CPSC 406: Computational Optimization

LINEAR CONSTRAINTS

February 13, 2018

revised February 26, 2020

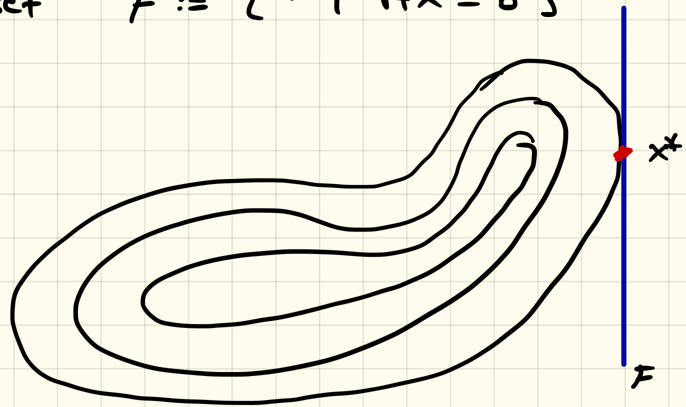
revised March 2, 2023

Linearly-Constrained Optimization

$$(P) \quad \underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) \quad \text{subj to} \quad Ax = b$$

- $f : \mathbb{R}^n \to \mathbb{R}$ continuously differentiable (once or twice, as needed)
- $A = \underbrace{\boxed{}}_{n}\, m$ with $\operatorname{rank}(A) = m$
- minimum value finite and attained

Feasible set $\mathcal{F} := \{ x \mid Ax = b \}$



Eliminating Constraints

Equivalent representation of the feasible set:

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

where
- $\bar{x}$ is any particular feasible solution, ie, $A\bar{x} = b$
- Z is a basis for null(A), ie, $Z = \underbrace{\begin{bmatrix} \; \\ \; \\ \; \end{bmatrix}}_{n-m} n$ st $AZ = 0$

Reduced problem is unconstrained in $n-m$ variables:

$$\underset{p \in \mathbb{R}^{n-m}}{\text{minimize}} \quad f(\bar{x} + Zp)$$

- apply any unconstrained method to this problem to obtain p^*
- obtain solution x^* from original problem as
$$x^* = \bar{x} + Zp^*.$$

EXAMPLE (IN CLASS)

minimize $\frac{1}{2}(x_1^2 + x_2^2)$ subj to $x_1 + x_2 = 1$

OPTIMALITY CONDITIONS

Define "reduced" objective for any particular sol'n $\bar{x}$ and basis Z:

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

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

Let p^* be solution and set $x^* = \bar{x} + Zp^*$. Then p^* is optimal only if

$$\nabla f_Z(p^*) = 0 \iff Z^T \nabla f(x^*) = 0 \iff \nabla f(x^*) \in \text{null}(Z^T)$$

Fundamental subspaces in $\mathbb{R}^n$ associated with A and Z:

$$\begin{array}{ccccc} \text{range}(A^T) & \oplus & \text{null}(A) & = & \mathbb{R}^n \\ \rotatebox{90}{$\equiv$} & & \rotatebox{90}{$\equiv$} & & \\ \text{null}(Z^T) & \oplus & \text{range}(Z) & = & \mathbb{R}^n \end{array}$$

Thus, $\nabla f(x^*) \in \text{null}(Z^T) \iff \nabla f(x^*) \in \text{range}(A^T)$

$$\iff \exists\, y \text{ st } \nabla f(x^*) = A^T y$$

FIRST-ORDER NECESSARY CONDITIONS

A point x^* is a local minimizer of (P) <u>only if</u>

there exists an m-vector y such that

[optimality] $\nabla f(x^*) = A^T y \equiv \sum_{i=1}^{m} a_i y_i$ $A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}$

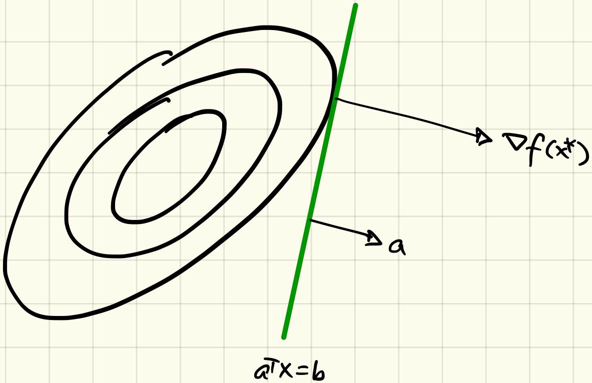
[feasibility] $Ax^* = b$

Equivalent to

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

$$Ax^* = b$$

The vector y is sometimes referred to as "Lagrange multipliers".



Second-Order Optimality

$$f_z(p) := f(\bar{x}+Zp), \quad \nabla f_z(p) = Z^T \nabla f(\bar{x}+Zp), \quad \nabla^2 f_z(p) = Z^T \nabla^2 f(\bar{x}+Zp) Z$$

<u>Necessary</u> 2^{nd}-order optimality: x^* is a local minimizer <u>only if</u>

$$\left.\begin{array}{c} Ax^* = b \\ Z^T \nabla f(x^*) = 0 \\ Z^T \nabla^2 f(x^*) Z \succeq 0 \end{array}\right\} \overset{\text{or}}{\iff} \left\{\begin{array}{l} Ax^* = b \\ \nabla f(x^*) = A^T y \text{ for some vector } y \\ p^T \nabla^2 f(x^*) p \geq 0 \quad \forall p \in \text{null}(A) \end{array}\right.$$

<u>Sufficient</u> 2^{nd}-order optimality: x^* is a local minimizer if

$$\left.\begin{array}{c} Ax^* = b \\ Z^T \nabla f(x^*) = 0 \\ Z^T \nabla^2 f(x^*) Z \succ 0 \end{array}\right\} \overset{\text{or}}{\iff} \left\{\begin{array}{l} Ax^* = b \\ \nabla f(x^*) = A^T y \text{ for some vector } y \\ p^T \nabla^2 f(x^*) p > 0 \quad \forall\, 0 \neq p \in \text{null}(A) \end{array}\right\}$$

EXAMPLE : LEAST-NORM SOLUTIONS

$$\text{minimize } \|x\|_2 \text{ subj to } Ax=b \quad \text{(underdetermined)}$$

Can take $f(x)=\frac{1}{2}\|x\|^2$. First-order optimality:

$$\left.\begin{array}{l} x = A^Ty \text{ for some } y \\ Ax=b \end{array}\right\} \Longleftrightarrow \begin{pmatrix} -I & A^T \\ A & \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}$$

EXAMPLE Find a minimal-norm solution to

$$x_1 + x_2 + \cdots + x_n = 1$$

$$\begin{pmatrix} -I & e \\ e^T & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\begin{array}{ll} -x + ey = 0 & \Rightarrow \quad x = ey \\ e^Tx \quad = 1 & \end{array}$$

$$\Rightarrow e^Tey = ny = 1 \quad y = 1/n \quad \Rightarrow \underline{x = \tfrac{1}{n}e}.$$

Reduced Gradient Method

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{st} & Ax = b \end{array} \iff \text{minimize } f_z(\bar{x} + Zp)$$

x_0 feasible ; Z basis for null(A)

for $k = 0, 1, 2, \ldots$

- Compute $g := \nabla f(x_k)$
- compute $H := \nabla^2 f(x_k)$
- solve $Z^T H Z p_k = -Z^T g$

  ie $\nabla^2 f_z(x_k) p_k = -\nabla f_z(x_k)$
- linesearch on $f(x_k + \alpha Z p_k)$
- $x_{k+1} = x_k + \alpha_k Z p_k$

Obtaining a Null-Space Basis

Permute variables (columns of A) so that

$$A = \begin{array}{c} \begin{bmatrix} B & N \end{bmatrix} \\ \;\; m \quad n-m \end{array} \qquad \text{where } B \text{ nonsingular}$$

$B \equiv$ "Basic"
$N \equiv$ "Non-basic"

Then feasibility requires

$$Ax = b \iff \begin{bmatrix} B & N \end{bmatrix} \begin{pmatrix} x_B \\ x_N \end{pmatrix} = b \iff Bx_B + Nx_N = b$$

Basic (x_B) and Non-Basic (x_N) variables:

- x_N free to move
- x_B uniquely determined by x_N, ie, $x_B = B^{-1}(b - Nx_N)$

Constructing a null-space matrix:

$$Z = \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} \implies AZ = \begin{bmatrix} B & N \end{bmatrix} \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix} = 0$$