\overline{x})$ is positive definite, then there is an $\alpha > 0$ such that $f(x) \geq f(\overline{x}) + \alpha\|x - \overline{x}\|^2$ for all x near $\overline{x}$.

More on Second-Order Sufficient Conditions

Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable at $\overline{x} \in \mathbb{R}^n$. If $\nabla f(\overline{x}) = 0$ and $\nabla^2 f(\overline{x})$ is positive definite, then there is an $\alpha > 0$ such that $f(x) \geq f(\overline{x}) + \alpha \|x - \overline{x}\|^2$ for all x near $\overline{x}$.

To use this sufficiency condition we need a method for testing for positive definiteness. Of course, we could compute the eigenvalues. But this requires solving for the roots of an nth degree polynomial (the eigenvalues). We look at an alternative approach that can sometimes be simpler.

Classification of Critical Points

Let $H \in \mathbb{R}^{n \times n}$ be symmetric. We define the kth principal minor of H, denoted $\Delta_k(H)$, to be the determinant of the upper-left $k \times k$ submatrix of H. Then

Classification of Critical Points

Let $H \in \mathbb{R}^{n \times n}$ be symmetric. We define the kth principal minor of H, denoted $\Delta_k(H)$, to be the determinant of the upper-left $k \times k$ submatrix of H. Then

1. H is positive definite if and only if
   $\Delta_k(H) > 0, \;\; k = 1, 2, \ldots, n.$

Classification of Critical Points

Let $H \in \mathbb{R}^{n \times n}$ be symmetric. We define the kth principal minor of H, denoted $\Delta_k(H)$, to be the determinant of the upper-left $k \times k$ submatrix of H. Then

1. H is positive definite if and only if $\Delta_k(H) > 0, \; k = 1, 2, \ldots, n.$
2. H is negative definite if and only if $(-1)^k \Delta_k(H) > 0, \; k = 1, 2, \ldots, n.$

Classification of Critical Points

$$H = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & 1 \\ -1 & 1 & 4 \end{bmatrix}.$$

Classification of Critical Points

$$H = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & 1 \\ -1 & 1 & 4 \end{bmatrix}.$$

We have

$$\Delta_1(H) = 1, \quad \Delta_2(H) = \begin{vmatrix} 1 & 1 \\ 1 & 5 \end{vmatrix} = 4, \quad \text{and} \quad \Delta_3(H) = \det(H) = 8.$$

Classification of Critical Points

$$H = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & 1 \\ -1 & 1 & 4 \end{bmatrix}.$$

We have

$$\Delta_1(H) = 1, \quad \Delta_2(H) = \begin{vmatrix} 1 & 1 \\ 1 & 5 \end{vmatrix} = 4, \quad \text{and} \quad \Delta_3(H) = \det(H) = 8.$$

Therefore, H is positive definite.

Classification of Critical Points

If the symmetric matrix H is neither positive or negative semi-definite, we say that it is indefinite.

Classification of Critical Points

If the symmetric matrix H is neither positive or negative semi-definite, we say that it is indefinite.

Definition: *A critical point that is neither a local maximum or minimum is called a* **saddle point***.*

Classification of Critical Points

If the symmetric matrix H is neither positive or negative semi-definite, we say that it is indefinite.

Definition: *A critical point that is neither a local maximum or minimum is called a* **saddle point**.

Theorem: *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable at $\bar{x}$. If $\nabla f(\bar{x}) = 0$ and $\nabla^2 f(\bar{x})$ is indefinite, then $\bar{x}$ is a saddle point of f.*

Classification of Critical Points

If the symmetric matrix H is neither positive or negative semi-definite, we say that it is indefinite.

Definition: *A critical point that is neither a local maximum or minimum is called a* **saddle point**.

Theorem: *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable at $\bar{x}$. If $\nabla f(\bar{x}) = 0$ and $\nabla^2 f(\bar{x})$ is indefinite, then $\bar{x}$ is a saddle point of f.*

Theorem: *Let $H \in \mathbb{R}^{n \times n}$ be symmetric. If H is neither positive or negative definite and all of its principal minors are non-zero, then H is indefinite.*

Example

Compute and classify the critical poinits of

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4.$$

Example

Compute and classify the critical poinits of

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4.$$

$$\nabla f(x) = \begin{pmatrix} 2x_1 + 16x_2 \\ 4x_2^3 + 16x_1 \end{pmatrix} \qquad \nabla^2 f(x) = \begin{bmatrix} 2 & 16 \\ 16 & 12x_2^2 \end{bmatrix}$$

Example

Compute and classify the critical poinits of

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4.$$

$$\nabla f(x) = \begin{pmatrix} 2x_1 + 16x_2 \\ 4x_2^3 + 16x_1 \end{pmatrix} \qquad \nabla^2 f(x) = \begin{bmatrix} 2 & 16 \\ 16 & 12x_2^2 \end{bmatrix}$$

$$\nabla f(x) = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x_1 = -8x_2 \\ x_2^3 = -4x_1 \end{pmatrix} \quad \Leftrightarrow \quad x_1 = x_2 = 0 \text{ or } x_2^2 = 2^5$$

Example

Compute and classify the critical poinits of

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4.$$

$$\nabla f(x) = \begin{pmatrix} 2x_1 + 16x_2 \\ 4x_2^3 + 16x_1 \end{pmatrix} \qquad \nabla^2 f(x) = \begin{bmatrix} 2 & 16 \\ 16 & 12x_2^2 \end{bmatrix}$$

$$\nabla f(x) = 0 \quad \Leftrightarrow \quad \begin{pmatrix} x_1 = -8x_2 \\ x_2^3 = -4x_1 \end{pmatrix} \quad \Leftrightarrow \quad x_1 = x_2 = 0 \text{ or } x_2^2 = 2^5$$

The critical points are

$$x^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \; x^2 = \begin{pmatrix} -32\sqrt{2} \\ 4\sqrt{2} \end{pmatrix}, \; x^3 = \begin{pmatrix} 32\sqrt{2} \\ -4\sqrt{2} \end{pmatrix}$$

Example

Next observe that the function f is coercive since

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4 = (x_1 + 8x_2)^2 + x_2^4 - 64x_2^2.$$

Example

Next observe that the function f is coercive since

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4 = (x_1 + 8x_2)^2 + x_2^4 - 64x_2^2.$$

Therefore a global minimizer must exist, and it must be one of the critical points.

Example

Next observe that the function f is coercive since

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4 = (x_1 + 8x_2)^2 + x_2^4 - 64x_2^2.$$

Therefore a global minimizer must exist, and it must be one of the critical points.
It must be the critical point having smallest function value.

Example

Next observe that the function f is coercive since

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4 = (x_1 + 8x_2)^2 + x_2^4 - 64x_2^2.$$

Therefore a global minimizer must exist, and it must be one of the critical points.
It must be the critical point having smallest function value.

$$f(x^1) = 0 \qquad f(x^2) = f(x^3) = 2^1 1 - 2^4 2^{11/2} 2^{5/2} + 2^{10} = -2^{10}.$$

Example

Next observe that the function f is coercive since

$$f(x) = x_1^2 + 16x_1x_2 + x_2^4 = (x_1 + 8x_2)^2 + x_2^4 - 64x_2^2.$$

Therefore a global minimizer must exist, and it must be one of the critical points.
It must be the critical point having smallest function value.

$$f(x^1) = 0 \qquad f(x^2) = f(x^3) = 2^1 1 - 2^4 2^{11/2} 2^{5/2} + 2^{10} = -2^{10}.$$

So x^2 and x^3 are global minimizers.

Example

Let us also look at this problem using second-order information.

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 16 \\ 16 & 12x_2 \end{bmatrix}$$

So

$$|\nabla^2 f(x^1)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 0 \end{bmatrix} \right| = -2^8, \quad |\nabla^2 f(x^2)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 3 \cdot 2^7 \end{bmatrix} \right| = 2^9 = |\nabla^2 f(x^3)|$$

Example

Let us also look at this problem using second-order information.

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 16 \\ 16 & 12x_2 \end{bmatrix}$$

So

$$|\nabla^2 f(x^1)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 0 \end{bmatrix} \right| = -2^8, \quad |\nabla^2 f(x^2)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 3 \cdot 2^7 \end{bmatrix} \right| = 2^9 = |\nabla^2 f(x^3)|$$

$\nabla^2 f(x^1)$ is indefinite, so x^1 is a saddle point.

Example

Let us also look at this problem using second-order information.

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 16 \\ 16 & 12x_2 \end{bmatrix}$$

So

$$|\nabla^2 f(x^1)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 0 \end{bmatrix} \right| = -2^8, \quad |\nabla^2 f(x^2)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 3 \cdot 2^7 \end{bmatrix} \right| = 2^9 = |\nabla^2 f(x^3)|$$

$\nabla^2 f(x^1)$ is indefinite, so x^1 is a saddle point.
$\nabla^2 f(x^2)$ and $\nabla^2 f(x^3)$ are positive definite, so x^2 and x^3 are local minimizers.

Example

Let us also look at this problem using second-order information.

$$\nabla^2 f(x) = \begin{bmatrix} 2 & 16 \\ 16 & 12x_2 \end{bmatrix}$$

So

$$|\nabla^2 f(x^1)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 0 \end{bmatrix} \right| = -2^8, \quad |\nabla^2 f(x^2)| = \left| \begin{bmatrix} 2 & 16 \\ 16 & 3 \cdot 2^7 \end{bmatrix} \right| = 2^9 = |\nabla^2 f(x^3)|$$

$\nabla^2 f(x^1)$ is indefinite, so x^1 is a saddle point.
$\nabla^2 f(x^2)$ and $\nabla^2 f(x^3)$ are positive definite, so x^2 and x^3 are local minimizers.

We have already established that x^2 and x^3 are global minimizers using the coercivity of f.

Operations that Preserve Convexity

Let $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex functions for $i = 1, 2, \ldots, m$, and let $\lambda_i \geq 0,\ i = 1, \ldots, m$. Then the following functions are also convex.

Operations that Preserve Convexity

Let $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex functions for $i = 1, 2, \ldots, m$, and let $\lambda_i \geq 0, \; i = 1, \ldots, m$. Then the following functions are also convex.

1. $f(x) := \phi(f_1(x))$, where $\phi : \mathbb{R} \to \mathbb{R}$ is any non-decreasing function on $\mathbb{R}$.

Operations that Preserve Convexity

Let $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex functions for $i = 1, 2, \ldots, m$, and let $\lambda_i \geq 0,\ i = 1, \ldots, m$. Then the following functions are also convex.

1. $f(x) := \phi(f_1(x))$, where $\phi : \mathbb{R} \to \mathbb{R}$ is any non-decreasing function on $\mathbb{R}$.
2. $f(x) := \sum_{i=1}^{m} \lambda_i f_i(x)$ (Non-negative linear combinations)

Operations that Preserve Convexity

Let $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex functions for $i = 1, 2, \ldots, m$, and let $\lambda_i \geq 0,\ i = 1, \ldots, m$. Then the following functions are also convex.

1. $f(x) := \phi(f_1(x))$, where $\phi : \mathbb{R} \to \mathbb{R}$ is any non-decreasing function on $\mathbb{R}$.
2. $f(x) := \sum_{i=1}^{m} \lambda_i f_i(x)$ (Non-negative linear combinations)
3. $f(x) := \max\{f_1(x), f_2(x), \ldots, f_m(x)\}$ (pointwise max)

Operations that Preserve Convexity

Let $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex functions for $i = 1, 2, \ldots, m$, and let $\lambda_i \geq 0,\ i = 1, \ldots, m$. Then the following functions are also convex.

1. $f(x) := \phi(f_1(x))$, where $\phi : \mathbb{R} \to \mathbb{R}$ is any non-decreasing function on $\mathbb{R}$.
2. $f(x) := \sum_{i=1}^m \lambda_i f_i(x)$ (Non-negative linear combinations)
3. $f(x) := \max\{f_1(x), f_2(x), \ldots, f_m(x)\}$ (pointwise max)
4. $f(x) := \inf\left\{\sum_{i=1}^m f_i(x^i) \,\middle|\, x = \sum_{i=1}^m x^i\right\}$ (infimal convolution)

Operations that Preserve Convexity

Let $f_i : \mathbb{R}^n \to \bar{\mathbb{R}}$ be convex functions for $i = 1, 2, \ldots, m$, and let $\lambda_i \geq 0,\ i = 1, \ldots, m$. Then the following functions are also convex.

1. $f(x) := \phi(f_1(x))$, where $\phi : \mathbb{R} \to \mathbb{R}$ is any non-decreasing function on $\mathbb{R}$.
2. $f(x) := \sum_{i=1}^m \lambda_i f_i(x)$ (Non-negative linear combinations)
3. $f(x) := \max\{f_1(x), f_2(x), \ldots, f_m(x)\}$ (pointwise max)
4. $f(x) := \inf\left\{\sum_{i=1}^m f_i(x^i) \,\middle|\, x = \sum_{i=1}^m x^i\right\}$ (infimal convolution)
5. $f_1^*(y) := \sup_{x\in\mathbb{R}^n}[y^T x - f_1(x)]$ (convex conjugation)

More Examples of Convex Functions

Let $C \subset \mathbb{R}^n$ be a closed convex set, and let $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a convex function.

More Examples of Convex Functions

Let $C \subset \mathbb{R}^n$ be a closed convex set, and let $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a convex function.

The convex indicator of C:

$$\delta_C(x) := \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{otherwise.} \end{cases}$$

More Examples of Convex Functions

Let $C \subset \mathbb{R}^n$ be a closed convex set, and let $h : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a convex function.

The convex indicator of C:

$$\delta_C(x) := \begin{cases} 0, & \text{if } x \in C \\ +\infty, & \text{otherwise.} \end{cases}$$

The support function of C:

$$\sigma_C(x) := \sup \left\{ x^T y \ | y \in C \right\}$$

The distance function to C:

$$d_C(x) := \operatorname{dist}(x|C) := \inf \{ \|x - y\| \ | y \in C \}$$