16. Linear Constraint

- Reduced gradient
- Optimality conditions

Linear Constrained Optimization

$$\min_{x \in \mathbb{R}^n} \quad f(x) \qquad \text{s.t.} \quad Ax = b,$$

where • $A \in \mathbb{R}^{m \times n}$, $m \leq n$

• $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable

level sets of f

$x^* \leftarrow$ optimal point

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

$\mathcal{F}$ is the feasible set (set of points that satisfy const.)

Variable reduction

Reformulate the linear constrained problem into an unconstrained problem. (Change of variable).

Let

1. $\bar{x} \in \mathbb{R}^n$ be any point with $A\bar{x} = b$.
   $\bar{x}$ is a particular solution of $Ax = b$.
2. Z be a basis for $\mathrm{Null}(A)$.

$$N(A) \oplus R(A^T) = \mathbb{R}^n$$

If $A \in \mathbb{R}^{m \times n}$ is full row rank, $Z \in ?$

Feasible set:

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

Reduced Unconstrained problem

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad Ax = b \iff \min_{p \in \mathbb{R}^{n-m}} f(\bar{x} + Zp).$$

- Reduced the linearly constrained problem to an unconstrained problem in $n-m$ variables.
- Use any unconstrained opt. method to get p^*.
- Solution to constrained problem $x^* = \bar{x} + Zp^*$.

2nd order sufficient cond.

A point $x^* \in \mathbb{R}^n$ is a strict local minizer of $f(x)$ if

① $\nabla f(x^*) = 0$

② $\nabla^2 f(x^*) \succ 0$.

Optimality conditions.

Reduced objective function

$$f_z(p) = f(\bar{x} + Zp) = f\left(\bar{x} + \sum_{i=1}^{n-m} z_i p_i\right)$$

The gradient of f_z at p is:

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

Stationary:

A point p^* is stationary if

$$\nabla f_z(p^*) = 0 \iff Z^T \nabla f(\bar{x} + Zp^*) = 0$$

$$\iff \nabla f(\underbrace{\bar{x} + Zp^*}_{:= x^*}) \in \mathrm{Null}(Z^T)$$

First Order Necessary Cond.

Show $Null(Z^T) = Range(A^T)$. $A \in \mathbb{R}^{m \times n}$, $Z \in \mathbb{R}^{n \times (n-m)}$

$N(Z^T) \oplus R(Z) = \mathbb{R}^n$ and $R(A^T) \oplus N(A) = \mathbb{R}^n$

- Z is a basis for $Null(A)$

  $\Rightarrow N(Z^T) = R(A^T)$.

So, $\nabla f_z(x^*) \in N(Z^T) \iff \exists$ a $y \in \mathbb{R}^m$ such that $\nabla f(x^*) = A^T y$

First order necessary cond. A point x^* is a local minimizer of f ① only if there exists a point $y \in \mathbb{R}^m$ s.t.

① Optimality: $\nabla f(x^*) = A^T y$

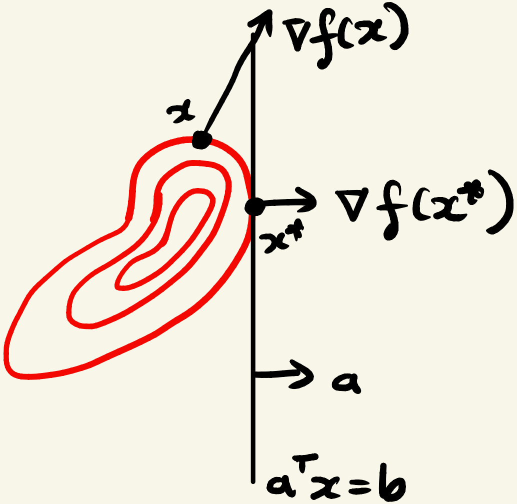
② Feasibility: $Ax^* = b$.

Integretation

The optimality cond: There exists $y \in \mathbb{R}^m$ s.t.

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

dual variable, Lagrange multiplier



require $\nabla f(x^*)$ aligned with normal vector a.

In general, require $\nabla f(x^*)$ to be in $R(A^T)$.

Second order necessary cond.

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

The Hessian of f_Z at p is:

$$\nabla^2 f_Z(p) = Z^T \nabla^2 f(\bar{x} + Zp) Z \in \mathbb{R}^{m \times m}$$

Second order necessary cond. A point x^* is a minimizer only if

① (Curvature) $Z^T \nabla^2 f(x^*) Z \succeq 0$

② (Feasibility) $Ax^* = b$.

Optimality condition.

Second order sufficient condition.

A point $x^* \in \mathbb{R}^n$ is a strict local minimizer of ① if it satisfies

1. (Feasibility) $Ax^* = b$
2. (Stationary) $\nabla f(x^*) \in \mathcal{N}(A^T)$

   $\Leftrightarrow \nabla f(x^*) = A^T y$ for some $y \in \mathbb{R}^m$

3. (Positivity) $Z^T \nabla^2 f(x^*) Z > 0$

   $\Leftrightarrow p^T \nabla^2 f(x^*) p > 0$ for all $p \in R(Z) \setminus \{0\}$

   $\Leftrightarrow p^T \nabla^2 f(x^*) p > 0$ for all $p \in \mathrm{Null}(A) \setminus \{0\}$.

Example

Consider
$$\min_{x \in \mathbb{R}^n} \|x\|_2 \quad \text{s.t.} \quad Ax = b.$$

WLOG: assume $f(x) = \frac{1}{2}\|x\|_2^2$ as objective function.

so $\nabla f(x) = x$

First order optimality:

① $\nabla f(x) \in R(A^T) \Rightarrow x = A^T y$ for some y.

② $Ax = b$.

$$\Leftrightarrow \begin{bmatrix} -I & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ b \end{bmatrix}$$

Example contd.

So, if $Ax = b$ is $e^T x = 1$, we get

$$\begin{bmatrix} -I & e \\ e^T & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Rightarrow \begin{cases} -x + e \cdot y = 0 \\ e^T x = 1 \end{cases}$$

$$\Rightarrow \quad -e^T x + e^T e y = 0$$

$$\Rightarrow \quad e^T e y = 1$$

$$\Rightarrow \quad y = \frac{1}{n}$$

$$\Rightarrow \quad x^* = \frac{e}{n}$$