Lecture 25
Nonlinear Programming

November 9, 2009
Outline

• Nonlinear Programming
  • Another example of NLP problem
  • What makes these problems complex
• Scalar Function
  □ Unconstrained Problem
  □ Local and global optima: definition, characterization
  □ Convex and Concave Functions
• Multivariate Function
  □ Unconstrained Problem
  □ Local and global optima: definition, characterization
  □ Convex and Concave Functions
Another Example: Portfolio Selection with Risky Securities

Consider investing in $n$ stocks, knowing the following information

- The current price of each share of stock $j$ is $p_j$
- Our budget for buying the shares is $B$
- The expected return of each share is $\mu_j$
- Being risk averse, we want the return to be at least some amount $L$
- The mutual risk of investment in stocks $i$ and $j$ is $\sigma_{ij}$

We want to decide on the number of shares of each stock $j$, so as to minimize the total risk of the investment. Formulate the problem as NLP.
• Let $x_j$ be the number of shares of stock $j$

• Budget constraint

$$\sum_{j=1}^{n} p_j x_j \leq B$$

• Risk averse constraint
  
  • Expected return is $\sum_{j=1}^{n} \mu_j x_j$
  
  • Expected return should be at least $L$

$$\sum_{j=1}^{n} \mu_j x_j \geq L$$

• Total variance (risk) of the investment

$$V(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j$$
Putting this all together, we have

\[
\text{minimize } V(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j
\]

subject to

\[
\sum_{j=1}^{n} p_j x_j \leq B
\]

\[
\sum_{j=1}^{n} \mu_j x_j \geq L
\]

\[x_j \geq 0 \text{ for all } j\]

William Sharpe and Harry Markowitz received Nobel Prize in Economics 1990 for developing (nlp) model to study the effects of asset risk, return, correlation and diversification of investment portfolio returns.
Why are NLP problems complex

- Nonlinearities itself are not that bad
- Complexity comes from “nonconvexities”
- Nonconvexity gives rise to existence of multiple local and global minima
- Lack of efficient (practical) mathematical tools to distinguish among them
Unconstrained Optimization: Scalar Function

Given scalar function $f : \mathbb{R} \to \mathbb{R}$, determine

$$\min_{x \in \mathbb{R}} f(x)$$

The function may have local and global minima. These can cause difficulties when solving the problem.
Local and Global Minima

A point $x^* \in \mathbb{R}$ is a **local minimum** if there exists an open interval $(a, b)$ containing $x^*$ and such that

$$f(x^*) \leq f(x) \quad \text{for all } x \in (a, b)$$

A point $x^* \in \mathbb{R}$ is a **global (absolute) minimum** if

$$f(x^*) \leq f(x) \quad \text{for all } x \in \mathbb{R}$$

![Graph showing local and global minima](image)
In the figure, suppose $f(x)$ continues to go up for $x < 0$ and stays on level 0 for $x \geq 5$.

Then, $x^* = 1$ and $x^* = 3$ are local minima, while all the points in the interval $[5, +\infty)$ are global minima.

Suppose now that $f(x)$ starts growing after $x = 5$.

Then, $x^* = 1$ and $x^* = 3$ are local minima, while $x^* = 5$ is a unique global minimum.

If $f(x)$ keeps decreasing after $x = 5$, then a global minimum does not exist.
Local and Global Maxima

A point \( x^* \in \mathbb{R} \) is a **local maximum** if there exists an open interval \((a, b)\) containing \( x^* \) and such that

\[
f(x^*) \geq f(x) \quad \text{for all } x \in (a, b)
\]

A point \( x^* \in \mathbb{R} \) is a **global (absolute) maximum** if

\[
f(x^*) \geq f(x) \quad \text{for all } x \in \mathbb{R}
\]
Concave and Convex Functions

Concave and convex functions do not have local and global minima

All maxima are global for a concave function

All minima are global for convex function
Formal Characterization of Minima/Maxima

Let $f'$ denote the derivative of $f$
If $x^*$ is a local (or global) minimum it satisfies the following relation

$$f'(x^*) = 0$$

If $x^*$ is a local (or global) maximum it satisfies the same relation
The points $x^*$ that are solution to $f'(x^*) = 0$ are referred to as **stationary points** of $f$.
The stationary points include local and global minima and maxima, but may also include the points that are none of the above such as **inflection points**

**Example**: $f(x) = x^3$. The only solution to $f'(x) = 0$ is $x = 0$, which is an inflection point.
How do we find the minima and maxima in this case?

We find all the solutions of the equation $f'(x) = 0$ and then

- We check the second derivative at each of these points $x^*$:
  - If $f''(x^*) < 0$, then $x^*$ is a local maximum
  - If $f''(x^*) > 0$, then $x^*$ is a local minimum
  - If $f''(x^*) = 0$, then we need higher order derivatives (until the sign is nonzero)

- $f^{(m)} > 0$ and $m$ even, then $x^*$ is a local minimum; otherwise it is inflection point
- $f^{(m)} < 0$ and $m$ even, then $x^*$ is a local maximum; otherwise it is inflection point

where $f^{(m)}$ denotes the $m$-th derivative
Once we have all local minima (maxima), we find the global as the $x^*$ with the smallest of the values $f(x^*)$, and making sure that the function is bounded below (above).

Bounded below means: $f(x)$ does not decrease to infinity when $x \to -\infty$ or $x \to +\infty$. 

![Graph showing a function with local minima and maxima.](image-url)
Formal Characterization of Convex and Concave Scalar Functions

Function $f$ is concave over $\mathbb{R}$ under any of the following conditions:

- $f'(x)$ is decreasing function over $\mathbb{R}$
- $f''(x) \leq 0$ for all $x \in \mathbb{R}$
Function \( f \) is convex over \( \mathbb{R} \) under any of the following conditions:

- \( f'(x) \) is increasing function over \( \mathbb{R} \)
- \( f''(x) \geq 0 \) for all \( x \in \mathbb{R} \)

Every function \( f \) is either

- Convex
- Concave, or
- Nether convex nor concave

Example Characterize the following functions as convex, concave, or neither convex nor concave

\[
\begin{align*}
  f(x) &= x^2 \\
  f(x) &= 1 - 2x^2 \\
  f(x) &= x^3 - 1
\end{align*}
\]
Unconstrained Optimization: Multivariate Function
Unconstrained Problem

Given scalar function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, determine

$$\min_{x \in \mathbb{R}^n} f(x)$$

The function may have **local and global minima**, which cause difficulties when solving the problem.
Local and Global Minima and Maxima

A point $x^* \in \mathbb{R}^n$ is a **local minimum** if there exists an open ball $B$ centered at $x^*$ and such that

$$f(x^*) \leq f(x) \quad \text{for all } x \in B$$

A point $x^* \in \mathbb{R}^n$ is a **global (absolute) minimum** if

$$f(x^*) \leq f(x) \quad \text{for all } x \in \mathbb{R}^n$$

A point $x^* \in \mathbb{R}^n$ is a **local maximum** if there exists an open ball $B$ centered at $x^*$ and such that

$$f(x^*) \geq f(x) \quad \text{for all } x \in B$$

A point $x^* \in \mathbb{R}^n$ is a **global (absolute) maximum** if

$$f(x^*) \geq f(x) \quad \text{for all } x \in \mathbb{R}^n$$
Concave and Convex Functions

Concave and convex functions do not have local and global minima

For a concave function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, every maximum is global

For convex function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, every minimum is global
Formal Characterization of Minima/Maxima

Let $\nabla f$ denote the gradient of $f$, i.e.,

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f(x)}{\partial x_1} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{bmatrix}$$

Every (local and global) minimum $x^*$ satisfies the following relation

$$\nabla f(x^*) = 0$$

Also, every (local or global) maximum $x^*$ satisfies the above relation

The points $x^*$ that satisfy $\nabla f(x^*) = 0$ are **stationary points** of $f$. 
The stationary points include local and global minima and maxima, but may also include the points that are none of the above such as saddle points.

**Example:** \( f(x_1, x_2) = x_1 x_2 \). The only solution to \( \nabla f(x) = 0 \) over \( \mathbb{R}^n \) is

\[
\begin{align*}
x_1^* &= 0 \\
x_2^* &= 0
\end{align*}
\]

The function can grow to \(+\infty\) and decrease to \(-\infty\), so \( f \) does not have any local (or global) minima or maxima.

The point \( x^* = (0, 0) \) is a saddle point.
How do we characterize the minima and maxima in this case?

In theory, we find all the solutions of the equation $\nabla f(x) = 0$ and then we rely on the second order information, the **Hessian of the function** $f$

The Hessian of $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is the $n \times n$ matrix, denoted $\nabla^2 f(x)$ and with entries given by

$$[\nabla^2 f(x)]_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$
Example of Hessians: The Hessian of $f(x_1, x_2) = x_1 x_2$ is

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The Hessian of $f(x_1, x_2) = x_1^3 - x_2^4$ is

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 6x_1 & 0 \\ 0 & -12x_2^2 \end{bmatrix}$$
Having the solutions $x^\ast$ to $\nabla f(x) = 0$, to determine the nature of the points $x^\ast$, we do the following:

- We check the Hessian $\nabla^2 f(x^\ast)$ at each of the points $x^\ast$:
  - If the matrix $\nabla^2 f(x^\ast)$ is negative definite [$\nabla^2 f(x^\ast) < 0$], then $x^\ast$ is a local maximum
  - If the matrix $\nabla^2 f(x^\ast)$ is positive definite [$\nabla^2 f(x^\ast) > 0$], then $x^\ast$ is a local minimum
  - If $\nabla^2 f(x^\ast) = 0$, then we do not know what is happening

What helps us here in some situations is that the function is “nice” for the given optimization problem:

- $f$ is convex and we need to minimize $f$
- $f$ is concave and we need to maximize $f$
Formal Characterization of Convex and Concave Multivariate Functions

Function $f$ is **concave over** $\mathbb{R}^n$ when:

- The Hessian matrix $\nabla^2 f(x)$ is negative semidefinite for all $x$,

$$\nabla^2 f(x) \leq 0 \quad \text{for all } x \in \mathbb{R}^n$$
Function $f$ is **convex over** $\mathbb{R}^n$ when

- its Hessian is positive semidefinite for all $x$:

$$\nabla^2 f(x) \leq 0 \quad \text{for all } x \in \mathbb{R}^n$$

**Every function** $f : \mathbb{R}^n \to \mathbb{R}$ **is either**

- Convex
- Concave, or
- Nether convex nor concave

**Example** Characterize the following functions as convex, concave, or neither convex nor concave

(a) $f(x) = x_1 x_2$

(b) $f(x) = x_1^3 - x_2^4$
(a) The Hessian of $f(x_1, x_2) = x_1 x_2$ is

$$\nabla^2 f(x_1, x_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

The Hessian does not depend on $x$ in this case.

The eigenvalues of the Hessian are $\lambda_1 = 1$ and $\lambda_2 = -1$. The Hessian is undetermined in sign. This function is neither convex nor concave.
(b) The Hessian of \( f(x_1, x_2) = x_1^3 + x_2^4 \) is

\[
\nabla^2 f(x_1, x_2) = \begin{bmatrix}
6x_1 & 0 \\
0 & 12x_2^2
\end{bmatrix}
\]

The Hessian depends on \( x \).

The eigenvalues are also function of \( x \), and are given by

\[
\lambda_1(x) = 6x_1, \quad \lambda_2(x) = 12x_2^2
\]

The Hessian is positive semidefinite for \( x_1 \geq 0 \), and undetermined in sign if \( x_1 < 0 \).

This function is neither convex nor concave over \( \mathbb{R}^n \).