Chapter 6 Optimization

LECTURE 10.

- objective (function)
- constrained vs. unconstrained
  - feasible choices (or points)

duality

e.g. $\begin{cases} \text{min weight} \\ \text{s.t. strength} \geq \end{cases} \leftrightarrow \begin{cases} \text{max strength.} \\ \text{s.t. weight} \leq \end{cases}$

$\begin{cases} \text{min cost} \\ \text{s.t. nutrition} \geq \end{cases} \leftrightarrow \begin{cases} \text{max nutrition} \\ \text{s.t. cost} \leq \end{cases}$

$f: \mathbb{R}^n \to \mathbb{R}$, $S \subseteq \mathbb{R}^n$

find x^* in S x^* "minimizer"
"minimum"

(max f is min of $-f$ $\rightarrow$ consider only minimization)

f objective function (linear or nonlinear
usually differentiable).

S constraints
inequalities, or equalities.

$x \in S \Rightarrow x$ "feasible"

$S = \mathbb{R}^n \Rightarrow$ "unconstrained"

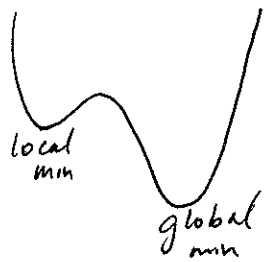
$$\min_x f(x)$$
subj. to $g(x) = 0$, and
$h(x) \leq 0$.

CLASSIFICATION

f, g, h linear or affine $\rightarrow$ linear programming
any of f, g, h nonlinear $\Rightarrow$ nonlinear programming

$f(x^*) \leq x \qquad \forall \; x \in S$ global minimum

$f(x^*) \leq x \qquad x \subset N(x^*) \subseteq S$ local minimum



Unless <u>special problem</u>, usually can't guarantee <u>global min</u>

- could, e.g., try many different starting points.

→ <u>Convex</u> programming problems

"discrete optimization" integer programming

§6.2.2 Unconstrained Optimality Conditions

Scalar case:

$f'(x) = 0$ $\quad$ $f''(x) > 0$ $\quad$ min

$\qquad\qquad\quad$ $f''(x) < 0$ $\quad$ max

$\qquad\qquad\quad$ ~~$f''(x) = 0$ $\quad$ inflection pt.~~

$\qquad\qquad\quad$ f''(x) = 0 $\quad$ inconclusive $\quad$ E.g., x^3 (inflection point), x^4 (minimum), -x^4 (maximum)

Vector case:

$f(x) \quad , \quad x \in \mathbb{R}^n$

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

gradient of f.

∇f points uphill

$-\nabla f$ points downhill

$$\boxed{\begin{array}{l} f(x+s) = f(x) + \nabla f(x+\alpha s)^T s \quad \text{for some } \alpha \in (0,1) \\ \text{choose } s = -\nabla f \end{array}}$$

First order necessary condition

Taylor's theorem

$$f(x+s) = f(x) + \nabla f(x)^T s + \frac{1}{2} s^T H s + \cdots \qquad \alpha \in (0,1)$$

let $s = -\alpha \nabla f(x)$

(stationary pt.
equilibrium pt.

$$f(x - \alpha \nabla f) = f(x) - \alpha \nabla f^T \nabla f + \underbrace{\frac{\alpha^2}{2} \nabla f^T H \nabla f}_{} + \cdots$$

$< f(x)$ for some $\alpha \in (0,1)$.

$\underline{\nabla f(x) = 0}$ first-order necessary condition

system of nonlinear equations.

x is a "critical point" ← necessary, but <u>not</u> sufficient

- x may be min, max, or neither (saddle pt.).

$f : \mathbb{R}^n \to \mathbb{R}$ twice differentiable <u>Hessian matrix of f</u>

$H_f : \mathbb{R}^n \to \mathbb{R}^{n \times n}$

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & & \ddots & \\ \vdots & & & \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

if 2^{nd} partial derivs of f continuous, then H_f <u>symmetric</u>

Let x^* be a critical pt. of f. + that $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable.

Taylor's theorem, $s \in \mathbb{R}^n$

$$f(x^* + s) = f(x) + \nabla f(x^*)^T s + \frac{1}{2} s^T H_f(x^* + \alpha s) s, \quad \alpha \in (0,1)$$

$\underline{H_f(x^*) \succ 0}$ second-order sufficient condition

$\nabla f(x^*) = 0$, + $H_f(x^*)$ is

CLASSIFICATION

- pos. def ⟹ x^* is a <u>min</u> of f
- neg. def ⟹ x^* is a <u>max</u> of f
- indef ⟹ x^* is a <u>saddle pt.</u> of f

Note: $H_f(x^*) \succ 0$ then f is <u>convex</u> in some nbhd of x^*.

Test for positive definiteness:

1. try to compute Cholesky factorization
2. LDL^T

} simple + cheap

3. eigenvalues — expensive!

Example. 6.5 Classifying Critical Pts.

$$f(x) = 2x_1^3 + 3x_1^2 + 12x_1x_2 + 3x_2^2 - 6x_2 + 6$$

$$\nabla f(x) = \begin{pmatrix} 6x_1^2 + 6x_1 + 12x_2 \\ 12x_1 + 6x_2 - 6 \end{pmatrix} = 0$$

Solving $\nabla f(x) = 0$, get $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -3 \end{pmatrix}$ critical points

$$H_f(x) = \begin{pmatrix} 12x_1 + 6 & 12 \\ 12 & 6 \end{pmatrix}$$ symmetric ✓

saddle $H_f\left(\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 12+6 & 12 \\ 12 & 6 \end{pmatrix} = \begin{pmatrix} 18 & 12 \\ 12 & 6 \end{pmatrix}$ not p.def ✗, $\lambda \cong 25.4, -1.4$

local min $H_f\left(\begin{pmatrix} 2 \\ -3 \end{pmatrix}\right) = \begin{pmatrix} 30 & 12 \\ 12 & 6 \end{pmatrix}$ pos def ✓, $\lambda \cong 35.0, 1.0$