Basic Concepts in Optimization. Part III: Continuous and Constrained Optimization

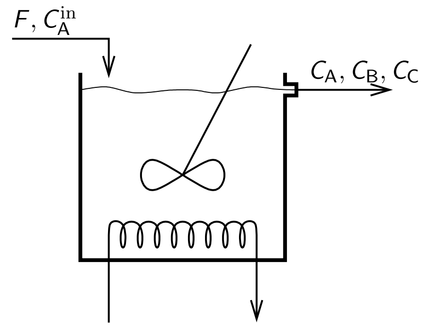
Benoît Chachuat <benoit@mcmaster.ca>

McMaster University
Department of Chemical Engineering

ChE 4G03: Optimization in Chemical Engineering

Building Experience in Constrained Optimization

Class Exercise: Formulate an optimization model for the following problem:



- Isothermal CFSTR with series reaction (1st-order kinetics):

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

- The goal is to maximize C_B in the effluent at steady state
- The flow rate F and inlet concentration $C_A^{\rm in}$ are manipulated, with maximum levels F^U and $C_A^{{\rm in},U}$
- The concentration of A in the effluent should not exceed C_A^U

Building Experience in Constrained Optimization

- Mass-balance equations:

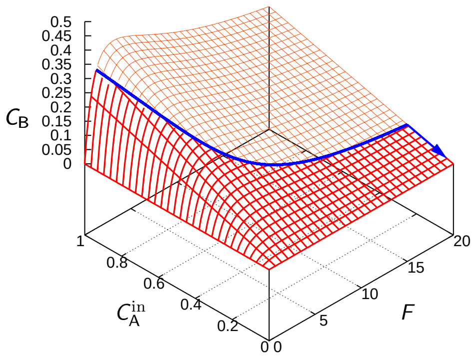
$$0 = F\left(C_A^{\rm in} - C_A\right) - k_1 C_A V$$
$$0 = -FC_B + (k_1 C_A - k_2 C_B)V$$

- Expressions of C_A and C_B:

$$C_A = C_A^{\rm in}\frac{F}{F + k_1 V}$$
$$C_B = C_A^{\rm in}\frac{Fk_1 V}{(F + k_1 V)(F + k_2 V)}$$

- Optimization model:

Illustration for:
$V = 1\ {\rm m}^3$, $F^U = 20\ {\rm m}^3\ {\rm s}^{-1}$,
$C_A^{{\rm in},U} = 1\ {\rm mol\ m}^{-3}$, $C_A^U = 0.2\ {\rm mol\ m}^{-3}$,
$k_1 = 5\ {\rm s}^{-1}$, $k_2 = 2\ {\rm s}^{-1}$



Where is the optimum?

Outline

- Important concepts for the optimization of systems with continuous variables and nonlinear equations
- The topic is now extended to fully constrained problems, so the emphasis will be on the objective function and the constraints

Contents:

For additional details, see Rardin (1998), Chapter 14.3-14.4

Linear Equality Constrained Optimization

- Consider the optimization problem with linear equality constraints

$$\begin{aligned} \text{minimize:} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned}$$

with $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$

- Provided that **A has full rank**, $\text{rank}(\mathbf{A}) = m$, one can *always* partition $\mathbf{x}$ into $\mathbf{x}_\text{d} \in \mathbb{R}^m$ and $\mathbf{x}_\text{i} \in \mathbb{R}^{n-m}$ such that:

$$\mathbf{x}_\text{d} = -\mathbf{A}_\text{d}^{-1}\left[\mathbf{A}_\text{i}\mathbf{x}_\text{i} - \mathbf{b}\right],$$

with $\mathbf{A}_\text{d} \in \mathbb{R}^{m \times m}$ and $\mathbf{A}_\text{i} \in \mathbb{R}^{m \times (n-m)}$ the related partitions of $\mathbf{A}$

- The original optimization problem is thus equivalent to the reduced unconstrained problem

$$\boxed{\text{minimize:} \quad F(\mathbf{x}_\text{i}) \triangleq f\left(-\mathbf{A}_\text{d}^{-1}\left[\mathbf{A}_\text{i}\mathbf{x}_\text{i} - \mathbf{b}\right], \mathbf{x}_\text{i}\right)}$$

Solving Linear Equality Constrained Optimization

Class Exercise: Convert the following constrained optimization to an unconstrained optimization, and solve

$$\begin{aligned} \min_{\mathbf{x}} \ & f(\mathbf{x}) \triangleq 4x_1^2 + 5x_2^2 \\ \text{s.t. } & 2x_1 + 3x_2 = 6 \end{aligned}$$

General Equality Constraints

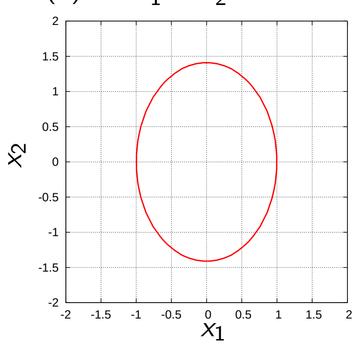
- An nonlinear equality constraint, $h(\mathbf{x}) = 0$, with $\mathbf{x} \in \mathbb{R}^n$, defines a $(n-1)$-dimensional set,

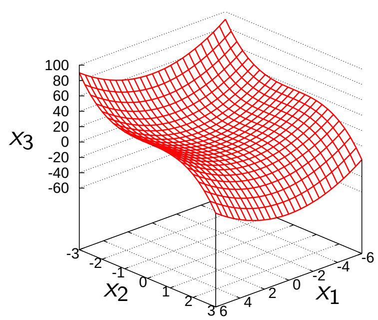
$$S \triangleq \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = 0\} \subsetneq \mathbb{R}^n$$

A 1-d curve in the case $n = 2$, a 2-d surface for $n = 3$, etc.

Examples:

$h(\mathbf{x}) \triangleq 2x_1^2 + x_2^2 - 2 = 0$ $\qquad$ $h(\mathbf{x}) \triangleq x_1^2 - 2x_2^3 - x_3 = 0$



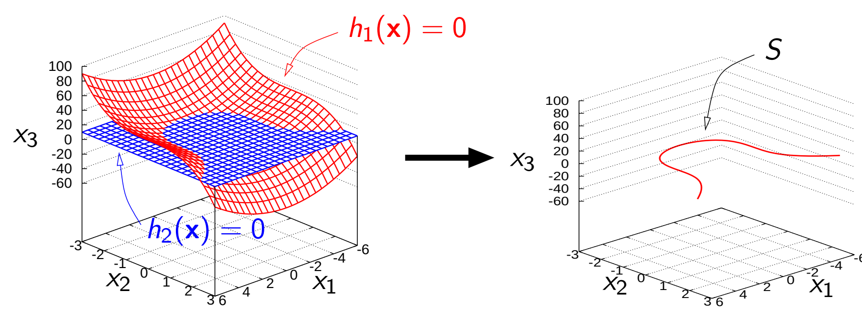


Systems of General Equality Constraints

- When considering $m \geq 1$ equality constraints, $h_1(\mathbf{x}), \ldots, h_m(\mathbf{x})$, their intersection forms a set,

$$S \triangleq \{\mathbf{x} \in \mathbb{R}^n : h_1(\mathbf{x}) = \cdots = h_m(\mathbf{x}) = 0\} \subsetneq \mathbb{R}^n$$

Example: Consider the constraints $h_1(\mathbf{x}) \triangleq x_1^2 - 2x_2^3 - x_3 = 0$ and $h_2(\mathbf{x}) \triangleq x_3 - 10 = 0$

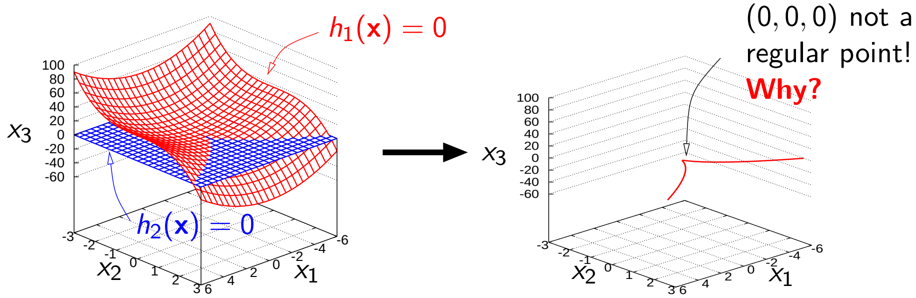


Regularity of General Equality Constraints

- A point $\bar{\mathbf{x}} \in S \stackrel{\Delta}{=} \{\mathbf{x} \in \mathbb{R}^n : h_1(\mathbf{x}) = \cdots = h_m(\mathbf{x}) = 0\}$ is said to be a regular point if the gradient vectors $\nabla h_1(\bar{\mathbf{x}}), \ldots, \nabla h_m(\bar{\mathbf{x}})$ are linearly independent:

$$\text{rank}\left(\nabla h_1(\bar{\mathbf{x}}) \quad \ldots \quad \nabla h_m(\bar{\mathbf{x}}) \right) = m$$

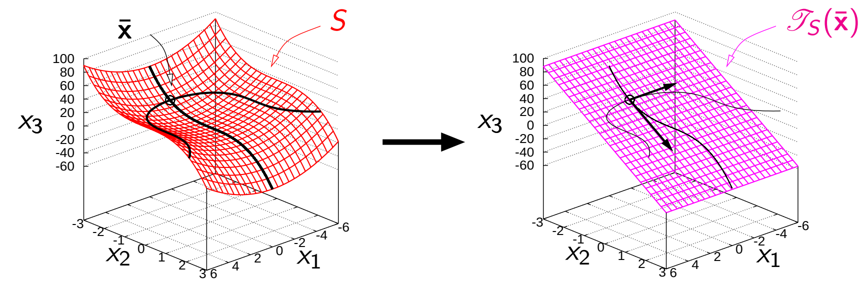
Class Exercise: Consider the constraints $h_1(\mathbf{x}) \stackrel{\Delta}{=} x_1^2 - 2x_2^3 - x_3 = 0$ and $h_2(\mathbf{x}) \stackrel{\Delta}{=} x_3 = 0$ — Check regularity of points $(0,0,0)$ and $(2,1,0)$



Tangent Set to General Equality Constraints

- Consider the collection of *all* smooth curves passing through a point $\bar{\mathbf{x}} \in S$. The tangent set of S at $\bar{\mathbf{x}}$, denoted by $\mathscr{T}_S(\bar{\mathbf{x}})$, is the collection of the vectors tangent to *all* these curves at $\bar{\mathbf{x}}$
- At a regular point $\bar{\mathbf{x}} \in S$, $\mathscr{T}_S(\bar{\mathbf{x}}) \stackrel{\Delta}{=} \{\mathbf{d} \in \mathbb{R}^n : \nabla h_j(\bar{\mathbf{x}})^{\mathsf{T}}\mathbf{d} = 0, \forall j\}$

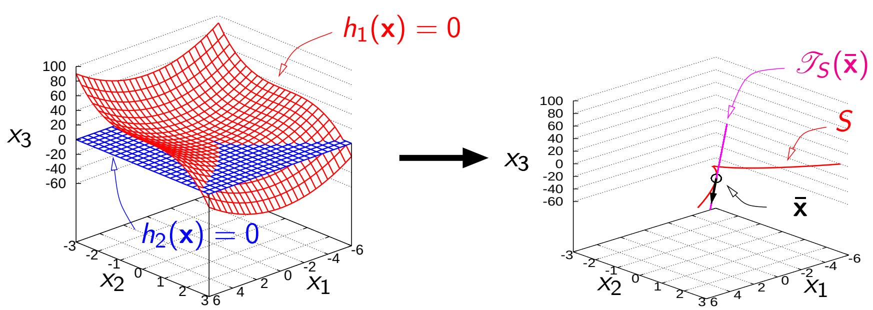
Class Exercise: Determine equation(s) for the tangent set to the constraint $h_1(\mathbf{x}) \stackrel{\Delta}{=} x_1^2 - 2x_2^3 - x_3 = 0$ at $\bar{\mathbf{x}} \stackrel{\Delta}{=} (0, -2, 16)$



Tangent Set to General Equality Constraints

- Consider the collection of *all* smooth curves passing through a point $\bar{\mathbf{x}} \in S$. The tangent set of S at $\bar{\mathbf{x}}$, denoted by $\mathscr{T}_S(\bar{\mathbf{x}})$, is the collection of the vectors tangent to *all* these curves at $\bar{\mathbf{x}}$
- At a regular point $\bar{\mathbf{x}} \in S$, $\mathscr{T}_S(\bar{\mathbf{x}}) \stackrel{\Delta}{=} \{\mathbf{d} \in \mathbb{R}^n : \nabla h_j(\bar{\mathbf{x}})^{\mathsf{T}}\mathbf{d} = 0, \forall j\}$

Class Exercise: Determine equation(s) for the tangent set to the constraints $h_1(\mathbf{x}) \stackrel{\Delta}{=} x_1^2 - 2x_2^3 - x_3 = 0$ and $h_2(\mathbf{x}) \stackrel{\Delta}{=} x_3 = 0$ at $\bar{\mathbf{x}} \stackrel{\Delta}{=} (2,1,0)$



Geometric Optimality Condition

A point $\mathbf{x}^* \in \mathbb{R}^n$ is a local optimum of a real-value function f in the feasible set $S \subset \mathbb{R}^n$, if sufficiently small neighborhoods surrounding it contain no feasible points that are superior in objective value

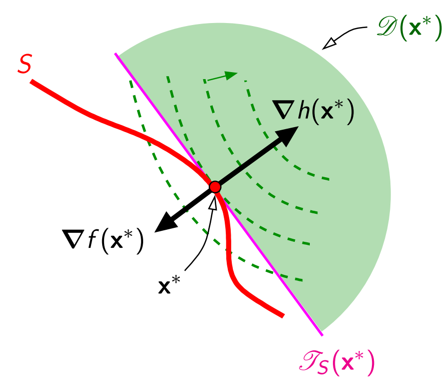
Geometric Optimality Condition

Let f in $\mathcal{C}^1$, and let
$S \stackrel{\Delta}{=} \{\mathbf{x} : h_1(\mathbf{x}) = \cdots = h_m(\mathbf{x}) = 0\}$.
If $\mathbf{x}^*$ is an optimum for f on S, then

$$\mathscr{D}(\mathbf{x}^*) \cap \mathscr{T}_S(\mathbf{x}^*) = \emptyset,$$

with:

- $\mathscr{D}(\mathbf{x}^*)$, set of improving directions
- $\mathscr{T}_S(\mathbf{x}^*)$, tangent set



$$\nexists \mathbf{d} \in \mathbb{R}^n : \nabla f(\mathbf{x}^*)^{\mathsf{T}}\mathbf{d} < 0 \ \underline{\text{and}}\ \nabla h_i(\mathbf{x}^*)^{\mathsf{T}}\mathbf{d} = 0$$

First-Order Necessary Conditions for Optimality

First-Order Necessary Conditions

Let $f, h_1, \ldots, h_m$ in $\mathcal{C}^1$. If $\mathbf{x}^*$ is a (local) optimum for f s.t. $h_i(\mathbf{x}) = 0$, $i = 1, \ldots, m$, **and** regular, there is a unique vector $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that:

$$\begin{cases} h_1(\mathbf{x}^*) = \cdots = h_m(\mathbf{x}^*) = 0 \\ \nabla f(\mathbf{x}^*) = \sum_{i=1}^m \lambda_i^* \nabla h_i(\mathbf{x}^*) \end{cases}$$

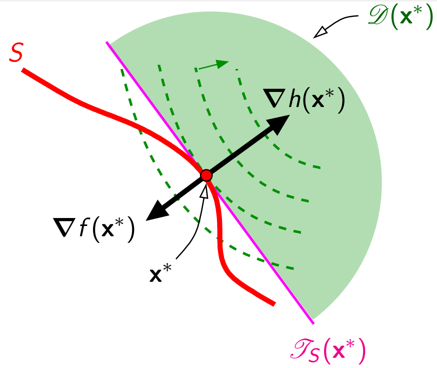
Square system: $(n+m)$ conditions in $(n+m)$ variables $(\mathbf{x}, \boldsymbol{\lambda})$

Lagrange multipliers: $\lambda_i \leftrightarrow h_i$

Lagrangian stationarity:

$$\nabla \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$$

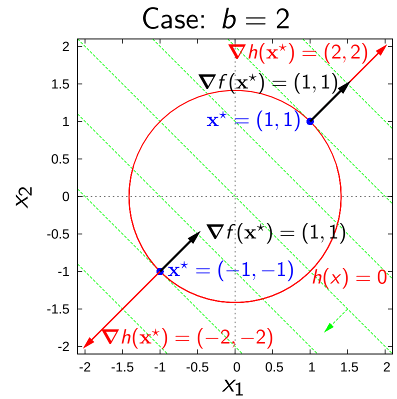
where $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) \triangleq f(\mathbf{x}) - \boldsymbol{\lambda}^\mathsf{T}\mathbf{h}(\mathbf{x})$ is called the Lagrangian



Solving Nonlinear Equality Constrained Optimization

Class Exercise: Use the 1st-order necessary conditions for optimality to single out candidate optimal points (with $b \geq 0$)

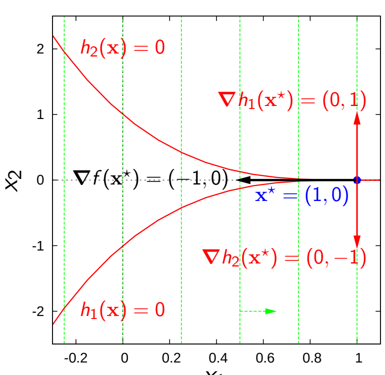
$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq x_1 + x_2$$
$$\text{s.t. } h(\mathbf{x}) \triangleq x_1^2 + x_2^2 = b,$$



Solving Nonlinear Equality Constrained Optimization

Class Exercise: Identify the optimum point (visually) and check the 1st-order necessary conditions for optimality at that point

$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq -x_1$$
$$\text{s.t. } h_1(\mathbf{x}) \triangleq (1 - x_1)^3 + x_2 = 0$$
$$h_2(\mathbf{x}) \triangleq (1 - x_1)^3 - x_2 = 0$$



What's going on here?

Interpretation of the Lagrange Multipliers

Optimization Problem:

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{x}) = \mathbf{b}$$

Lagrangian: $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^m \lambda_i [h_i(\mathbf{x}) - b_i]$

$$\text{NCO: } \begin{cases} h_j(\mathbf{x}^*) - b_j = 0, \quad \forall j = 1, \ldots, m \\ \nabla f(\mathbf{x}^*) - \sum_{i=1}^m \lambda_i^* \nabla h_i(\mathbf{x}^*) = \mathbf{0} \end{cases}$$

- A variation in RHS b_j affects the optimal solution $\longrightarrow \mathbf{x}^*(b_j), \boldsymbol{\lambda}^*(b_j)$
- Constraint j must remain satisfied for *every* value of b_j:

$$0 = \frac{\partial [h_j(\mathbf{x}^*(b_j)) - b_j]}{\partial b_j} = \nabla h_j(\mathbf{x}^*) \frac{\partial \mathbf{x}^*}{\partial b_j} - 1$$

- Rate of change in optimal solution value w.r.t. b_j:

$$\frac{\partial f(\mathbf{x}^*(b_j))}{\partial b_j} = \nabla f(\mathbf{x}^*) \frac{\partial \mathbf{x}^*}{\partial b_j} = \lambda_j^* + \underbrace{\left[\nabla f(\mathbf{x}^*) - \sum_{i=1}^m \lambda_i^* \nabla h_i(\mathbf{x}^*)\right]}_{=\mathbf{0} \text{ (NCO)}} \frac{\partial \mathbf{x}^*}{\partial b_j}$$

Interpretation of the Lagrange Multipliers

Optimization Problem:

$$\min_{\mathbf{x}} f(\mathbf{x})$$
$$\text{s.t. } h(\mathbf{x}) = \mathbf{b}$$

Lagrangian: $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i [h_i(\mathbf{x}) - b_i]$

NCO: $\begin{cases} h_j(\mathbf{x}^*) - b_j = 0, \quad \forall j = 1, \ldots, m \\ \nabla f(\mathbf{x}^*) - \sum_{i=1}^{m} \lambda_i^* \nabla h_i(\mathbf{x}^*) = \mathbf{0} \end{cases}$

The optimal Lagrange multiplier, λ_j^*, associated with constraint $h_j(\mathbf{x}) = b_j$ can be interpreted as the rate of change in optimal value for infinitesimal change in RHS b_j

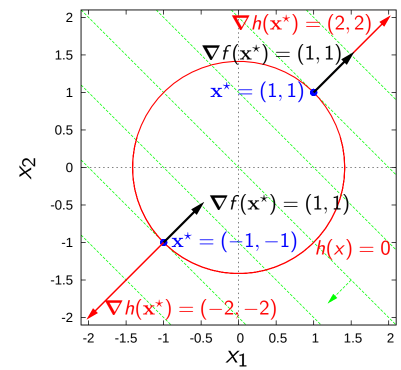
- How does this relate to sensitivity analysis for LP?
- How does sensitivity analysis for NLP differ from LP?

Checking our Interpretation of the Lagrange Multiplier

Class Exercise: Based on the analytical solutions obtained previously, calculate the change in optimal objective for a change in RHS b; compare with the optimal Lagrange multiplier value

$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq x_1 + x_2$$
$$\text{s.t. } x_1^2 + x_2^2 = b,$$

Case: $b = 2$



Inequality Constrained Optimization

- Consider the optimization problem with inequality constraints

$$\begin{aligned} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize:}} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \end{aligned}$$

- In general, multiple inequality constraints define a (possibly unbounded) n-dimensional set,

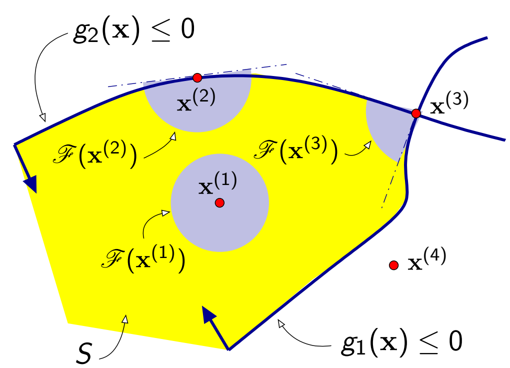
$$S \triangleq \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m\} \subseteq \mathbb{R}^n$$

- At a feasible point $\bar{\mathbf{x}}$, the ith constraints is said to be binding or active if $g_i(\bar{\mathbf{x}}) = 0$; it is said to be inactive if $g_i(\bar{\mathbf{x}}) < 0$
- The set of active constraints at a feasible point $\bar{\mathbf{x}}$ is

$$\mathcal{A}(\bar{\mathbf{x}}) \triangleq \{i : g_i(\bar{\mathbf{x}}) = 0\}$$

Inequality Constrained Optimization (cont'd)

Illustration:



Active Sets:

- $\mathcal{A}(\mathbf{x}^{(1)}) = \emptyset$
- $\mathcal{A}(\mathbf{x}^{(2)}) = \{2\}$
- $\mathcal{A}(\mathbf{x}^{(3)}) = \{1, 2\}$
- $\mathbf{x}^{(4)}$ infeasible!

Characterizing Feasible Directions

- Consider the feasible domain $S \triangleq \{\mathbf{x} : g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m\}$
- The set of feasible directions at a point $\bar{\mathbf{x}} \in S$ is

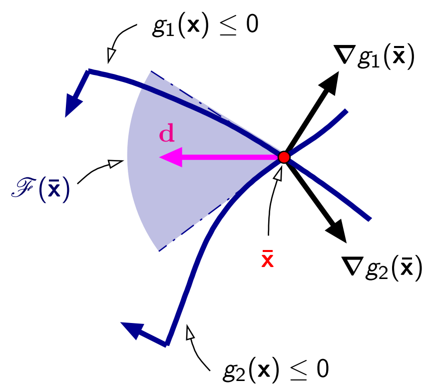
$$\mathscr{F}(\bar{\mathbf{x}}) \triangleq \{\mathbf{d} \neq \mathbf{0} : \exists \varepsilon > 0 \text{ such that } \bar{\mathbf{x}} + \alpha \mathbf{d} \in S, \forall \alpha \in (0, \epsilon)\}$$

Algebraic Characterization

Let $g_1, \ldots, g_m$ in $\mathcal{C}^1$, and let $\bar{\mathbf{x}} \in S$. Any direction $\mathbf{d} \in \mathbb{R}^n$ such that

$$\nabla g_i(\bar{\mathbf{x}})^{\mathsf{T}} \mathbf{d} < 0, \quad \forall i \in \mathcal{A}(\bar{\mathbf{x}})$$

is a feasible direction



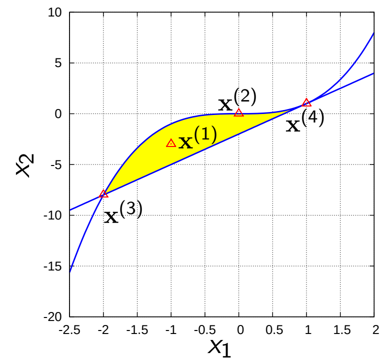
- This condition is sufficient, yet not necessary!
- What if $\mathcal{A}(\bar{\mathbf{x}}) = \emptyset$?
- What if $g_i(\mathbf{x}) \geq 0$?

Regularity of General Inequality Constraints

- A feasible point $\bar{\mathbf{x}} \in S$ is called regular point if the gradient vectors $\nabla g_i(\bar{\mathbf{x}})$ for the all the active constraints are linearly independent:

$$\text{rank}\left(\ \nabla g_i(\bar{\mathbf{x}}), i \in \mathcal{A}(\bar{\mathbf{x}}) \ \right) = \text{card}\left(\mathcal{A}(\bar{\mathbf{x}})\right)$$

Class Exercise: Consider the constraints $g_1(\mathbf{x}) \triangleq -x_1^3 + x_2 \leq 0$ and $g_2(\mathbf{x}) \triangleq 3x_1 - x_2 - 2 \leq 0$



Check regularity at:

- $\mathbf{x}^{(1)} = (-1, -3)$
- $\mathbf{x}^{(2)} = (0, 0)$
- $\mathbf{x}^{(3)} = (-2, -8)$
- $\mathbf{x}^{(4)} = (1, 1)$

Geometric Optimality Condition

A point $\mathbf{x}^* \in \mathbb{R}^n$ is a local optimum of a real-value function f in the feasible set $S \subset \mathbb{R}^n$, if sufficiently small neighborhoods surrounding it contain no feasible points that are superior in objective value

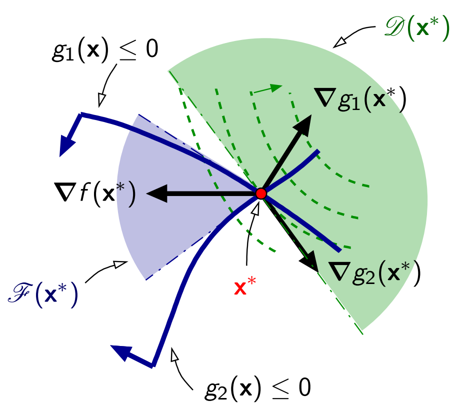
Geometric Optimality Condition

Let f in $\mathcal{C}^1$, and let $S \triangleq \{\mathbf{x} : g_i(\mathbf{x}) \leq 0, i = 1, \ldots, m\}$. If $\mathbf{x}^*$ is an optimum for f on S, then

$$\mathscr{D}(\mathbf{x}^*) \cap \mathscr{F}(\mathbf{x}^*) = \emptyset,$$

with:

- $\mathscr{D}(\mathbf{x}^*)$, set of improving directions
- $\mathscr{F}(\mathbf{x}^*)$, set of feasible directions



$$\nexists \mathbf{d} \in \mathbb{R}^n : \nabla f(\mathbf{x}^*)^{\mathsf{T}} \mathbf{d} < 0 \ \underline{\text{and}} \ \nabla g_i(\mathbf{x}^*)^{\mathsf{T}} \mathbf{d} < 0, \forall i \in \mathcal{A}(\mathbf{x}^*)$$

KKT Points

- Let f and $g_1, \ldots, g_m$ in $\mathcal{C}^1$, and consider the NLP problem

$$\begin{aligned} \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize:}} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m \end{aligned}$$

- A point $(\bar{\mathbf{x}}, \bar{\boldsymbol{\nu}}) \in \mathbb{R}^n \times \mathbb{R}^m$ is called a KKT point if it satisfied

$$\begin{aligned} &\text{Primal Feasibility:} && g_i(\bar{\mathbf{x}}) \leq 0, \quad i = 1, \ldots, m \\ &\text{Dual Feasibility:} && \begin{cases} \nabla f(\bar{\mathbf{x}}) = \sum_{i=1}^m \bar{\nu}_i \nabla g_i(\bar{\mathbf{x}}) \\ \bar{\nu}_i \leq 0, \quad i = 1, \ldots, m \end{cases} \\ &\text{Complementarity Slackness:} && \bar{\nu}_i g_i(\bar{\mathbf{x}}) = 0, \quad i = 1, \ldots, m \end{aligned}$$

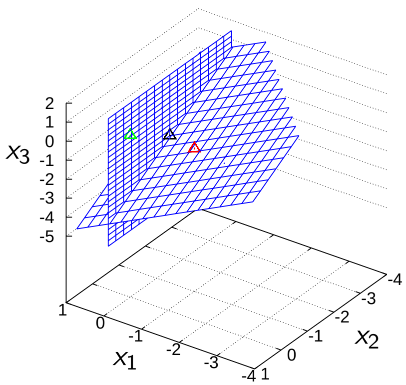
- Complementarity slackness models the disjunction

$$\bar{\nu}_i = 0 \ \vee \ g_i(\bar{\mathbf{x}}) = 0, \quad \text{for each constraint } i = 1, \ldots, m$$

Checking the KKT Conditions

Class Exercise: Consider the inequality constrained NLP

$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)$$
$$\text{s.t. } g_1(\mathbf{x}) \triangleq x_1 + x_2 + x_3 \leq -3,$$
$$g_2(\mathbf{x}) \triangleq x_1 \leq 0,$$

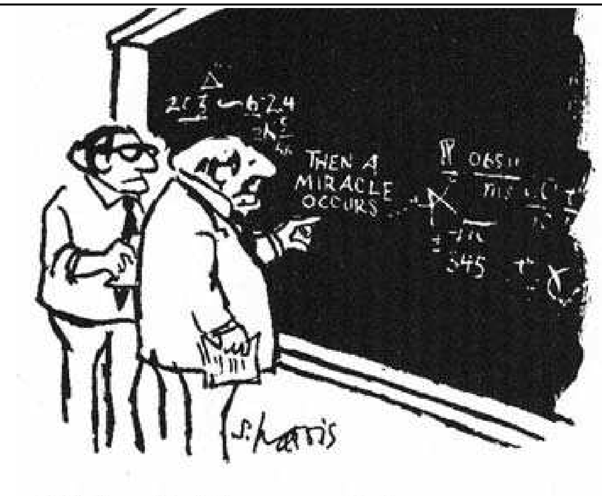


Write and check the KKT conditions at:

- $\mathbf{x}^{(1)} = (-1,-1,-1)$, $\boldsymbol{\nu}^{(1)} = (-1,0)$
- $\mathbf{x}^{(2)} = (0,-\frac{3}{2},-\frac{3}{2})$, $\boldsymbol{\nu}^{(2)} = (-\frac{3}{2},\frac{3}{2})$
- $\mathbf{x}^{(3)} = (0,0,0)$, $\boldsymbol{\nu}^{(3)} = (0,0)$

From Geometric to Algebraic Optimality Conditions

$$\nexists \mathbf{d} \in \mathbb{R}^n : \nabla f(\mathbf{x}^*)^{\mathsf{T}}\mathbf{d} < 0 \text{ **and** } \nabla g_i(\mathbf{x}^*)^{\mathsf{T}}\mathbf{d} < 0, \forall i \in \mathcal{A}(\mathbf{x}^*)$$



"I think you should be more explicit here in step two."

First-Order Necessary Conditions for Optimality

First-Order Necessary Conditions

Let $f, g_1, \ldots, g_m$ be $\mathcal{C}^1$. If $\mathbf{x}^*$ is a (local) optimum for f s.t. $g_i(\mathbf{x}) \leq 0$, $i = 1, \ldots, m$, **and** regular, there is a unique vector $\boldsymbol{\nu}^* \in \mathbb{R}^m$ such that $(\mathbf{x}^*, \boldsymbol{\nu}^*)$ is a KKT point:

$$g_i(\mathbf{x}^*) \leq 0, \quad \nu_i^* g_i(\mathbf{x}^*) = 0, \quad i = 1, \ldots, m$$
$$\nabla f(\mathbf{x}^*) = \sum_{i=1}^m \nu_i^* \nabla g_i(\mathbf{x}^*), \quad \boldsymbol{\nu}^* \leq \mathbf{0}$$

KKT Multipliers (Minimize):

- $g_i(\mathbf{x}) \leq 0 \leftrightarrow \nu_i^* \leq 0$
- $g_i(\mathbf{x}) \geq 0 \leftrightarrow \nu_i^* \geq 0$

Interpretation:

- $g_i(\mathbf{x}) \leq b_i \Rightarrow \nu_i^* \triangleq \frac{\partial f^*}{\partial b_i}$

Active Set Selection:

1. Pick-up active set (a priori)
2. Calculate KKT point $(\mathbf{x}^*, \boldsymbol{\nu}^*)$ (if any)

$\rightarrow$ Repeat for **all** possible active sets!

Solving Nonlinear Inequality Constrained Optimization

Class Exercise: Use the 1st-order necessary conditions for optimality to single out candidate optimal points

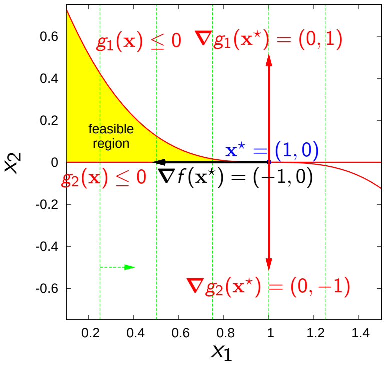
$$\min_{\mathbf{x}} f(\mathbf{x}) \triangleq \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)$$
$$\text{s.t. } g_1(\mathbf{x}) \triangleq x_1 + x_2 + x_3 \leq -3,$$
$$g_2(\mathbf{x}) \triangleq x_1 \leq 0,$$

- g_1, g_2 inactive
- g_1 active, g_2 inactive
- g_1 inactive, g_2 active
- g_1, g_2 active

Solving Nonlinear Inequality Constrained Optimization

Class Exercise: Identify the optimum point (visually) and check the 1st-order necessary conditions for optimality at that point

$$\min_{\mathbf{x}} \; f(\mathbf{x}) \triangleq -x_1$$
$$\text{s.t. } g_1(\mathbf{x}) \triangleq -(1-x_1)^3 + x_2 \leq 0$$
$$g_2(\mathbf{x}) \triangleq -x_2 \leq 0$$



What's going on here?

First-Order Necessary Conditions for Optimality

- Let $f, g_1, \ldots, g_{m_i}, h_1, \ldots, h_{m_e}$ be $\mathcal{C}^1$
- Suppose that $\mathbf{x}^*$ is a (local) optimum point for

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize:}} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m_i \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m_e \end{aligned}$$

and a regular point for the equality and active inequality constraints

- Then, there exist (unique) multiplier vectors $\boldsymbol{\nu}^* \in \mathbb{R}^{m_i}$, $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_e}$ such that $(\mathbf{x}^*, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*)$ satisfy:

$$\boxed{\begin{gathered} g_i(\mathbf{x}^*) \leq 0, \quad \nu_i^* g_i(\mathbf{x}^*) = 0, \quad \nu_i \leq 0, \quad i = 1, \ldots, m_i \\ h_i(\mathbf{x}^*) = 0, \quad i = 1, \ldots, m_e \\ \nabla f(\mathbf{x}^*) = \sum_{i=1}^{m_i} \nu_i^* \nabla g_i(\mathbf{x}^*) + \sum_{i=1}^{m_e} \lambda_i^* \nabla h_i(\mathbf{x}^*) \end{gathered}}$$

First-Order Sufficiency Conditions for Optimality

- Consider the NLP problem

$$\begin{aligned} \underset{\mathbf{x}\in\mathbb{R}^n}{\text{minimize:}} \quad & f(\mathbf{x}) \\ \text{subject to:} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1, \ldots, m_i \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \ldots, m_e \end{aligned}$$

where:
  - $f, g_1, \ldots, g_{m_i}, h_1, \ldots, h_{m_e}$ are $\mathcal{C}^1$
  - $f, g_1, \ldots, g_{m_i}$ are convex and $h_1, \ldots, h_{m_e}$ are affine on $\mathbb{R}^n$

- Suppose that $(\mathbf{x}^*, \boldsymbol{\nu}^*, \boldsymbol{\lambda}^*)$ satisfy:

$$\boxed{\begin{gathered} g_i(\mathbf{x}^*) \leq 0, \quad \nu_i^* g_i(\mathbf{x}^*) = 0, \quad \nu_i \leq 0, \quad i = 1, \ldots, m_i \\ h_i(\mathbf{x}^*) = 0, \quad i = 1, \ldots, m_e \\ \nabla f(\mathbf{x}^*) = \sum_{i=1}^{m_i} \nu_i^* \nabla g_i(\mathbf{x}^*) + \sum_{i=1}^{m_e} \lambda_i^* \nabla h_i(\mathbf{x}^*) \end{gathered}}$$

- Then, $\mathbf{x}^*$ is a **global** optimum point

Applying First-Order Sufficiency Conditions

Class Exercise: Use the 1st-order sufficient conditions for optimality to determine a global optimum

$$\min_{\mathbf{x}} \; f(\mathbf{x}) \triangleq \frac{1}{2}\left(x_1^2 + x_2^2 + x_3^2\right)$$
$$\text{s.t. } g_1(\mathbf{x}) \triangleq x_1 + x_2 + x_3 - 3 = 0,$$
$$g_2(\mathbf{x}) \triangleq x_1 \leq 0,$$

- g_2 active
- g_2 inactive

The Final Words

Optimality conditions for NLP with equality and/or inequality constraints:

- **1st-Order Necessary Conditions**: A local optimum of a (differentiable) NLP must be a KKT point **if**:
  - that point is regular; or
  - all the (active) constraints are linear
- **1st-Order Sufficiency Conditions**: A KKT point of a convex (differentiable) NLP is a global optimum optimum of a convex program
- **2nd-Order Necessary and Sufficiency Conditions** can also be derived for constrained NLP, similar to unconstrained NLP...

"Everything should be made as simple as possible, but no simpler."

— Albert Einstein.