Lecture 10 - Unconstrained Optimization

ling

Section 1

Motivation

Motivation

Imagine you want to find the extrema of a function, what can you do?

$$\max_{\mathbf{x}} f(\mathbf{x}) = x_1^2 + x_2^2 - x_1 x_2$$

- Silliest way: compare every possible value
- Case 1:

$$\mathbf{x} \in \left[\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right]$$

- Case 2: $\mathbf{x} \in \mathbb{Z}^2$
- Case 3: $\mathbf{x} \in R^2$

Section 2

General steps for finding extrema of functions

General steps for finding extrema of functions

1. Consider boundary
    - Close and Bounded sets
    - Limits
2. Consider interior point
    - First Order Necessary condition (FOC)
    - Second Order Necessary and Sufficient condition (SOC)
3. Consider non-differentiable points
    - Check if limit of Newton's quotient exist

$$\lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}$$

    - Tips: check (a) piecewise function, (b) common un-differentiable functions such as $|x|$

Note

This procedure works for all optimization problems. However, modifications are required when solving (1) single variable (2) multi-variable (3) unconstrained (4) constrained problems.

Section 3

Consider boundary

Consider boundary

- If boundary is closed and bounded, then simply check the points at boundary ($\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$)
- unbounded single variable:
  - Check limits $\lim_{x \to \pm\infty} f(x)$
- unbounded multi-variable:
  - Much more complicated as we need to consider multi-dimension borders.
  - If interested, search up on Coercive functions (a special class of function that ensure the existence of global minimizer)

Section 4

Consider interior points

Consider interior points

- Interior points are points falls within the domain S

Consider interior point: FOC

- Single variable:

$$f'(x) = 0$$

- Multi-variable:

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} = \mathbf{0}$$

  - Requires solving simultaneous equation

Note on $P \Rightarrow Q$

Why is it called necessary ($\Rightarrow$) and not sufficient condition ($\Leftarrow$)?

- If a point is a minima/maxima (P), then the point has $\nabla f(\mathbf{x}) = 0$ (Q)
- If a point has $\nabla f(\mathbf{x}) = 0$, it could be either (1) minima (2) maxima (3) neither ($Q \not\Rightarrow P$)

Consider interior point: SOC

- Single variable:
  - local min: $f''(x^*) > 0$
  - local max: $f''(x^*) < 0$
  - no result: $f''(x^*) = 0$
    - Inflection point require further changing sign test of $f''(x)$
    - An inflection point might not be a stationary point (i.e. $f'(x) \neq 0$)
    - If inflection point is also a stationary point, then it is a saddle point
- Multi-variable:
  - local min: $H_f(\mathbf{x}^*) \succ 0$ (positive definite)
  - local max: $H_f(\mathbf{x}^*) \prec 0$ (negative definite)
  - saddle point: indefinite $H_f(\mathbf{x}^*)$ (neither positive nor negative definite)
    - saddle point: neither local minimizer nor a local maximizer

Principal minor: technique to find definiteness of a matrix

Principal minors

The kth principal minor Δ_k of $\mathbf{A}_{n\times n}$ is the determinant of the kth principal submatrix of $\mathbf{A}$, i.e.

$$\Delta_k = det \begin{bmatrix} a_{11} \cdots a_{1k} \\ \vdots \ddots \vdots \\ a_{k1} \cdots a_{kk} \end{bmatrix}$$

(Leading Principal minor test) Suppose $\mathbf{A} \in \mathbf{R}^{n\times n}$ is a symmetric matrix

- $\mathbf{A}$ is positive definite if and only if $\Delta_k > 0$
    - principal minor is always positive
- $\mathbf{A}$ is negative definite if and only if $(-1)^k \Delta_k > 0$
    - principal minor is alternating in sign, with $\Delta_1 < 0$
- Note: no result if $\Delta_k \geq 0$

EC2104 technique

Form $A = f''_{11}(x_0, y_0), B = f''_{12}(x_0, y_0), C = f''_{22}(x_0, y_0)$

- strict local max: $A < 0, AC - B^2 > 0$
- strict local max: $A > 0, AC - B^2 > 0$
- saddle point: $AC - B^2 < 0$
- no result: $AC - B^2 = 0$

Observe:

- Hessian matrix is

$$H_f(\mathbf{x}) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$

- Principal minors are
  - $\Delta_1 = A$
  - $\Delta_2 = AC - B^2$

Section 5

Comment on finding global optimizer

Comment on finding global optimizer

A few cases you can (theoretically) guarantee to find a global optimizer

1. Closed and Bounded problem
2. Convex programming

Beyond these cases, you have to compare all critical points to determine the global optimizer (which is a very challenging task)

Special case: convex function and convex set

- Only if (1) function is convex (2) domain D is convex
  - convex function: $H_f(\mathbf{x}) \succeq 0 \ \forall \mathbf{x} \in D$
  - convex set: for any two points $\mathbf{x}, \mathbf{y}$ in D, the line segment joining $\mathbf{x}, \mathbf{y}$ also lies in D ($\mathbf{x}, \mathbf{y} \in D \Rightarrow \lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \in D \ \forall \lambda \in [0, 1]$)
- $\mathbf{x}^*$ is a local minimizer $\Rightarrow$ $\mathbf{x}^*$ is a global minimizer

Technique in EC2104

$A = f''_{11}(x, y), B = f''_{12}(x, y), C = f''_{22}(x, y)$

- global maximum: $A \leq 0, C \leq 0, AC - B^2 \geq 0$
- global minimum: $A \geq 0, C \geq 0, AC - B^2 \geq 0$

Observe:

- Hessian matrix is
$$H_f(\mathbf{x}) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}$$
- Principal minors are
  - $\Delta_1 = A$
  - $\Delta_2 = AC - B^2$

Difference between global and local min/max?

- Must hold true for all $\mathbf{x}$, not just the minimizer $\mathbf{x}^*$
- Allow for some "slack" (include equal to)

Section 6

Comparative Statics

Comparative Statics

Consider

$$\max_x f(x) = e^x r - x$$

$$x^* = \log(r)$$

- What happens if I change hyper-parameter at the optimal solution?

$$\frac{\partial f(x^*, r)}{\partial r}$$

- Envelop Theorem: a shortcut to investigate comparative statics
  1. Substitution: solve $\frac{\partial}{\partial r} f^*(r) = \frac{\partial}{\partial r} e^{\log(r)} r - \log(r)$
  2. Envelop Theorem: solve $\frac{\partial}{\partial r} f(x^*(r), r) = \frac{\partial}{\partial r} e^{x^*} r - x^*$

Section 7

Increasing Transformation

Increasing Transformation

If a transformation is positive monotonic transformation, or strictly increasing transformation, then the optimisation result has

1. Same optimal point x^*
2. Different optimal functional value $f(x^*)$

Useful for:

1. Simplify original optimization problem
2. Simplify the transformed optimization problem

Consider:

$$\max_p \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$