THE UNIVERSITY OF CHICAGO

19 Interior Point Algorithms

Outline

- Same idea as in the case of the interior-point method for QP.
- Create a path that is interior with respect to the Lagrange multipliers and the slacks that depends on a smoothing parameter mu.
- Drive mu to 0.

Interior –point, “smoothing” parth

- Formulation (with slacks) :

$$\begin{aligned} \min_{x,s} \quad & f(x) \\ \text{subject to} \quad & c_E(x) = 0, \\ & c_I(x) - s = 0, \\ & s \geq 0. \end{aligned}$$

- Interior-point (smoothing path; mu=0: KKT)

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

$$\begin{aligned} c_E(x) &= 0, \\ c_I(x) - s &= 0, \end{aligned}$$

Barrier interpretation

- The nonlinear equation is the same as the KKT point of the barrier function:

$$\begin{aligned} \min_{x,s} \quad & f(x) - \mu \sum_{i=1}^{m} \log s_i \\ \text{subject to} \quad & c_{\mathcal{E}}(x) = 0, \\ & c_{\mathcal{I}}(x) - s = 0, \end{aligned}$$

Newton Method:

- Linearization for fixed mu:

$$\begin{bmatrix} \nabla^2_{xx}\mathcal{L} & 0 & -A_{\mathrm{E}}^T(x) & -A_{\mathrm{I}}^T(x) \\ 0 & Z & 0 & S \\ A_{\mathrm{E}}(x) & 0 & 0 & 0 \\ A_{\mathrm{I}}(x) & -I & 0 & 0 \end{bmatrix} \begin{bmatrix} p_x \\ p_s \\ p_y \\ p_z \end{bmatrix} = - \begin{bmatrix} \nabla f(x) - A_{\mathrm{E}}^T(x)y - A_{\mathrm{I}}^T(x)z \\ Sz - \mu e \\ c_{\mathrm{E}}(x) \\ c_{\mathrm{I}}(x) - s \end{bmatrix},$$

$$\mathcal{L}(x, s, y, z) = f(x) - y^T c_{\mathrm{E}}(x) - z^T(c_{\mathrm{I}}(x) - s).$$

Choose the step (Eqns 19.9)

- The new iteration:

$$x^+ = x + \alpha_s^{\max} p_x, \quad s^+ = s + \alpha_s^{\max} p_s,$$

- Where:

$$y^+ = y + \alpha_z^{\max} p_y, \quad z^+ = z + \alpha_z^{\max} p_z,$$

- And,

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

$$\tau = 0.99 - 0.995$$

How do I measure progress?

- Merit function:

$$E(x, s, y, z; \mu) = \max \{ \|\nabla f(x) - A_E(x)^T y - A_I(x)^T z\|, \|Sz - \mu e\|, \\ \|c_E(x)\|, \|c_I(x) - s\| \},$$

- If small enough, then I adjust mu

Basic Interior-Point Algorithm

Algorithm 19.1 (Basic Interior-Point Algorithm).

Choose x_0 and $s_0 > 0$, and compute initial values for the multipliers y_0 and $z_0 > 0$. Select an initial barrier parameter $\mu_0 > 0$ and parameters $\sigma, \tau \in (0, 1)$. Set $k \leftarrow 0$.

repeat until a stopping test for the nonlinear program (19.1) is satisfied
  repeat until $E(x_k, s_k, y_k, z_k; \mu_k) \le \mu_k$
    Solve (19.6) to obtain the search direction $p = (p_x, p_s, p_y, p_z)$;
    Compute $\alpha_s^{\max}, \alpha_z^{\max}$ using (19.9);
    Compute $(x_{k+1}, s_{k+1}, y_{k+1}, z_{k+1})$ using (19.8);
    Set $\mu_{k+1} \leftarrow \mu_k$ and $k \leftarrow k + 1$;
  end
  Choose $\mu_k \in (0, \sigma \mu_k)$;
end

How to solve the linear system

- Rewriting the Newton Direction:

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

- Can use indefinite factorization LDLT.
- Or, projected CG (since it is in saddle-point form)

$$\Sigma = S^{-1}Z.$$

Linear System, part II

- Or, we can eliminate p_s and use LDLT

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

- And even p_z:

$$\begin{bmatrix} \nabla^2_{xx}\mathcal{L} + A_I^T \Sigma A_I & A_E^T(x) \\ A_E(x) & 0 \end{bmatrix}$$

How do we deal with nonconvexity and non-LICQ?

- Regularization

$$\begin{bmatrix} \nabla^2_{xx}\mathcal{L} + \delta I & 0 & A_E(x)^T & A_I(x)^T \\ 0 & \Sigma & 0 & -I \\ A_E(x) & 0 & -\gamma I & 0 \\ A_I(x) & -I & 0 & 0 \end{bmatrix}.$$

- Choose *delta* so that signature of the matrix corresponds to positive definiteness of reduced matrix:
- For signature, can use LDLT

$$(n+m, l+m, 0)$$

- If there are zero eigenvalues, we increase *gamma* (appendix B describes the procedure)

I can instead use a line search ?

- Backtracking search for merit function (based on barrier interpretation) :

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

$$\alpha_s \in (0, \alpha_s^{\max}], \qquad \alpha_z \in (0, \alpha_z^{\max}],$$

- Directional derivative (for line search)

$$\frac{\partial}{\partial p}\|c(x)\| = \frac{\partial}{\partial p}\sqrt{c(x)^T c(x)} = \begin{cases} \dfrac{c(x)}{\|c(x)\|}\nabla c(x) p & c(x) \neq 0 \\ \dfrac{\nabla c(x) p}{\|\nabla c(x) p\|}\nabla c(x) p & c(x) = 0, \nabla c(x) p \neq 0 \\ 0 & \text{otherwise} \end{cases}$$

How do we update barrier parameter?

- Decrease of barrier (example):

$$\mu_{k+1} = \sigma_k \mu_k, \quad \text{with} \quad \sigma_k \in (0, 1).$$

$$\sigma_k = 0.1 \min\left(0.05 \frac{1 - \xi_k}{\xi_k}, 2\right)^3, \quad \text{where} \quad \xi_k = \frac{\min_i [s_k]_i [z_k]_i}{(s^k)^T z^k / m}.$$

- Step update:

$$\begin{aligned} x^+ &= x + \alpha_s p_x, \quad & s^+ &= s + \alpha_s p_s, \\ y^+ &= y + \alpha_z p_y, \quad & z^+ &= z + \alpha_z p_z. \end{aligned}$$

A practical interior-point algorithm

Algorithm 19.2 (Line Search Interior-Point Algorithm).

Choose x_0 and $s_0 > 0$, and compute initial values for the multipliers y_0 and $z_0 > 0$. If a quasi-Newton approach is used, choose an $n \times n$ symmetric and positive definite initial matrix B_0. Select an initial barrier parameter $\mu > 0$, parameters $\eta, \sigma \in (0, 1)$, and tolerances ϵ_μ and ϵ_{TOL}. Set $k \leftarrow 0$.

repeat until $E(x_k, s_k, y_k, z_k; 0) \le \epsilon_{\text{TOL}}$
    repeat until $E(x_k, s_k, y_k, z_k; \mu) \le \epsilon_\mu$
        Compute the primal-dual direction $p = (p_x, p_s, p_y, p_z)$ from
            (19.12), where the coefficient matrix is modified as in
            (19.25), if necessary;
        Compute $\alpha_s^{\max}, \alpha_z^{\max}$ using (19.9); Set $p_w = (p_x, p_s)$;
        Compute step lengths α_s, α_z satisfying both (19.27) and
         $\phi_\nu(x_k + \alpha_s p_x, s_k + \alpha_s p_s) \le \phi_\nu(x_k, s_k) + \eta \alpha_s D\phi_\nu(x_k, s_k; p_w)$;
        Compute $(x_{k+1}, s_{k+1}, y_{k+1}, z_{k+1})$ using (19.28);
        if a quasi-Newton approach is used
            update the approximation B_k;
        Set $k \leftarrow k + 1$;
    end
    Set $\mu \leftarrow \sigma\mu$ and update ϵ_μ;
end