CS5321
Numerical Optimization

19 Interior-Point Methods for Nonlinear Programming

(Barrier Methods)

Problem formulation

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & c_E(x) = 0 \\ & c_I(x) - s = 0 \\ & s \geq 0 \end{aligned}$$

- c_E is a vector of c_i, $i \in \mathbf{E}$
- c_I is a vector of c_i, $i \in \mathbf{I}$
- s is the slack variables

- Add barrier functions for inequality constraints.

$$\begin{aligned} \min_x \quad & f(x) - \mu \sum_{i=1}^{m} \ln s_i \\ \text{s.t.} \quad & c_E(x) = 0 \\ & c_I(x) - s = 0 \end{aligned}$$

- μ is the barrier parameter, which may be reduced iteratively.

KKT conditions

- The Lagrangian of the original problem is

$$L(x,s,y,z) = f(x) - \mu \sum_{i=1}^{m} \ln s_i - y^T c_E(x) - z^T (c_I(x) - s)$$

  - y and z are the Lagrangian multipliers.
  - Let A_E and A_I denote the Jacobian of c_E and c_I.
- The KKT condition of the barrier function is

$$\begin{aligned} \nabla f(x) - A_E^T(x)y - A_I^T(x)z &= 0 \\ Sz - \mu e &= 0 \\ c_E(x) &= 0 \\ c_I(x) - s &= 0 \end{aligned}$$

Line-search algorithm

- Solve KKT conditions by the Newton's method
  - The primal-dual system

$$\begin{bmatrix} \nabla^2_{xx}L & 0 & -A_E^T & -A_I^T \\ 0 & Z & 0 & S \\ A_E & 0 & 0 & 0 \\ A_I & -I & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_s \\ p_y \\ p_z \end{bmatrix} = - \begin{bmatrix} \nabla f - A_E^T y - A_I^T z \\ Sz - \mu e \\ c_E \\ c_I - s \end{bmatrix}$$

- Update (x, s, y, z)+=$(\alpha_x p_x, \alpha_s p_s, \alpha_y p_y, \alpha_z p_z)$ and μ_{k+1}
  - The fraction to the boundary rule: for $\tau \in (0,1)$

$$\alpha_s = \max\{\alpha \in (0,1] : s + \alpha p_s \geq (1-\tau)s\}$$

$$\alpha_z = \max\{\alpha \in (0,1] : z + \alpha p_z \geq (1-\tau)z\}$$

Solving the primal-dual system

- Let $\Sigma = S^{-1}Z$ and rewrite the primal dual system as

$$\begin{bmatrix} \nabla^2_{xx}L & 0 & A_E^T & A_I^T \\ 0 & \Sigma & 0 & -I \\ A_E & 0 & 0 & 0 \\ A_I & -I & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_s \\ -p_y \\ -p_z \end{bmatrix} = - \begin{bmatrix} \nabla f - A_E^T y - A_I^T z \\ z - \mu S^{-1} e \\ c_E \\ c_I - s \end{bmatrix}$$

  - The matrix become symmetric.

- Eliminate p_s and reform the problem

$$\begin{bmatrix} \nabla^2_{xx}L & A_E^T & A_I^T \\ A_E & 0 & 0 \\ A_I & 0 & -\Sigma^{-1} \end{bmatrix} \begin{bmatrix} p_x \\ -p_y \\ -p_z \end{bmatrix} = - \begin{bmatrix} \nabla f - A_E^T y - A_I^T z \\ c_E \\ c_I - \mu Z^{-1} e \end{bmatrix}$$

Avoid singularity

- Singularity may caused by Σ(some elements goes to ∞) or ill-conditionness of $\nabla_{xx}{}^2 f$ and $A_{\mathbf{I}}$.
  1. Use projected Hessian and modified Σ

$$\begin{bmatrix} G & 0 & A_E^T & A_I^T \\ 0 & T & 0 & -I \\ A_E & 0 & 0 & 0 \\ A_I & -I & 0 & 0 \end{bmatrix}$$

  2. Add diagonal shift $\begin{bmatrix} \nabla_{xx}^2 L + \delta I & 0 & A_E^T & A_I^T \\ 0 & \Sigma & 0 & -I \\ A_E & 0 & -\gamma I & 0 \\ A_I & -I & 0 & 0 \end{bmatrix}$

Barrier param and step length

- Barrier parameters $\{\mu_k\}$ need converge to zero.

1. Static method $\mu_{k+1} = \sigma_k \mu_k$ for $\sigma_k \in (0,1)$
2. Adaptive methods
   a. In the linear programming, $\mu_{k+1} = \sigma \frac{s_k^T z_k}{m}$ (chap 14)

- Use merit function to decide whether a step is productive and should be accepted. (chap 18)

$$\phi(x, s) = f(x) - \mu \sum_{i=1}^{m} \ln s_i + \nu \|c_E\| + \nu \|c_I - s\|$$

Trust-region SQP method

- Two differences from the line search approach
  1. Solve primal (x,s) and dual problem (y, z) alternately
  2. Use scaling S^{-1} on p_s.
- The primal problem

$$
\begin{aligned}
\min_{p_x, p_s} \quad & \nabla f^T p_x + \tfrac{1}{2} p_x^T \nabla_{xx}^2 L p_x - \mu e^T S^{-1} p_s + \tfrac{1}{2} p_s^T \Sigma p_s \\
\text{subject to} \quad & A_E(x) p_x + c_E(x) = r_E \\
& A_I(x) p_x - p_s + (c_I(x) - s) = r_I \\
& \|(p_x, S^{-1} p_s)\| \le \Delta, \\
& p_s \ge -\tau s.
\end{aligned}
$$

Solving the dual problem

- Define $\hat{A} = \begin{bmatrix} A_E & 0 \\ A_I & -S \end{bmatrix}$. The dual problem is to solve

$$\hat{A}^T \begin{bmatrix} y \\ z \end{bmatrix} = \begin{bmatrix} \nabla f(x) \\ -\mu e \end{bmatrix}$$

- Using the least-square method

$$\begin{bmatrix} y \\ z \end{bmatrix} = (\hat{A}\hat{A}^T)^{-1}\hat{A} \begin{bmatrix} \nabla f(x) \\ -\mu e \end{bmatrix}$$

  - The solution cannot guarantee z to be positive. $\Rightarrow$ replaced by a small positive number if $z_i \leq 0$

$$z_i \rightarrow \min(10^{-3}, \mu/s_i)$$

Scaling

- Let $p_s' = S^{-1}p_s$. The problem can be rewritten as

$$\begin{aligned}
\min_{p_x, p_s} \quad & \nabla f^T p_x + \tfrac{1}{2} p_x^T \nabla_{xx}^2 L p_x - \mu e^T p_s' + \tfrac{1}{2} {p_s'}^T S \Sigma S p_s' \\
\text{s.t.} \quad & A_E(x) p_x + c_E(x) = r_E \\
& A_I(x) p_x - S p_s' + (c_I(x) - s) = r_I \\
& \|(p_x, p_s')\| \le \Delta, \\
& p_s' \ge -\tau.
\end{aligned}$$

- Parameter r_E and r_I can be computed by solving

$$r_E = A_E(x) v_x + c_E(x), r_I = A_I(x) v_x - S v_s + c_I(x) - s$$

$$\min_v \|r_E\|_2^2 + \|r_I\|_2^2$$

$$\text{s.t.} \|(v_x, v_s)\| \le 0.8\Delta, v_s \ge V(\tau/2)e$$

Convergence theory

- Theorem 19.1: Let $\{x_k\}$ be the iterative points of the basic algorithm and $\{\mu_k\}\rightarrow 0$. Suppose f and c are continuously differentiable. The all limit points x^* of $\{x_k\}$ are feasible. Moreover, if any x^* satisfies LICQ, the KKT conditions are satisfied.