02610
Optimization and Data Fitting

Week 12: Introduction to Constrained Optimization

Yiqiu Dong

DTU Compute
Technical University of Denmark

23 November 2020

Constrained optimization problems

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & c_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{E} \\ & c_i(\mathbf{x}) \geq 0 \text{ for } i \in \mathcal{I} \end{aligned}$$

- $\mathbf{x}$ is the unknown vector.
- f is the objective function.
- c_i are constraint functions, and $\mathcal{E}$ and $\mathcal{I}$ are sets of indices for equality and inequality constraints, respectively.
- **Assumption:** f and c_i are all smooth.
- Feasible set: The set of all possible $\mathbf{x}$, i.e., the points satisfy all constraints.

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n : c_i(\mathbf{x}) = 0, i \in \mathcal{E}; c_i(\mathbf{x}) \geq 0, i \in \mathcal{I}\}.$$

What is a solution?

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & c_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{E} \\ & c_i(\mathbf{x}) \geq 0 \text{ for } i \in \mathcal{I} \end{aligned}$$

We would most like to have global minimizer, i.e.,

A points $\mathbf{x}^*$ is a global minimizer if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \Omega$ with $\Omega = \{\mathbf{x} \in \mathbb{R}^n : c_i(\mathbf{x}) = 0, i \in \mathcal{E}; c_i(\mathbf{x}) \geq 0, i \in \mathcal{I}\}$.

But the global minimizer can be difficult to find due to limited knowledge of f, most algorithms are able to find only a local minimizer, i.e.,

A points $\mathbf{x}^*$ is a local minimizer if $\mathbf{x}^* \in \Omega$ and there is a neighborhood $\mathcal{N}$ of $\mathbf{x}^*$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{N} \cap \Omega$.

What is a solution?

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & c_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{E} \\ & c_i(\mathbf{x}) \geq 0 \text{ for } i \in \mathcal{I} \end{aligned}$$

We would most like to have global minimizer, i.e.,

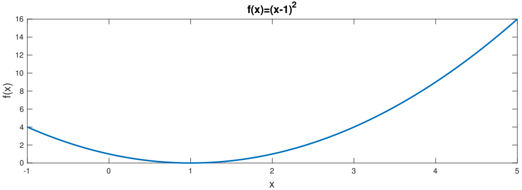
A points $\mathbf{x}^*$ is a global minimizer if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \Omega$ with $\Omega = \{\mathbf{x} \in \mathbb{R}^n : c_i(\mathbf{x}) = 0, i \in \mathcal{E}; c_i(\mathbf{x}) \geq 0, i \in \mathcal{I}\}$.

But the global minimizer can be difficult to find due to limited knowledge of f, most algorithms are able to find only a local minimizer, i.e.,

A points $\mathbf{x}^*$ is a local minimizer if $\mathbf{x}^* \in \Omega$ and there is a neighborhood $\mathcal{N}$ of $\mathbf{x}^*$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{N} \cap \Omega$.

Example 1

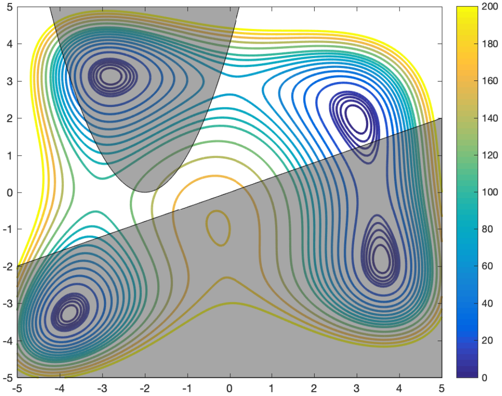
$$\min_{x\in\mathbb{R}} f(x) = (x-1)^2$$



- With the constraint $x \geq 0$, the minimizer is $x^* = 1$.
- With the constraint $x - 2 \geq 0$, we have $\Omega = [2, +\infty)$ and the minimizer is $x^* = 2$.
- With the constraint $x - 2 \geq 0$ and $3 - x \geq 0$, we have $\Omega = [2, 3]$ and the minimizer is $x^* = 2$.
- With the constraint $3 - x = 0$, we have $\Omega = \{3\}$ and the minimizer is $x^* = 3$.
- With the constraint $3 - x \geq 0$ and $x - 4 \geq 0$, we have $\Omega = \emptyset$ and the minimizer does not exist.

Example 2

$$
\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^2} \quad & f(\mathbf{x}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7) \\
\text{subject to} \quad & c_1(\mathbf{x}) = (x_1+2)^2 - x_2 \geq 0 \\
& c_2(\mathbf{x}) = -4x_1 + 10x_2 \geq 0
\end{aligned}
$$

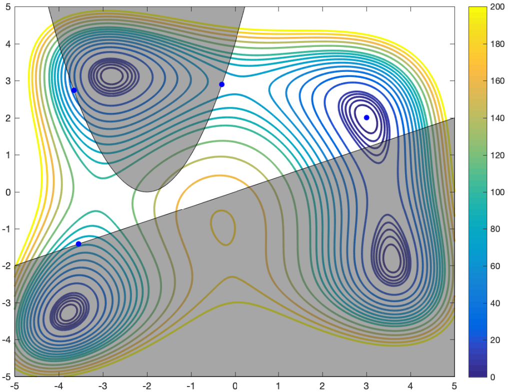


Example 2

$$\min_{\mathbf{x}\in\mathbb{R}^2} \quad f(\mathbf{x}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)$$

$$\text{subject to} \quad c_1(\mathbf{x}) = (x_1+2)^2 - x_2 \geq 0$$

$$c_2(\mathbf{x}) = -4x_1 + 10x_2 \geq 0$$



Active set

At a feasible point $\mathbf{x}$, the inequality constraint c_i ($i \in \mathcal{I}$) is:

- *active* iff $c_i(\mathbf{x}) = 0$ ($\mathbf{x}$ is on the boundary for that constraint);
- *inactive* iff $c_i(\mathbf{x}) > 0$ ($\mathbf{x}$ is interior point for that constraint).

The active set $\mathcal{A}(\mathbf{x})$ is the set of indices of equality constraints and active inequality constriants:

$$\mathcal{A}(\mathbf{x}) = \mathcal{E} \cup \{i \in \mathcal{I} : c_i(\mathbf{x}) = 0\}.$$

- An inequality constraint which is inactive at $\mathbf{x}$ has no influence on the optimization problem in a neighborhood of $\mathbf{x}$.

Convex optimization problems

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & c_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{E} \\ & c_i(\mathbf{x}) \geq 0 \text{ for } i \in \mathcal{I} \end{aligned}$$

- The feasible set Ω must be convex.
- The objective function f must be convex on Ω.
- The equality constraint functions c_i with $i \in \mathcal{E}$ must be linear.
- The inequality constraint functions c_i with $i \in \mathcal{I}$ must be concave.

Optimality condition

If Ω is bounded and convex and if f is convex on Ω, then

- any local minimizer $\mathbf{x}^* \in \Omega$ is a global minimizer.
- If f is strictly convex, then the global minimizer is unique.

Convex optimization problems

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & c_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{E} \\ & c_i(\mathbf{x}) \geq 0 \text{ for } i \in \mathcal{I} \end{aligned}$$

- The feasible set Ω must be convex.
- The objective function f must be convex on Ω.
- The equality constraint functions c_i with $i \in \mathcal{E}$ must be linear.
- The inequality constraint functions c_i with $i \in \mathcal{I}$ must be concave.

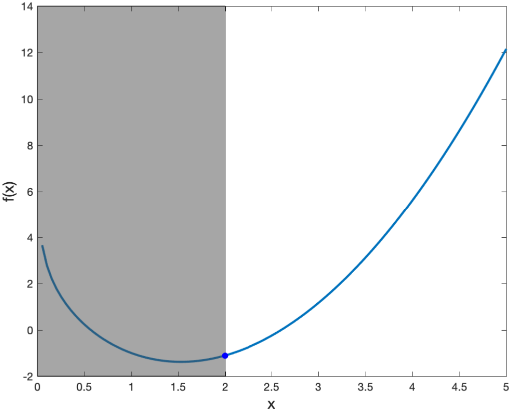
Optimality condition

If Ω is bounded and convex and if f is convex on Ω, then

- any local minimizer $\mathbf{x}^* \in \Omega$ is a global minimizer.
- If f is strictly convex, then the global minimizer is unique.

Example: Convex problem 1

$$\min_{x} \; f(x) = (x-1)^2 - \sqrt{x} - \ln(x)$$
$$\text{s.t. } c(x) = x - 2 \geq 0$$



Example: Convex problem 2

$$\min_{\mathbf{x}\in\mathbb{R}^2_{++}} \; f(x) = (x_1-1)^2 + (x_2-1)^2 - \sqrt{x_1+x_2} - \ln(x_1) - \ln(x_2)$$

$$\text{s.t. } c_1(\mathbf{x}) = x_1 + x_2 - 4 \geq 0$$

$$c_2(\mathbf{x}) = -x_1^2 - x_2^2 + 16 \geq 0$$

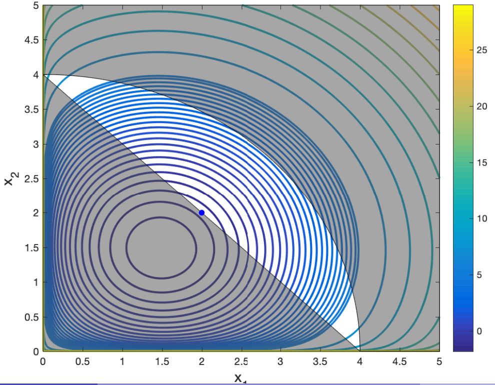


Illustration of optimality conditions

- **Case 0:** No equality constraints, no active inequality constraints.

  We can consider it as unconstrained problem and move in a descent direction. If the step length is not too large, then the constraints are of no consequence.

  Interior local minimizer should satisfy $\nabla f(\mathbf{x}^*) = 0$.

- **Case 1:** One equality constraint (no inequality constraints).
  - $\mathbf{x}^*$ may be a local, constrained minimizer and it satisfies

$$\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*) \text{ with } \lambda \in \mathbb{R}.$$

- **Case 2:** A single inequality constraint (no equality constraints).

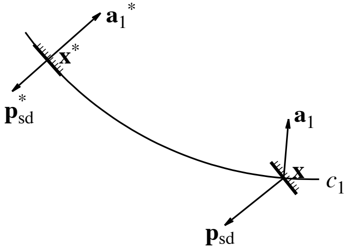
  If no decrease of f is possible:
  - Interior point $c_1(\mathbf{x}^*) > 0$ (inactive constraint): $\nabla f(\mathbf{x}^*) = 0$.
  - Boundary point $c_1(\mathbf{x}^*) = 0$: There cannot be a direction $\mathbf{p}$ such that $\nabla f(\mathbf{x}^*)^T \mathbf{p} < 0$ and $\nabla c_1(\mathbf{x}^*)^T \mathbf{p} \geq 0$.
    $\mathbf{x}^*$ should satisfy $\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*)$ with $\lambda \geq 0$.

Illustration of optimality conditions

- **Case 0:** No equality constraints, no active inequality constraints.

  Interior local minimizer should satisfy $\nabla f(\mathbf{x}^*) = 0$.

- **Case 1:** One equality constraint (no inequality constraints).



  - The feasible set is $\Omega = \{\mathbf{x} : c_1(\mathbf{x}) = 0\}$.
  - $\mathbf{p}_{\mathrm{sd}} = -\nabla f(\mathbf{x})$ is the steepest descent direction.
  - $\mathbf{a}_1 = \nabla c_1(\mathbf{x})$ is the constraint gradient.
  - $\mathbf{x}^*$ may be a local, constrained minimizer and it satisfies

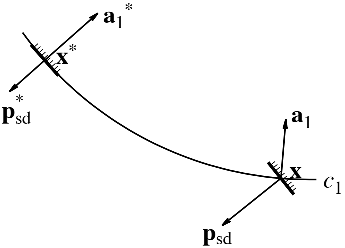
$$\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*) \text{ with } \lambda \in \mathbb{R}.$$

Illustration of optimality conditions

- **Case 0:** No equality constraints, no active inequality constraints.

  Interior local minimizer should satisfy $\nabla f(\mathbf{x}^*) = 0$.

- **Case 1:** One equality constraint (no inequality constraints).



  - The feasible set is $\Omega = \{\mathbf{x} : c_1(\mathbf{x}) = 0\}$.
  - $\mathbf{p}_{\mathrm{sd}} = -\nabla f(\mathbf{x})$ is the steepest descent direction.
  - $\mathbf{a}_1 = \nabla c_1(\mathbf{x})$ is the constraint gradient.
  - $\mathbf{x}^*$ may be a local, constrained minimizer and it satisfies

$$\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*) \text{ with } \lambda \in \mathbb{R}.$$

Illustration of optimality conditions

- **Case 0:** No equality constraints, no active inequality constraints.

  Interior local minimizer should satisfy $\nabla f(\mathbf{x}^*) = 0$.

- **Case 1:** One equality constraint (no inequality constraints).
  - $\mathbf{x}^*$ may be a local, constrained minimizer and it satisfies

$$\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*) \text{ with } \lambda \in \mathbb{R}.$$

- **Case 2:** A single inequality constraint (no equality constraints).

  If no decrease of f is possible:
  - Interior point $c_1(\mathbf{x}^*) > 0$ (inactive constraint): $\nabla f(\mathbf{x}^*) = 0$.
  - Boundary point $c_1(\mathbf{x}^*) = 0$:
    $\mathbf{x}^*$ should satisfy $\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*)$ with $\lambda \geq 0$.

Illustration of optimality conditions

- **Case 3:** Two inequality constraints (no equality constraints).

  If no decrease of f is possible:

  - Both are inactive: $\nabla f(\mathbf{x}^*) = 0$.
  - One active ($c_1(\mathbf{x}^*) = 0$) and the other inactive ($c_2(\mathbf{x}^*) > 0$): $\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*)$ with $\lambda \geq 0$.
  - Both active ($c_1(\mathbf{x}^*) = c_2(\mathbf{x}^*) = 0$):

Illustration of optimality conditions

- **Case 3:** Two inequality constraints (no equality constraints).

  If no decrease of f is possible:

  - Both are inactive: $\nabla f(\mathbf{x}^*) = 0$.
  - One active ($c_1(\mathbf{x}^*) = 0$) and the other inactive ($c_2(\mathbf{x}^*) > 0$): $\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*)$ with $\lambda \geq 0$.
  - Both active ($c_1(\mathbf{x}^*) = c_2(\mathbf{x}^*) = 0$):

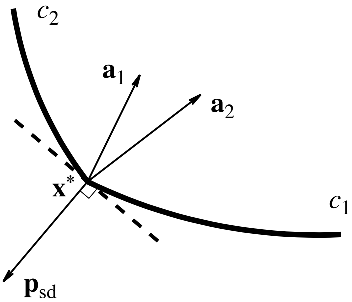


Illustration of optimality conditions

- **Case 3:** Two inequality constraints (no equality constraints).

  If no decrease of f is possible:

  - Both are inactive: $\nabla f(\mathbf{x}^*) = 0$.
  - One active ($c_1(\mathbf{x}^*) = 0$) and the other inactive ($c_2(\mathbf{x}^*) > 0$): $\nabla f(\mathbf{x}^*) = \lambda \nabla c_1(\mathbf{x}^*)$ with $\lambda \geq 0$.
  - Both active ($c_1(\mathbf{x}^*) = c_2(\mathbf{x}^*) = 0$): There cannot be a direction $\mathbf{p}$ such that $\nabla f(\mathbf{x}^*)^T \mathbf{p} < 0$ and $\nabla c_i(\mathbf{x}^*)^T \mathbf{p} \geq 0$ with $i \in \mathcal{I}$.

    $\mathbf{x}^*$ should satisfy

$$\nabla f(\mathbf{x}^*) = \lambda_1 \nabla c_1(\mathbf{x}^*) + \lambda_2 \nabla c_2(\mathbf{x}^*) \quad \text{with} \quad \lambda_1, \lambda_2 \geq 0$$

Lagrangian function

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & c_i(\mathbf{x}) = 0 \text{ for } i \in \mathcal{E} \\ & c_i(\mathbf{x}) \geq 0 \text{ for } i \in \mathcal{I} \end{aligned} \tag{1}$$

Lagrangian function

The Lagrangian function for the problem (1) is defined by

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i c_i(\mathbf{x}),$$

where $\{\lambda_i\}$ are the *Lagrangian multipliers*.

- The gradient of $\mathcal{L}$ with respect to $\mathbf{x}$ is denoted as

$$\nabla_{\mathbf{x}}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \nabla f(\mathbf{x}) - \sum_{i\in\mathcal{E}\cup\mathcal{I}} \lambda_i \nabla c_i(\mathbf{x}).$$

First-order necessary conditions (KKT conditions)

Theorem

Suppose that

1. $\mathbf{x}^*$ is a local minimizer of (1), where f and c_i are continuously differentiable;
2. either all active constrains c_i are linear,
   or the gradients $\{\nabla c_i(\mathbf{x}^*)\}$ for all active constraints are linearly independent *(LICQ)*.

Then there exists a Lagrangian multiplier vector $\boldsymbol{\lambda}^*$ such that

1. $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$, *(stationary)*
2. $c_i(\mathbf{x}^*) = 0$ for $i \in \mathcal{E}$, *(feasibility)*
3. $c_i(\mathbf{x}^*) \geq 0$ for $i \in \mathcal{I}$, *(feasibility)*
4. $\lambda_i^* \geq 0$ for $i \in \mathcal{I}$,
5. $\lambda_i^* c_i(\mathbf{x}^*) = 0$, for $i \in \mathcal{E} \cup \mathcal{I}$. *(complementarity)*

KKT conditions

- For an equality constraint $c_i(\mathbf{x}^*) = 0$, λ_i^* can have any sign.
- For an active inequality constraint $c_i(\mathbf{x}^*) = 0$, λ_i^* is nonnegative.
- For an inactive inequality constraint $c_i(\mathbf{x}^*) > 0$, we must have $\lambda_i^* = 0$.
- If $\mathcal{I} = \emptyset$, then KKT is reduced to $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \mathbf{0}$ together with equality constraints.
- $f(\mathbf{x}^*) = \mathcal{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*)$.
- **Strict complementarity:** Exactly one of λ_i^* and $c_i(\mathbf{x}^*)$ is zero for each $i \in \mathcal{I}$.
- **Constrained stationary point:** $\mathbf{x}^*$ is feasible and $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ satisfies the KKT conditions.

Example: Equality constrained problem

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}), \quad f \text{ convex} \\ \text{subject to} \quad & A\mathbf{x} = \mathbf{b} \end{aligned}$$

KKT conditions for this problem are

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + A^T \boldsymbol{\lambda}^* = \mathbf{0}$$
$$A\mathbf{x}^* - \mathbf{b} = \mathbf{0}$$

We can use Newton's method to solve this system of nonlinear equations:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k^{\mathbf{x}}, \quad \boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \alpha_k \mathbf{p}_k^{\boldsymbol{\lambda}},$$

where $\mathbf{p}_k^{\mathbf{x}}$ and $\mathbf{p}_k^{\boldsymbol{\lambda}}$ is obtained by solving

$$\begin{bmatrix} \nabla_{\mathbf{x}}^2 f(\mathbf{x}_k) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p}_k^{\mathbf{x}} \\ \mathbf{p}_k^{\boldsymbol{\lambda}} \end{bmatrix} = - \begin{bmatrix} \nabla_{\mathbf{x}} f(\mathbf{x}_k) + A^T \boldsymbol{\lambda}_k \\ A\mathbf{x}_k - \mathbf{b} \end{bmatrix}$$

Example: Equality constrained problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad f(\mathbf{x}), \quad f \text{ convex}$$
$$\text{subject to} \quad A\mathbf{x} = \mathbf{b}$$

KKT conditions for this problem are

$$\nabla_{\mathbf{x}} f(\mathbf{x}^*) + A^T \boldsymbol{\lambda}^* = \mathbf{0}$$
$$A\mathbf{x}^* - \mathbf{b} = \mathbf{0}$$

We can use Newton's method to solve this system of nonlinear equations:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k^{\mathbf{x}}, \quad \boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \alpha_k \mathbf{p}_k^{\boldsymbol{\lambda}},$$

where $\mathbf{p}_k^{\mathbf{x}}$ and $\mathbf{p}_k^{\boldsymbol{\lambda}}$ is obtained by solving

$$\begin{bmatrix} \nabla^2_{\mathbf{x}} f(\mathbf{x}_k) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p}_k^{\mathbf{x}} \\ \mathbf{p}_k^{\boldsymbol{\lambda}} \end{bmatrix} = - \begin{bmatrix} \nabla_{\mathbf{x}} f(\mathbf{x}_k) + A^T \boldsymbol{\lambda}_k \\ A\mathbf{x}_k - \mathbf{b} \end{bmatrix}$$

Example: Equality constrained problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad f(\mathbf{x}), \quad f \text{ convex}$$
$$\text{subject to} \quad A\mathbf{x} = \mathbf{b}$$

KKT conditions for this problem are

$$\begin{aligned} \nabla_{\mathbf{x}} f(\mathbf{x}^*) + A^T\boldsymbol{\lambda}^* &= \mathbf{0} \\ A\mathbf{x}^* - \mathbf{b} &= \mathbf{0} \end{aligned}$$

We can use Newton's method to solve this system of nonlinear equations:

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k^{\mathbf{x}}, \quad \boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_k + \alpha_k \mathbf{p}_k^{\boldsymbol{\lambda}},$$

where $\mathbf{p}_k^{\mathbf{x}}$ and $\mathbf{p}_k^{\boldsymbol{\lambda}}$ is obtained by solving

$$\begin{bmatrix} \nabla_{\mathbf{x}}^2 f(\mathbf{x}_k) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p}_k^{\mathbf{x}} \\ \mathbf{p}_k^{\boldsymbol{\lambda}} \end{bmatrix} = - \begin{bmatrix} \nabla_{\mathbf{x}} f(\mathbf{x}_k) + A^T\boldsymbol{\lambda}_k \\ A\mathbf{x}_k - \mathbf{b} \end{bmatrix}$$

Quadratic programs

The quadratic programming problem

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{g}^T\mathbf{x} \\ \text{s.t.} \quad & \mathbf{a}_i^T\mathbf{x} = b_i, \quad i \in \mathcal{E} \\ & \mathbf{a}_i^T\mathbf{x} \geq b_i, \quad i \in \mathcal{I} \end{aligned}$$

- If H is positive semidefinite, it is a convex QP.
- If H is positive definite, it is a strictly convex QP.
- $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{g}^T\mathbf{x} - \boldsymbol{\lambda}^T(A^T\mathbf{x} - \mathbf{b})$.

Quadratic programs

The quadratic programming problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{g}^T\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{a}_i^T\mathbf{x} = b_i, \quad i \in \mathcal{E}$$

- $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{g}^T\mathbf{x} - \boldsymbol{\lambda}^T(A^T\mathbf{x} - \mathbf{b})$.

If only equality constraint, KKT conditions are:

$$\begin{aligned} H\mathbf{x}^* + \mathbf{g} - A\boldsymbol{\lambda}^* &= \mathbf{0} \\ A^T\mathbf{x}^* - \mathbf{b} &= 0 \end{aligned}$$

Quadratic programs

The quadratic programming problem

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{g}^T\mathbf{x} \\ \text{s.t.} \quad & \mathbf{a}_i^T\mathbf{x} = b_i, \quad i \in \mathcal{E} \\ & \mathbf{a}_i^T\mathbf{x} \geq b_i, \quad i \in \mathcal{I} \end{aligned}$$

- $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = \frac{1}{2}\mathbf{x}^T H\mathbf{x} + \mathbf{g}^T\mathbf{x} - \boldsymbol{\lambda}^T(A^T\mathbf{x} - \mathbf{b})$.

In general case, KKT conditions are:

$$H\mathbf{x}^* + \mathbf{g} - \sum_{i\in\mathcal{A}(\mathbf{x}^*)} \lambda_i^* \mathbf{a}_i = \mathbf{0}$$

$$\begin{aligned} \mathbf{a}_i^T\mathbf{x}^* &= b_i \quad i \in \mathcal{A}(\mathbf{x}^*) \\ \mathbf{a}_i^T\mathbf{x} &\geq b_i \quad i \in \mathcal{I}\backslash\mathcal{A}(\mathbf{x}^*) \\ \lambda_i^* &\geq 0, \quad i \in \mathcal{I}\cap\mathcal{A}(\mathbf{x}^*) \end{aligned}$$

where $\mathcal{A}(\mathbf{x}^*) = \{i \in \mathcal{E}\cup\mathcal{I} : \mathbf{a}_i^T\mathbf{x}^* = b_i\}$ denotes the active set at $\mathbf{x}^*$.

Numerical methods for QP

Mainly three types:

- **Active-set methods.**
  - **Idea:** Identify the optimal active set from an initial guess for it, by repeatedly adding or subtracting one constraint each time.
  - Appropriate for small- or medium-scale problems; particularly for convex QP.
- **Gradient projection methods.**
  - **Idea:** Apply the steepest descent method but "bending" along the constraints.
  - Appropriate for large-scale problems; particularly simple with box constraints.
- **Interior-point methods.**
  - **Idea:** Apply Newton-like step on the KKT system.
  - Appropriate for large-scale problems.

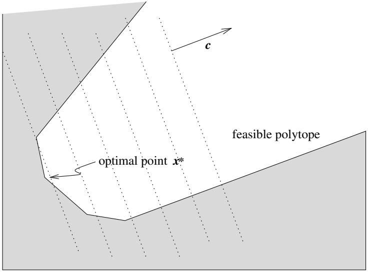
quadprog

```
QUADPROG Quadratic programming.
   X = QUADPROG(H,f,A,b) attempts to solve the quadratic programming
   problem:

            min 0.5*x'*H*x + f'*x   subject to:  A*x <= b
             x

   X = QUADPROG(H,f,A,b,Aeq,beq) solves the problem above while
   additionally satisfying the equality constraints Aeq*x = beq.

   X = QUADPROG(H,f,A,b,Aeq,beq,LB,UB) defines a set of lower and upper
   bounds on the design variables, X, so that the solution is in the
   range LB <= X <= UB. Use empty matrices for LB and UB if no bounds
   exist. Set LB(i) = -Inf if X(i) is unbounded below; set UB(i) = Inf if
   X(i) is unbounded above.

   X = QUADPROG(H,f,A,b,Aeq,beq,LB,UB,X0) sets the starting point to X0
```

Linear programs

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad f(\mathbf{x}) = \mathbf{g}^T\mathbf{x}$$
$$\text{s.t.} \quad A\mathbf{x} = \mathbf{b}$$
$$\mathbf{x} \geq 0$$

- The objective function is linear.
- The constraints are linear.
- It is convex as well as concave.
- The feasible set is a polytope.
- The number of minimizers: 0 (infeasible or unbounded), 1 (a vertex) or ∞ (a face).
- $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{s}) = \mathbf{g}^T\mathbf{x} - \boldsymbol{\lambda}^T(A\mathbf{x} - \mathbf{b}) - \mathbf{s}^T\mathbf{x}$.



Linear programs

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} \quad & f(\mathbf{x}) = \mathbf{g}^T\mathbf{x} \\ \text{s.t.} \quad & A\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

- $\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \mathbf{s}) = \mathbf{g}^T\mathbf{x} - \boldsymbol{\lambda}^T(A\mathbf{x} - \mathbf{b}) - \mathbf{s}^T\mathbf{x}$.

KKT conditions are:

$$\begin{aligned} A^T\boldsymbol{\lambda}^* + \mathbf{s}^* &= \mathbf{g} \\ A\mathbf{x}^* &= \mathbf{b} \\ \mathbf{x}^* &\geq 0 \\ \mathbf{s}^* &\geq 0 \\ x_i^* s_i^* &= 0, \quad i = 1, \cdots, n \end{aligned}$$

Linear programs

$$\min_{\mathbf{x}\in\mathbb{R}^n} \quad f(\mathbf{x}) = \mathbf{g}^T\mathbf{x}$$
$$\text{s.t.} \quad A\mathbf{x} = \mathbf{b}$$
$$\mathbf{x} \geq 0$$

Main type of methods:

- **Simplex method:** an active-set method.
- **Interior-point methods:** Apply Newton-like step on the KKT system. Appropriate for large-scale problems.

linprog

```
LINPROG Linear programming.
   X = LINPROG(f,A,b) attempts to solve the linear programming problem:

            min f'*x    subject to:   A*x <= b
             x

   X = LINPROG(f,A,b,Aeq,beq) solves the problem above while additionally
   satisfying the equality constraints Aeq*x = beq.

   X = LINPROG(f,A,b,Aeq,beq,LB,UB) defines a set of lower and upper
   bounds on the design variables, X, so that the solution is in
   the range LB <= X <= UB. Use empty matrices for LB and UB
   if no bounds exist. Set LB(i) = -Inf if X(i) is unbounded below;
   set UB(i) = Inf if X(i) is unbounded above.

   X = LINPROG(f,A,b,Aeq,beq,LB,UB,X0) sets the starting point to X0. This
   option is only available with the active-set algorithm. The default
   interior point algorithm will ignore any non-empty starting point.
```

fmincon

```
fmincon finds a constrained minimum of a function of several variables.
   fmincon attempts to solve problems of the form:
    min F(X)  subject to:  A*X  <= B, Aeq*X  = Beq (linear constraints)
     X                     C(X) <= 0, Ceq(X) = 0   (nonlinear constraints)
                              LB <= X <= UB        (bounds)

   fmincon implements four different algorithms: interior point, SQP,
   active set, and trust region reflective. Choose one via the option
   Algorithm: for instance, to choose SQP, set OPTIONS =
   optimoptions('fmincon','Algorithm','sqp'), and then pass OPTIONS to
   fmincon.

   X = fmincon(FUN,X0,A,B) starts at X0 and finds a minimum X to the
   function FUN, subject to the linear inequalities A*X <= B. FUN accepts
   input X and returns a scalar function value F evaluated at X. X0 may be
   a scalar, vector, or matrix.

   X = fmincon(FUN,X0,A,B,Aeq,Beq) minimizes FUN subject to the linear
   equalities Aeq*X = Beq as well as A*X <= B. (Set A=[] and B=[] if no
   inequalities exist.)

   X = fmincon(FUN,X0,A,B,Aeq,Beq,LB,UB) defines a set of lower and upper
   bounds on the design variables, X, so that a solution is found in
   the range LB <= X <= UB. Use empty matrices for LB and UB
   if no bounds exist. Set LB(i) = -Inf if X(i) is unbounded below;
   set UB(i) = Inf if X(i) is unbounded above.
```

Optimization related courses

- 02612 Constrained Optimization
  - Master course. Every spring.
- 02953 Convex optimization
  - PhD course. Every June.
- 02947 PDE constrained optimization
  - PhD course. Every odd year.
- Optimization technique is also applied in many other courses, such as model predictive control (02619), introduction to inverse problems (02624), machine learning for signal processing (02471), Mathematical modelling (02526), Image analysis (02502), etc..