# Linear Constraints

CPSC 406 – Computational Optimization

# Linear constraints

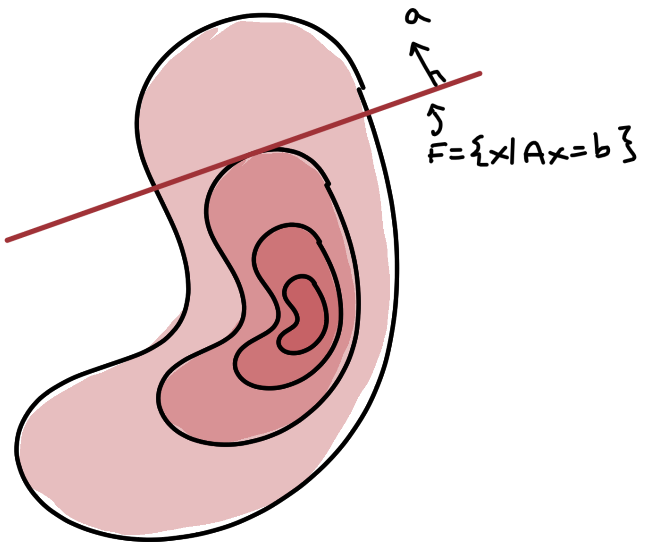
- underdetermined linear systems
- reduced-gradient methods

# Linearly-constrained optimization

$$\min_{x \in \mathbb{R}^n} \; \{ \, f(x) \mid Ax = b \, \}$$

- $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function
- $A$ is $m \times n$, $b$ is an $m$-vector, $m < n$ (underdetermined)
- assume throughout that $A$ has full row rank
- **feasible set**

$$\mathcal{F} = \{x \in \mathbb{R}^n \mid Ax = b\}$$



# Eliminating constraints

- equivalent representation of the feasible set

$$\mathcal{F} = \{x \in \mathbb{R}^n \mid Ax = b\} = \{\bar{x} + Zp \mid p \in \mathbb{R}^{n-m}\}$$

- $\bar{x}$ is a particular solution, ie, $A\bar{x} = b$
- $Z$ is a basis for the null space of $A$, ie, $AZ = 0$
- reduced problem is unconstrained in $n - m$ variables

$$\min_{p \in \mathbb{R}^{n-m}} \; f(\bar{x} + Zp)$$

- apply any unconstrained optimization method to solve to obtain solution $p^*$
- then a solution $x^* = \bar{x} + Zp^*$ is the solution to the original problem

# Example (in class)

$$\min_{x \in \mathbb{R}^2} \left\{ \tfrac{1}{2}(x_1^2 + x_2^2) \mid x_1 + x_2 = 1 \right\}$$

# Optimality conditions

- define **reduced** objective for any particular solution $\bar{x}$ and basis $Z$

$$f_Z(p) = f(\bar{x} + Zp)$$

$$\nabla f_Z(p) = Z^T \nabla f(\bar{x} + Zp) \quad \text{(reduced gradient)}$$

- let $p^*$ be a solution to the reduced problem and set $x^* = \bar{x} + Zp^*$

$$\nabla f_Z(p^*) = 0 \quad \Longleftrightarrow \quad Z^T \nabla f(x^*) = 0 \quad \Longleftrightarrow \quad \nabla f(x^*) \in \mathbf{null}(Z^T)$$

- funadamental subspaces of $A$ and $Z$ are orthogonal complements

$$\mathbf{null}(A) \equiv \mathbf{range}(Z) \quad \Longleftrightarrow \quad \mathbf{null}(Z^T) \equiv \mathbf{range}(A^T)$$

- thus,

$$\nabla f(x^*) \in \mathbf{null}(Z^T) \quad \Longleftrightarrow \quad \nabla f(x^*) \in \mathbf{range}(A^T) \quad \Longleftrightarrow \quad \exists y \text{ st } \nabla f(x^*) = A^T y$$

# First-order necessary conditions

A point $x^*$ is a local minimizer of the linearly-constrained problem **only if**

$$\exists\, y \in \mathbb{R}^m \text{ st } \nabla f(x^*) = A^T y \qquad [\text{optimality}]$$

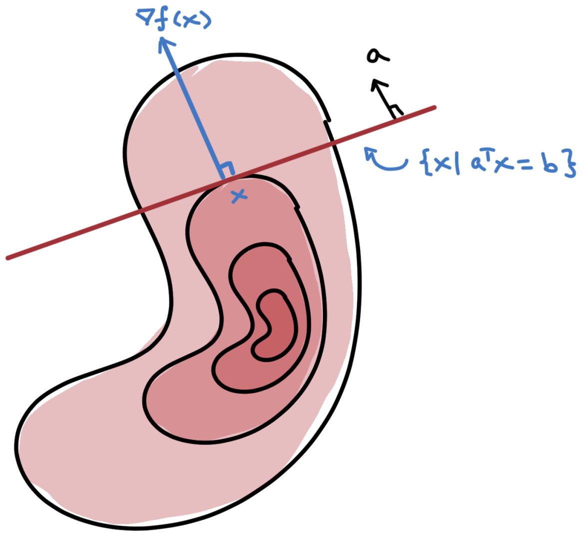
$$Ax^* = b \qquad [\text{feasibility}]$$

- optimality condition is equivalent to

$$Z^T \nabla f(x^*) = 0 \quad \iff \quad \nabla f(x^*)^\intercal p = 0 \quad \forall p \in \mathbf{null}(A)$$

- the $m$-vector $y$ contains the **Lagrange multipliers**

$$\nabla f(x^*) = A^T y = \sum_{i=1}^{m} y_i a_i$$



# Second-order optimality

$$f_Z(p) := f(\bar{x} + Zp) \qquad \nabla f_Z(p) := Z^T \nabla f(\bar{x} + Zp) \qquad \nabla^2 f_Z(p) := Z^T \nabla^2 f(\bar{x} + Zp) Z$$

**Necessary** 2nd-order optimality: $x^*$ is a local minimizer **only if**

$$\left.\begin{aligned} Ax^* &= b \\ Z^T \nabla f(x^*) &= 0 \\ Z^T \nabla^2 f(x^*) Z &\succeq 0 \end{aligned}\right\} \iff \left\{\begin{aligned} Ax^* &= b \\ \nabla f(x^*) &= A^T y && \text{for some } y \\ p^T \nabla^2 f_Z(p^*) p &\geq 0 \quad \forall p \in \mathbf{null}(A) \end{aligned}\right.$$

**Necessary and sufficient** 2nd-order optimality: $x^*$ is a local minimizer **if and only if**

$$\left.\begin{aligned} Ax^* &= b \\ Z^T \nabla f(x^*) &= 0 \\ Z^T \nabla^2 f(x^*) Z &\succ 0 \end{aligned}\right\} \iff \left\{\begin{aligned} Ax^* &= b \\ \nabla f(x^*) &= A^T y && \text{for some } y \\ p^T \nabla^2 f_Z(p^*) p &> 0 \quad \forall 0 \neq p \in \mathbf{null}(A) \end{aligned}\right.$$

# Example: Least norm solutions

$$\min_{x\in\mathbb{R}^n} \{ \|x\| \mid Ax = b \}$$

Take $f(x) = \frac{1}{2}\|x\|^2$ and apply first-order optimality conditions

$$\left.\begin{aligned} x &= A^T y \quad \text{for some } y \\ Ax &= b \end{aligned}\right\} \iff \begin{bmatrix} -I & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix}$$

A possible solution approach: - observe that multiplier $y$ satisfies

$$AA^T y = b$$

- factor $A^T = QR$ (thin QR factorization)
- multipliers: $y = R^{-1}R^{-T}b$
- solution: $x = A^T y = (QR)(R^{-1}R^{-T}b) = QR^{-T}b$

# Question

Find a minimimal norm solution to the linear system

$$\sum_{i=1}^{n} \xi_i = 1$$

If $n = 5$ and $x = (\xi_1, \xi_2, \xi_3, \xi_4, \xi_5)$, which of the following is a minimal norm solution?

a. $x = 1/5 \cdot (1, 1, 1, 1, 1)$
b. $x = 5 \cdot (5, 5, 5, 5, 5)$
c. $x = (1, 1, 1, 1, 1)$
d. $x = (1, 2, 3, 4, 5)$

# Question

What are the lagrange multipliers for the minimum-norm problem

$$\min_{x \in \mathbb{R}^n} \left\{ \|x\|^2 \mid Ax = b \right\}$$

where

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

# Reduced gradient method

$$\min_{x \in \mathbb{R}^n} \; \{f(x) \mid Ax = b\}$$

- choose $x^0$ and $Z$ such that $Ax^0 = b$ and $AZ = 0$
- for $k = 0, 1, 2, \ldots$
  - compute gradient $g^k = \nabla f(x^k)$
  - STOP if $\|Z^T g^k\|$ small
  - compute Hessian approximation $H^k \approx \nabla^2 f(x^k), H^k \succ 0$
  - solve $Z^T H^k Z p^k = -Z^T g^k$
  - linesearch on $f(x^k + \alpha Z p^k)$

# Obtaining a null-space basis

- assume variables (columns of $A$) permuted so that

$$A = [B \quad N] \quad \text{where} \quad B \quad \text{nonsingular}$$

- feasibility requires

$$b = Ax = Bx_B + Nx_N$$

- basic ($x_B$) and nonbasic ($x_N$) variables:
  - $x_N$ are "free"
  - $x_B = B^{-1}b - B^{-1}Nx_N$ uniquely determined by $x_N$
- constructing a null-space matrix

$$Z = \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} \quad \Longrightarrow \quad AZ = [B \quad N] \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} = 0$$