Basic Concepts in Optimization.
Part II: Continuous and Unconstrained Optimization

Benoît Chachuat <benoit@mcmaster.ca>

McMaster University
Department of Chemical Engineering

ChE 4G03: Optimization in Chemical Engineering

Outline

- Important concepts for the optimization of systems with continuous variables and nonlinear equations
- The topic is limited to unconstrained problems, so the emphasis will be on the objective function

"But no problem is unconstrained! So why is that needed?"

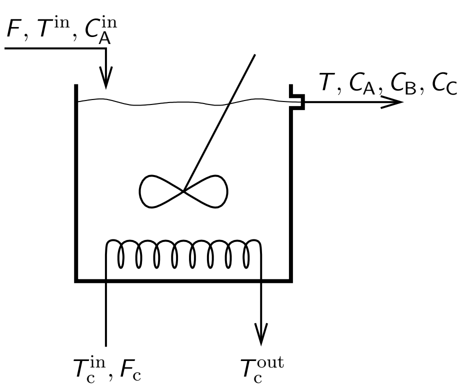
- Sometimes the solution does not lie on the constraints — unconstrained optima
- Unconstrained optimization is used in solution methods for constrained problems

Contents:

For additional details, see Rardin (1998), Chapter 13.3-13.4

Building Experience in Nonlinear Optimization

Class Exercise: Is this system linear or nonlinear?



- This is an isothermal CFSTR with the series reaction:

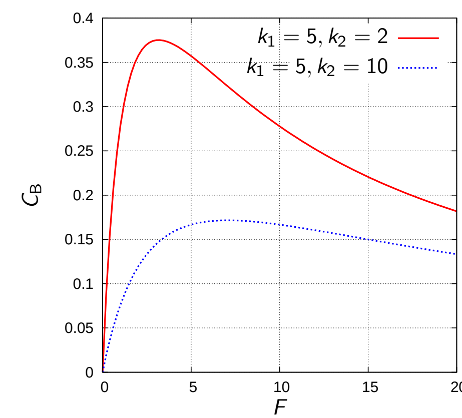
$$A \xrightarrow{k_1} B \xrightarrow{k_2} C$$

- The goal is to maximize C_B in the effluent at steady state
- Only the flow rate of feed F can be adjusted
- The reaction kinetics are first order

Building Experience in Nonlinear Optimization

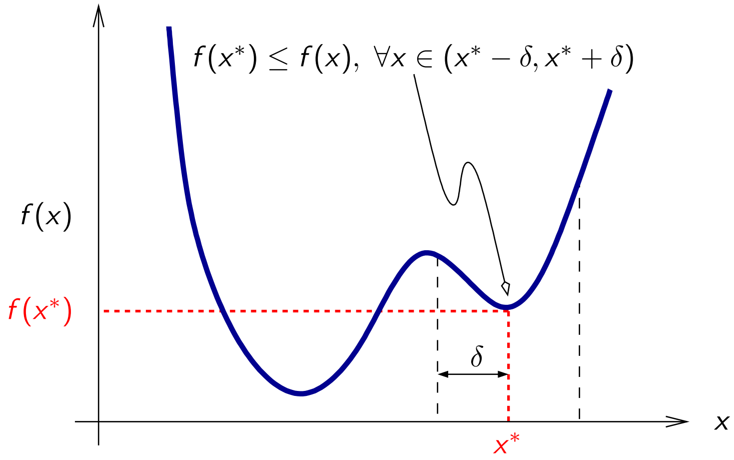
- Mass-balance equations:
- Expressions of C_A and C_B:
- Optimization problem:

Illustration for:
$V = 1\ {\rm m}^3$, $C_A^{\rm in} = 1\ {\rm mol\ s}^{-1}$, $k_1 = 5\ {\rm s}^{-1}$, $k_2 = 2\ {\rm s}^{-1}$



Local Optima in the Unconstrained, Single Variable Case

- A point $x^* \in \mathbb{R}$ is a local optimum of a real-value function f if sufficiently small neighborhoods surrounding it contain no points that are superior in objective value:



Algebraic Characterization of Local Optima

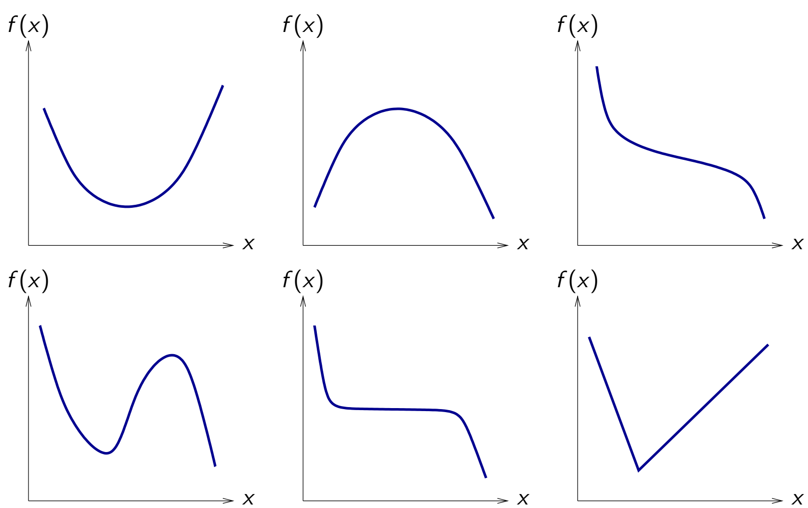
- The previous characterization is impractical to single out local optima (geometric characterization)
- Can we come up with a more practical algebraic characterization?
  - The derivative of f at x^* is defined as:
  $$f'(x^*) \triangleq \lim_{\epsilon \to 0} \frac{f(x^* + \epsilon) - f(x^*)}{\epsilon}$$
  - If this derivative exists and is continuous, then for a local minimum
  $$\frac{f(x^* + \epsilon) - f(x^*)}{\epsilon} \geq 0,\ \forall \epsilon \in (0, \delta) \quad \Rightarrow \quad f'(x^*) \geq 0 \quad \textbf{Why?}$$
  $$\frac{f(x^* + \epsilon) - f(x^*)}{\epsilon} \leq 0,\ \forall \epsilon \in (-\delta, 0) \quad \Rightarrow \quad f'(x^*) \leq 0 \quad \textbf{Why?}$$
  - Therefore, we obtain:

1st-Order Necessary Condition of Optimality

$$x^* \text{ local optimum} \quad \Rightarrow \quad f'(x^*) = 0,\ x^* \text{ stationary point}$$

Applying the 1st-Order Necessary Condition of Optimality

Class Exercise: Which points are stationary points? Do such points correspond to local optima? (local minimum or maximum)



Math Refresher: Taylor Series Approximations

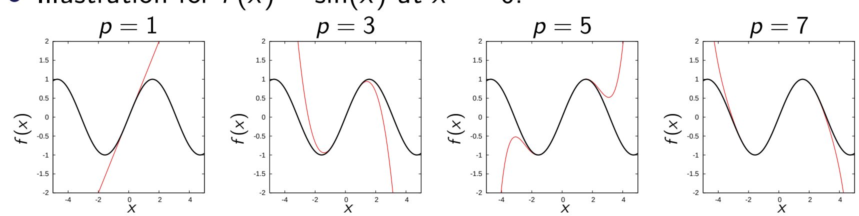
- Taylor's approximation represents the impact of a change δ for current point x° as:

$$f(x^\circ + h) \approx f(x^\circ) + hf'(x^\circ) + \frac{h^2}{2} f''(x^\circ) + \cdots + \frac{h^p}{p!} f^{(p)}(x^\circ) = \sum_{k=0}^{p} \frac{h^k}{k!} f^{(k)}(x^\circ)$$

- The approximation error can be shown to be:

$$f(x^\circ + \delta) - \sum_{k=0}^{p} \frac{h^k}{k!} f^{(k)}(x^\circ) = \frac{h^{p+1}}{(p+1)!} f^{(p+1)}(\xi), \quad \text{for some } x^\circ \leq \xi \leq x^\circ + h$$

- Illustration for $f(x) \triangleq \sin(x)$ at $x^\circ = 0$:



Algebraic Characterization of Local Optima (cont'd)

- Can we come up with a better algebraic characterization?
  - The Taylor's approximation up to $p = 1$ of f at x^* reads:

$$f(x^*+\epsilon) = f(x^*) + \underbrace{\epsilon f'(x^*)}_{=0} + \frac{\epsilon^2}{2} f''(\xi), \quad \text{for some } x^* \leq \xi \leq x^* + \epsilon$$

  - If the second derivative exists and is continuous, then for a local minimum

$$f''(\xi) = \frac{2}{\epsilon^2}\left[f(x^*+\epsilon) - f(x^*)\right] \geq 0 \quad \textbf{Why?}$$

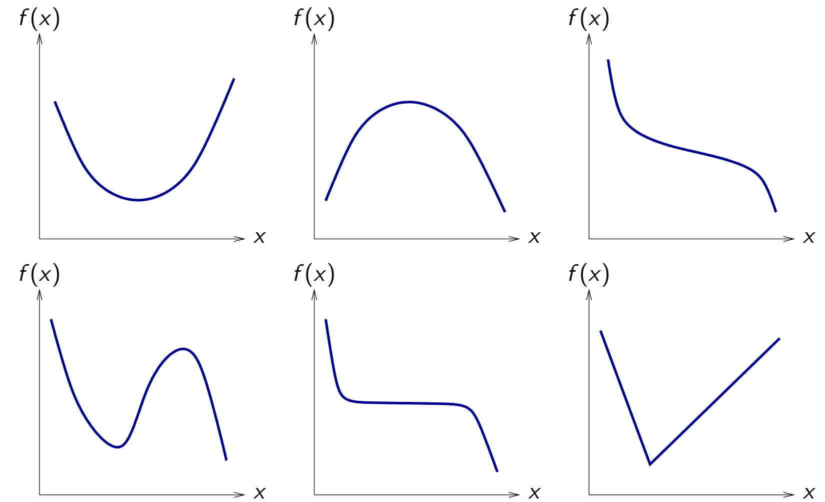
  - Finally, by taking the limit as $\epsilon \to 0$ (i.e., $f''(\xi) \to f''(x^*)$), we obtain:

2nd-Order Necessary Condition of Optimality

$$x^* \text{ local minimum} \quad \Rightarrow \quad f''(x^*) \geq 0$$
$$x^* \text{ local maximum} \quad \Rightarrow \quad f''(x^*) \leq 0$$

Applying the 2nd-Order Necessary Condition of Optimality

Class Exercise: Which points satisfy the first- and second-order necessary conditions? Do such points correspond to local maxima? to local minima?



Applying the 2nd-Order Necessary Condition of Optimality

Class Exercise: Single out candidate optima for the following unconstrained problems, and determine whether such points are candidate local minima or local maxima

1. $\min_x \; f(x) \triangleq 3 + 2x + 5x^2$

2. $\min_x \; f(x) \triangleq 3 + x^3$

Algebraic Characterization of Local Optima (cont'd)

- The previous conditions allow identification of candidate optima, but not to conclude with certainty
- Can such **sufficient conditions** be derived?
  - Consider the Taylor's approximation up to $p \triangleq 2n-1$ of f at x^*:

$$f(x^*+\epsilon) = f(x^*) + \epsilon f'(x^*) + \cdots + \frac{\epsilon^p}{p!} f^{(p)}(x^*) + \frac{\epsilon^{p+1}}{(p+1)!} f^{(p+1)}(\xi),$$

  for some $x^* \leq \xi \leq x^* + \epsilon$

  - Suppose that $f'(x^*) = \cdots = f^{(p)}(x^*) = 0$ and $f^{(p+1)}(x^*) > 0$. Then,

$$f(x^*+\epsilon) = f(x^*) + \frac{\epsilon^{p+1}}{(p+1)!} f^{(p+1)}(\xi)$$

  and

$$\exists \delta > 0 \text{ such that } f^{(p+1)}(\xi) > 0, \forall x^* \leq \xi \leq x^* + \delta \qquad \textbf{Why?}$$

Algebraic Characterization of Local Optima (cont'd)

- Therefore, for one such δ, we have:

$$f(x^* + \epsilon) > f(x^*), \quad \forall \epsilon \in (0, \delta] \qquad \textbf{Why?}$$

i.e., x^* is a strict local minimum!

2nd-Order Sufficient Conditions of Optimality

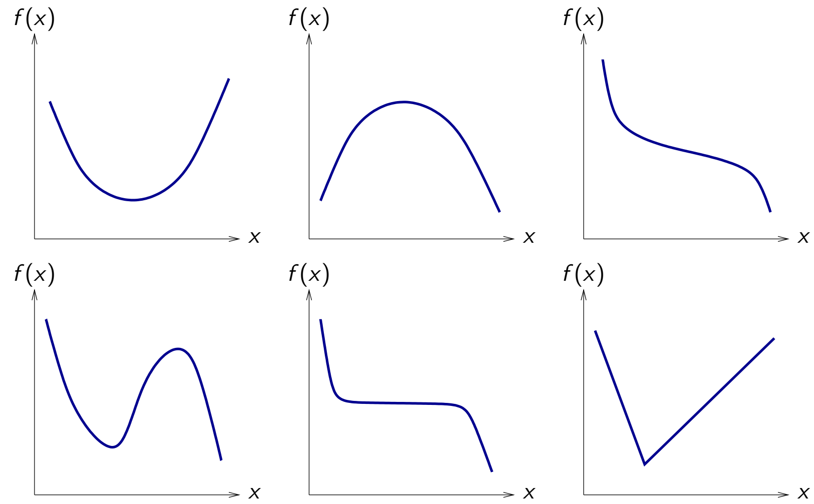
$$\left.\begin{array}{l} f'(x^*) = \cdots = f^{(2n-1)}(x^*) = 0 \\ f^{(2n)}(x^*) > 0 \end{array}\right\} \Rightarrow x^* \text{ strict local minimum}$$

$$\left.\begin{array}{l} f'(x^*) = \cdots = f^{(2n-1)}(x^*) = 0 \\ f^{(2n)}(x^*) < 0 \end{array}\right\} \Rightarrow x^* \text{ strict local maximum}$$

$$\left.\begin{array}{l} f'(x^*) = \cdots = f^{(2n)}(x^*) = 0 \\ f^{(2n+1)}(x^*) \neq 0 \end{array}\right\} \Rightarrow x^* \text{ saddle point}$$

Applying the 2nd-Order Sufficient Conditions of Optimality

Class Exercise: Which points satisfy the second-order sufficient conditions for a local strict minimum? for a strict local maximum? for a saddle point?



Applying the 2nd-Order Sufficient Conditions of Optimality

Class Exercise: Check whether the candidate optima determined previously correspond to strict local minima? To strict local maxima?

1. $\min_x \ f(x) \triangleq 3 + 2x + 5x^2$
2. $\min_x \ f(x) \triangleq 3 + x^3$
3. How about $\min_x \ f(x) \triangleq 1 + x^6$

Math Refresher: Multivariable Functions

Given a multivariable function $f : \mathbb{R}^n \to \mathbb{R}$ and a point $\mathbf{x}^\circ \in \mathbb{R}^n$,

- the gradient, $\nabla f(\mathbf{x}^\circ)$, is the vector of first partial derivatives at $\mathbf{x}^\circ$
- the Hessian, $\mathbf{H}(\mathbf{x}^\circ)$, the matrix of second partial derivatives

$$\nabla f(\mathbf{x}^\circ) \triangleq \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}^\circ) \\ \frac{\partial f}{\partial x_2}(\mathbf{x}^\circ) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}^\circ) \end{pmatrix} \qquad \mathbf{H}(\mathbf{x}^\circ) \triangleq \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^\circ) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^\circ) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}^\circ) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}^\circ) & \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}^\circ) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}^\circ) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}^\circ) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}^\circ) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\mathbf{x}^\circ) \end{pmatrix}$$

- the gradient describes the rate of change or slope of f with small increments in variables values $x_1, \ldots, x_n$ around $\mathbf{x}^\circ$
- the Hessian tells us about the change in slope or curvature of f in the neighborhood of $\mathbf{x}^\circ$
- the Hessian matrix is symmetric if f is twice continuously differentiable

Math Refresher: Multivariable Functions

Consider a $\mathcal{C}^2$ function $f : \mathbb{R}^n \to \mathbb{R}$, a point $\mathbf{x}^\circ \in \mathbb{R}^n$, a direction $\boldsymbol{\delta} \in \mathbb{R}^n$, and a step $\eta \in \mathbb{R}$,

- A 2nd-order (or quadratic) Taylor series approximation is:

$$f(\mathbf{x}^\circ + \alpha\boldsymbol{\delta}) \approx f(\mathbf{x}^\circ) + \alpha \boldsymbol{\nabla} f(\mathbf{x}^\circ)^\mathsf{T}\boldsymbol{\delta} + \frac{\alpha^2}{2}\boldsymbol{\delta}^\mathsf{T}\mathbf{H}(\mathbf{x}^\circ)\boldsymbol{\delta}$$
$$\approx f(\mathbf{x}^\circ) + \alpha\sum_{i=1}^{n}\delta_i \frac{\partial f}{\partial x_i}(\mathbf{x}^\circ) + \frac{\alpha^2}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\delta_i\delta_j\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}^\circ)$$

- Multivariable Taylor's theorem:

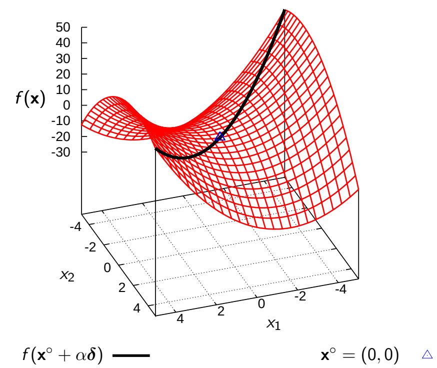
$$f(\mathbf{x}^\circ + \alpha\boldsymbol{\delta}) = f(\mathbf{x}^\circ) + \alpha \boldsymbol{\nabla} f(\mathbf{x}^\circ)^\mathsf{T}\boldsymbol{\delta} + \frac{\alpha^2}{2}\boldsymbol{\delta}^\mathsf{T}\mathbf{H}(\mathbf{x}^\circ + \xi\boldsymbol{\delta})\boldsymbol{\delta}$$

for some $0 \leq \xi \leq \alpha$

Math Refresher: Multivariable Functions

Class Exercise: Formulate a 2nd-order Taylor approximation of the following function at $\mathbf{x}^\circ = (0, 0)$ in the direction $\boldsymbol{\delta} = (1, 1)$,

$$f(\mathbf{x}) \triangleq \frac{3}{2}x_1^2 - x_2^2 + x_1 x_2$$



Algebraic Characterization of Local Optima

- Can we generalize the 1st-order necessary condition of optimality to multiple variable problems?
  - If the function f is continuously differentiable, then for a local minimum $\mathbf{x}^*$ and any direction $\boldsymbol{\delta}$,

$$\frac{f(\mathbf{x}^* + \epsilon\boldsymbol{\delta}) - f(\mathbf{x}^*)}{\epsilon} \geq 0, \ \forall \epsilon \in (0, \alpha) \quad \Rightarrow \quad \frac{d}{d\epsilon}[f(\mathbf{x}^* + \epsilon\boldsymbol{\delta})] \geq 0$$
$$\frac{f(\mathbf{x}^* + \epsilon\boldsymbol{\delta}) - f(\mathbf{x}^*)}{\epsilon} \leq 0, \ \forall \epsilon \in (-\alpha, 0) \quad \Rightarrow \quad \frac{d}{d\epsilon}[f(\mathbf{x}^* + \epsilon\boldsymbol{\delta})] \leq 0$$

  - Therefore, we have:

$$\frac{d}{d\epsilon}[f(\mathbf{x}^* + \epsilon\boldsymbol{\delta})] = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial f}{\partial x_n}(\mathbf{x}^*) \end{pmatrix}\boldsymbol{\delta} = 0, \quad \forall \boldsymbol{\delta}$$

  (in particular, choose $\delta_i = 1$ and $\delta_k = 0$, $k \neq i$)

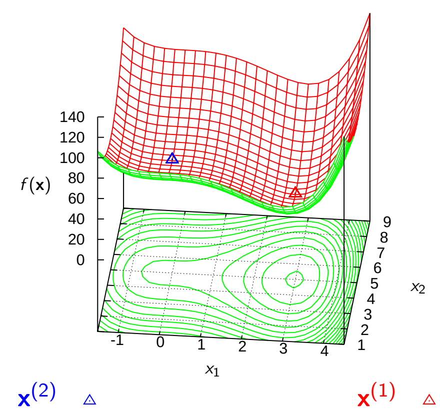
1st-Order Necessary Condition of Optimality

$$\mathbf{x}^* \text{ local optimum} \quad \Rightarrow \quad \boldsymbol{\nabla} f(\mathbf{x}^*) = \mathbf{0}, \ \mathbf{x}^* \text{ stationary point}$$

Applying the 1st-Order Necessary Conditions of Optimality

Class Exercise: Single out candidate optima (stationary points) for the following unconstrained optimization problem

$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2$$



Algebraic Characterization of Local Optima (cont'd)

- Can we refine our algebraic characterization of a local optimum $\mathbf{x}^*$?
  - Quadratic Taylor's approximation of f at $\mathbf{x}^*$ in direction $\boldsymbol{\delta}$:

$$f(\mathbf{x}^* + \epsilon\boldsymbol{\delta}) = f(\mathbf{x}^*) + \underbrace{\epsilon\nabla f(\mathbf{x}^*)^{\mathsf{T}}\boldsymbol{\delta}}_{=0} + \frac{\epsilon^2}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathbf{H}(\mathbf{x}^* + \xi\boldsymbol{\delta})\boldsymbol{\delta},$$

  for some $0 \leq \xi \leq \epsilon$
  - If the function f is $\mathcal{C}^2$, then for a local minimum

$$\boldsymbol{\delta}^{\mathsf{T}}\mathbf{H}(\mathbf{x}^* + \xi\boldsymbol{\delta})\boldsymbol{\delta} = \frac{2}{\epsilon^2}\left[f(\mathbf{x}^* + \epsilon\boldsymbol{\delta}) - f(\mathbf{x}^*)\right] \geq 0 \quad \textbf{Why?}$$

  - By taking the limit as $\epsilon \to 0$, we obtain the 2nd-order condition:

$$\boldsymbol{\delta}^{\mathsf{T}}\mathbf{H}(\mathbf{x}^*)\boldsymbol{\delta} \geq 0, \quad \forall\boldsymbol{\delta} \in \mathbb{R}^n$$

- **Can the foregoing condition be reformulated into more usable form?**

Positive and Negative (Semi) Definite Matrices

Positive (Semi) Definiteness

An $(n \times n)$ matrix $\mathbf{A}$ is said to be:
- **positive definite** if $\boldsymbol{\delta}^{\mathsf{T}}\mathbf{A}\boldsymbol{\delta} > 0, \ \forall\boldsymbol{\delta} \neq \mathbf{0}$
- **positive semi-definite** if $\boldsymbol{\delta}^{\mathsf{T}}\mathbf{A}\boldsymbol{\delta} \geq 0, \ \forall\boldsymbol{\delta}$

(**negative (semi) definite** matrices defined likewise)

For which class of matrices are the previous tests easy to check?

- Consider a diagonal matrix $\boldsymbol{\Lambda} \triangleq \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$. Then,

$$\begin{array}{rcl} \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Lambda}\boldsymbol{\delta} > 0, \ \forall\boldsymbol{\delta} \neq \mathbf{0} & \Leftrightarrow & \lambda_1, \ldots, \lambda_n > 0 \\ \boldsymbol{\delta}^{\mathsf{T}}\boldsymbol{\Lambda}\boldsymbol{\delta} \geq 0, \ \forall\boldsymbol{\delta} & \Leftrightarrow & \lambda_1, \ldots, \lambda_n \geq 0 \end{array}$$

Positive and Negative (Semi) Definite Matrices

Can I transform a matrix into diagonal form?

- Yes! Perform an **eigenvalue/eigenvector decomposition**...

Refresher: Eigenvalue Decomposition

Eigenvalues of a $(n \times n)$ square matrix $\mathbf{A}$ are the solutions λ of the following equation:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

Real eigenvalues are guaranteed to exist for a **real symmetric** matrix

Theorem: Characterization of Positive (Semi) Definiteness

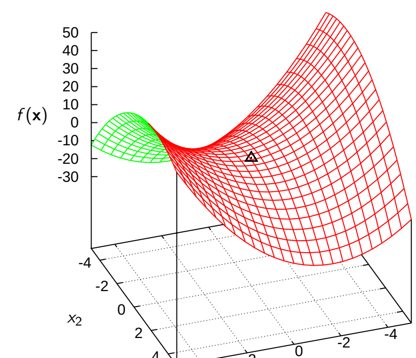
An $(n \times n)$ matrix $\mathbf{A}$ is:
- **positive definite** if and only if all its eigenvalues λ are > 0
- **positive semi-definite** if and only if all its eigenvalues λ are ≥ 0

Algebraic Characterization of Local Optima (cont'd)

2nd-Order Necessary Condition of Optimality

$$\begin{array}{rcl} \mathbf{x}^* \text{ local minimum} & \Rightarrow & \mathbf{H}(\mathbf{x}^*) \succeq 0 \text{ (positive semi-definite)} \\ \mathbf{x}^* \text{ local maximum} & \Rightarrow & \mathbf{H}(\mathbf{x}^*) \preceq 0 \text{ (negative semi-definite)} \\ \mathbf{x}^* \text{ saddle point} & \Rightarrow & \mathbf{H}(\mathbf{x}^*) \text{ indefinite} \end{array}$$

Class Exercise: Characterize $\mathbf{x}^\circ = (0,0)$ for $f(\mathbf{x}) \triangleq \frac{3}{2}x_1^2 - x_2^2 + x_1x_2$

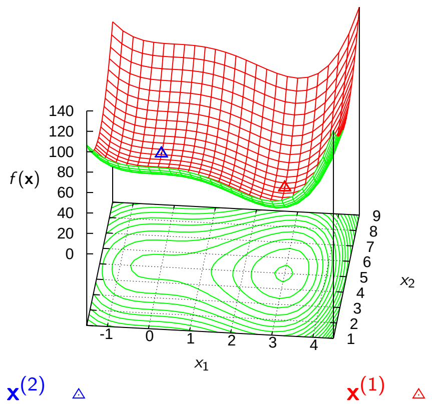


Applying the 2nd-Order Necessary Conditions of Optimality

Class Exercise: Determine whether each candidate optimum identified previously for the unconstrained optimization problem

$$\min_{\mathbf{x}} f(\mathbf{x}) \stackrel{\Delta}{=} 40 + x_1^3(x_1 - 4) + 3(x_2 - 5)^2$$

could be a local minimum? a local maximum? a saddle point?



Algebraic Characterization of Local Optima (cont'd)

- Can **sufficient conditions** be derived as well?
  - Consider the quadratic Taylor's approximation of f at $\mathbf{x}^*$:

$$f(\mathbf{x}^* + \epsilon\boldsymbol{\delta}) = f(\mathbf{x}^*) + \epsilon \nabla f(\mathbf{x}^*)^{\mathsf{T}}\boldsymbol{\delta} + \frac{\epsilon^2}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathbf{H}(\mathbf{x}^* + \xi\boldsymbol{\delta})\boldsymbol{\delta},$$

  for some $0 \leq \xi \leq \epsilon$

  - Suppose that f is $\mathcal{C}^2$, $\nabla f(\mathbf{x}^*) = \mathbf{0}$, and $\mathbf{H}(\mathbf{x}^*) \succ 0$ (definite positive). Then,

$$f(\mathbf{x}^* + \epsilon\boldsymbol{\delta}) = f(\mathbf{x}^*) + \frac{\epsilon^2}{2}\boldsymbol{\delta}^{\mathsf{T}}\mathbf{H}(\mathbf{x}^* + \xi\boldsymbol{\delta})\boldsymbol{\delta}$$

  and

$$\exists \alpha > 0 \text{ such that } \boldsymbol{\delta}^{\mathsf{T}}\mathbf{H}(\mathbf{x}^* + \xi\boldsymbol{\delta})\boldsymbol{\delta} > 0, \forall \xi \in [0, \alpha] \qquad \textbf{Why?}$$

Algebraic Characterization of Local Optima (cont'd)

- Therefore, for one such α, we have:

$$f(\mathbf{x}^* + \epsilon\boldsymbol{\delta}) > f(\mathbf{x}^*), \quad \forall \epsilon \in (0, \alpha] \qquad \textbf{Why?}$$

- Since this is true for every direction $\boldsymbol{\delta}$, x^* is a strict local minimum!

2nd-Order Sufficient Conditions of Optimality

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \text{ and } \mathbf{H}(\mathbf{x}^*) \succ 0 \quad \Rightarrow \quad x^* \text{ strict local minimum}$$

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \text{ and } \mathbf{H}(\mathbf{x}^*) \prec 0 \quad \Rightarrow \quad x^* \text{ strict local maximum}$$

No conclusion can be drawn in case $\mathbf{H}(\mathbf{x}^*)$ is indefinite!

Class Exercise: Apply the 2nd-order sufficient conditions to the points $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ in the previous exercise and conclude.

Refresher: Convexity

(Strictly) Convex Functions

A function $f : S \to \mathbb{R}$, defined on a convex set $S \subset \mathbb{R}^n$, is said to be:

- **convex on S** if

$$f(\gamma\mathbf{x}_1 + (1-\gamma)\mathbf{x}_2) \leq \gamma f(\mathbf{x}_1) + (1-\gamma) f(\mathbf{x}_2)$$

  for every $\mathbf{x}_1, \mathbf{x}_2 \in S$ and every step $\gamma \in [0, 1]$

- **strictly convex on S** if

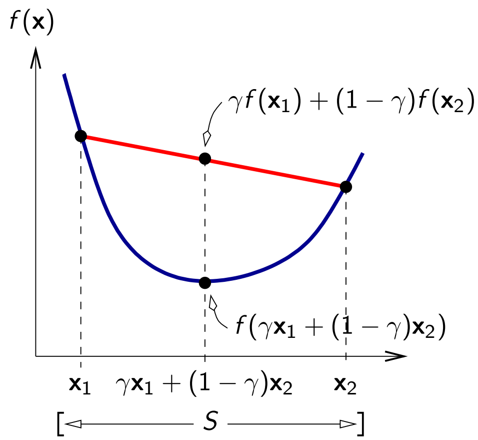
$$f(\gamma\mathbf{x}_1 + (1-\gamma)\mathbf{x}_2) < \gamma f(\mathbf{x}_1) + (1-\gamma) f(\mathbf{x}_2)$$

  for every $\mathbf{x}_1, \mathbf{x}_2 \in S$ and every step $\gamma \in (0, 1)$

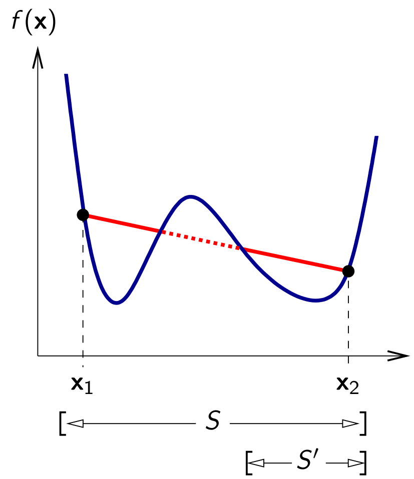
- **[strictly] concave on S** if $(-f)$ is [strictly] convex on S

Refresher: Convexity (cont'd)

Case of a convex function on the convex set S



Case of a nonconvex function on S, yet convex on the convex set S'



Convexity and Global Optimality

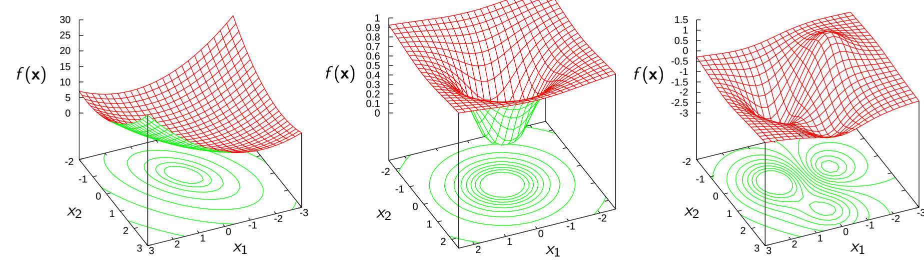
Sufficient Condition for Unconstrained **Global** Optima

Let $f : \mathbb{R}^n \to \mathbb{R}$ be [strictly] convex and let $\mathbf{x}^*$ be a local minimum of the unconstrained nonlinear program,

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

Then, $\mathbf{x}^*$ is also a [strict] global minimum.

Class Exercise: Discuss each problem in the light of the foregoing sufficient condition



How to Detect Convexity?

Gradient and Hessian Tests

A function $f : S \subset \mathbb{R}^n \to \mathbb{R}$ in $\mathcal{C}^2$ is convex on S if, at each $\mathbf{x}^\circ \in S$,

- **Gradient Test:** $f(\mathbf{x}) \geq f(\mathbf{x}^\circ) + \nabla f(\mathbf{x}^\circ)^{\mathsf{T}} (\mathbf{x} - \mathbf{x}^\circ), \quad \forall \mathbf{x} \in S$
- **Hessian Test:** $\mathbf{H}(\mathbf{x}^\circ) \succeq 0$ (positive semi-definite)

Strict convexity is detected by making the inequality signs strict

Positive definiteness of $\mathbf{H}$ again!

But **why?** Is it related to 2nd-order necessary/sufficiency conditions?

Class Exercise: Determine whether or not the following functions are convex:

- $(x_1, x_2) \mapsto x_1^2 + x_1x_2 + 2x_2 + 4$, for $(x_1, x_2) \in \mathbb{R}$
- $(x_1, x_2) \mapsto (x_1 + 1)^4 + x_1x_2 + (x_2 + 1)^4$, for $(x_1, x_2) \geq 0$

The Final Words

- Optimality conditions for unconstrained, single variable optimization
- Optimality conditions for unconstrained, multivariable optimization

$\Rightarrow$ The building blocks of many optimization algorithms and convergence tests

- Convexity and sufficiency for global optimality

$\Rightarrow$ Formulating optimization models that are convex is always desirable