Optimization

ht 2016

FORMULA SHEET

Line search

Golden section search:

$$\lambda_k = a_k + (1-\alpha)(b_k - a_k)$$
$$\mu_k = a_k + \alpha(b_k - a_k)$$

where $\alpha = \dfrac{\sqrt{5}-1}{2}$

Fibonacci search:

$$\lambda_k = a_k + (1-\alpha_{n,k})(b_k - a_k)$$
$$\mu_k = a_k + \alpha_{n,k}(b_k - a_k)$$

where $\alpha_{n,k} = \dfrac{F_{n-k}}{F_{n-k+1}}$

Fibonacci numbers: $\begin{cases} F_n = F_{n-1} + F_{n-2} \\ F_0 = 1,\ F_1 = 1 \end{cases}$

Newton's method:

$$\lambda_{k+1} = \lambda_k - \frac{F'(\lambda_k)}{F''(\lambda_k)}$$

Multidimensional search

$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k$, where λ_k solves $\min_\lambda f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k)$

Cyclic coordinates: $\boldsymbol{d}_i = \boldsymbol{e}_1, \ldots, \boldsymbol{e}_n, \boldsymbol{e}_1, \ldots$

Steepest descent: $\boldsymbol{d}_k = -\nabla f(\boldsymbol{x}_k)$

Newton's method: $\boldsymbol{d}_k = -\boldsymbol{H}(\boldsymbol{x}_k)^{-1}\nabla f(\boldsymbol{x}_k)$ and $\lambda_k = 1$

Modified Newton: $\boldsymbol{d}_k = -\big(\epsilon_k \boldsymbol{I} + \boldsymbol{H}(\boldsymbol{x}_k)\big)^{-1}\nabla f(\boldsymbol{x}_k)$

Quasi-Newton: (Davidon-Fletcher-Powell)

$\boldsymbol{y}_{j+1} = \boldsymbol{y}_j + \lambda_j \boldsymbol{d}_j$, where λ_j solves $\min_\lambda f(\boldsymbol{y}_j + \lambda \boldsymbol{d}_j)$ and where

$$\begin{aligned} \boldsymbol{d}_{j+1} &= -\boldsymbol{D}_{j+1}\nabla f(\boldsymbol{y}_{j+1}) \\ \boldsymbol{D}_{j+1} &= \boldsymbol{D}_j + \frac{\boldsymbol{p}_j \boldsymbol{p}_j^T}{\boldsymbol{p}_j^T \boldsymbol{q}_j} - \frac{\boldsymbol{D}_j \boldsymbol{q}_j \boldsymbol{q}_j^T \boldsymbol{D}_j}{\boldsymbol{q}_j^T \boldsymbol{D}_j \boldsymbol{q}_j} \quad \text{with} \\ & \boldsymbol{p}_j = \lambda_j \boldsymbol{d}_j \quad (\equiv \boldsymbol{y}_{j+1} - \boldsymbol{y}_j) \\ & \boldsymbol{q}_j = \nabla f(\boldsymbol{y}_{j+1}) - \nabla f(\boldsymbol{y}_j) \end{aligned}$$

Conjugate gradients: (Fletcher-Reeves)

$\boldsymbol{y}_{j+1} = \boldsymbol{y}_j + \lambda_j \boldsymbol{d}_j$, where λ_j solves $\min_\lambda f(\boldsymbol{y}_j + \lambda \boldsymbol{d}_j)$ and where

$$\boldsymbol{d}_{j+1} = -\nabla f(\boldsymbol{y}_{j+1}) + \beta_j \boldsymbol{d}_j, \qquad \beta_j = \frac{\|\nabla f(\boldsymbol{y}_{j+1})\|^2}{\|\nabla f(\boldsymbol{y}_j)\|^2}$$

If f quadratic function: $\beta_j = \dfrac{\boldsymbol{d}_j^T \boldsymbol{H}\, \nabla f(\boldsymbol{y}_{j+1})}{\boldsymbol{d}_j^T \boldsymbol{H} \boldsymbol{d}_j}$

Farkas' Theorem

Exactly one of $Ax \le 0,\ c^T x > 0$ and $A^T y = c,\ y \ge 0$ has a solution.

Linear Programming

$$\boldsymbol{A} = \left(\begin{array}{c|c} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \hline \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{array}\right) = \left(\begin{array}{c} \boldsymbol{A}_{J'} \\ \hline \boldsymbol{A}_{J''} \end{array}\right) = \left(\begin{array}{c|c} \boldsymbol{A}_{K'} & \boldsymbol{A}_{K''} \end{array}\right)$$

$$\boldsymbol{b} = \left(\begin{array}{c} \boldsymbol{b}_{J'} \\ \hline \boldsymbol{b}_{J''} \end{array}\right) \qquad \boldsymbol{c} = \left(\begin{array}{c|c} \boldsymbol{c}_{K'} & \boldsymbol{c}_{K''} \end{array}\right)$$

$$\begin{array}{lll} P: & \text{Minimize} & \boldsymbol{c}^T \boldsymbol{x} \\ & \text{subject to} & \boldsymbol{A}_{J'}\boldsymbol{x} \ge \boldsymbol{b}_{J'} \\ & & \boldsymbol{A}_{J''}\boldsymbol{x} = \boldsymbol{b}_{J''} \\ & & \boldsymbol{x}_{K'} \ge \boldsymbol{0} \\ & & \boldsymbol{x}_{K''} \text{ unrestricted} \end{array} \qquad \begin{array}{lll} D: & \text{Maximize} & \boldsymbol{b}^T \boldsymbol{y} \\ & \text{subject to} & \boldsymbol{A}_{K'}^T\boldsymbol{y} \le \boldsymbol{c}_{K'} \\ & & \boldsymbol{A}_{K''}^T\boldsymbol{y} = \boldsymbol{c}_{K''} \\ & & \boldsymbol{y}_{J'} \ge \boldsymbol{0} \\ & & \boldsymbol{y}_{J''} \text{ unrestricted} \end{array}$$

KKT conditions

Lagrange function: $L(\boldsymbol{x};\boldsymbol{u},\boldsymbol{v}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} u_i g_i(\boldsymbol{x}) + \sum_{j=1}^{l} v_j h_j(\boldsymbol{x})$

$$\text{CQ:} \quad \begin{cases} \sum_{i=1}^{m} \lambda_i \nabla g_i(\bar{\boldsymbol{x}}) + \sum_{j=1}^{l} \mu_j \nabla h_j(\bar{\boldsymbol{x}}) = \boldsymbol{0} \\ \lambda_i g_i(\bar{\boldsymbol{x}}) = 0 \quad \text{for all } i \\ \lambda_i \ge 0 \quad \text{for all } i \end{cases} \implies \begin{cases} \lambda_i = 0 \quad \text{for all } i \\ \mu_j = 0 \quad \text{for all } j. \end{cases}$$

Necessary conditions: $\nabla_{\boldsymbol{x}} L(\bar{\boldsymbol{x}};\boldsymbol{u},\boldsymbol{v}) = \boldsymbol{0}$

$$u_i g_i(\bar{\boldsymbol{x}}) = 0, \quad i = 1,\dots,m$$

$$u_i \ge 0, \qquad v_j \text{ unrestricted}$$

Sufficient conditions: $\bar{\boldsymbol{x}}$ KKT point and

$$\boldsymbol{d}^T \nabla^2_{\boldsymbol{xx}} L(\bar{\boldsymbol{x}};\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})\boldsymbol{d} > 0 \text{ for all } \boldsymbol{d} \ne \boldsymbol{0} \text{ such that}$$

$$\nabla g_i(\bar{\boldsymbol{x}})^T \boldsymbol{d} = 0, \quad i \in I^+,$$

$$\nabla g_i(\bar{\boldsymbol{x}})^T \boldsymbol{d} \le 0, \quad i \in I^0,$$

$$\nabla h_j(\bar{\boldsymbol{x}})^T \boldsymbol{d} = 0, \quad j = 1,\dots,l.$$

Duality

$P: \quad \min f(x) \quad \text{subject to } \ x \in X,\ \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0},\ \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$

Lagrange dual

$D: \quad \max \theta(\boldsymbol{u},\boldsymbol{v}) \quad \text{subject to } \ \boldsymbol{u} \ge \boldsymbol{0}, \qquad \text{where } \theta(\boldsymbol{u},\boldsymbol{v}) = \inf_{x \in X} L(\boldsymbol{x};\boldsymbol{u},\boldsymbol{v})$

$$q(\boldsymbol{x}) = \tfrac{1}{2}\boldsymbol{x}^T \boldsymbol{H}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} \quad \implies \quad \nabla q(\boldsymbol{x}) = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{c}.$$