Lecture 15
Primal-dual interior-point method

- primal-dual central path equations
- infeasible primal-dual method

Optimality conditions

primal and dual problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax + s = b \\ & s \geq 0 \end{array} \qquad\qquad \begin{array}{ll} \text{maximize} & -b^T z \\ \text{subject to} & A^T z + c = 0 \\ & z \geq 0 \end{array}$$

optimality conditions

$$\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}$$

$$s \geq 0, \qquad z \geq 0, \qquad s \circ z = 0$$

$s \circ z$ is component-wise (Hadamard) vector product:

$$s \circ z = (s_1 z_1,\, s_2 z_2,\, \ldots,\, s_m z_m)$$

Central path equations

$$\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}$$

$$s \geq 0, \qquad z \geq 0, \qquad s \circ z = \frac{1}{t}\mathbf{1}$$

- a continuous deformation of the optimality conditions
- solution x, z, s is

$$x = x^\star(t), \qquad s = b - Ax^\star(t), \qquad z = z^\star(t)$$

- $m + n$ linear, m nonlinear equations, and $2m$ simple inequalities

Interpretation of barrier method

- write central path equations as

$$Ax + s = b, \qquad A^T z + c = 0, \qquad z_i - \frac{1}{ts_i} = 0, \quad i = 1, \ldots, m$$

- linearize around strictly feasible $\hat{x}$, $\hat{z}$, $\hat{s}$:

$$A\Delta x + \Delta s = 0, \quad A^T \Delta z = 0, \quad \Delta z_i + \frac{\Delta s_i}{t\hat{s}_i^2} = -\hat{z}_i + \frac{1}{t\hat{s}_i}, \quad i = 1, \ldots, m$$

- eliminating Δs and Δz gives an equation in Δx (with $S = \mathbf{diag}(\hat{s})$):

$$A^T S^{-2} A \Delta x = -tc - A^T S^{-1} \mathbf{1}$$

this is exactly the centering Newton equation $\nabla^2 f_t(\hat{x}) \Delta x = -\nabla f_t(\hat{x})$

Primal-dual path-following methods

- use a different, symmetric linearization of central path
- update primal and dual variables x, z in each iteration
- update central path parameter t after every Newton step
- aggressive step sizes ($e.g.$, 0.99 of maximum step to the boundary)
- allow infeasible iterates
- add second-order terms to linearization of central path

used in most interior-point solvers

Basic primal-dual update

let $\hat{s}$, $\hat{x}$, $\hat{z}$ be the current iterates (with $\hat{s} > 0$, $\hat{z} > 0$)

- compute steps Δs, Δx, Δz by linearizing the central path equation

$$\begin{bmatrix} 0 \\ s \end{bmatrix} = \begin{bmatrix} 0 & A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} c \\ b \end{bmatrix}, \qquad s \circ z = \sigma\mu\,\mathbf{1}$$

  around $\hat{s}$, $\hat{x}$, $\hat{z}$, where $\mu = \hat{s}^T\hat{z}/m$ and $\sigma \in [0,1]$

- make an update

$$(\hat{x}, \hat{s}) := (\hat{x}, \hat{s}) + \alpha_{\mathrm{p}}(\Delta x, \Delta s), \qquad \hat{z} := \hat{z} + \alpha_{\mathrm{d}}\Delta z$$

  that preserves positivity of $\hat{s}$, $\hat{z}$

Linearized central path equation

central path equation (without inequalities)

$$Ax + s = b, \qquad A^T z + c = 0, \qquad s \circ z = \sigma\mu \mathbf{1}$$

linearization around $\hat{x}$, $\hat{s}$, $\hat{z}$

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Z \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} -(A\hat{x} + \hat{s} - b) \\ -(A^T\hat{z} + c) \\ \sigma\mu\mathbf{1} - \hat{s} \circ \hat{z} \end{bmatrix}$$

where $S = \mathbf{diag}(\hat{s})$, $Z = \mathbf{diag}(\hat{z})$

we assume $\hat{s} > 0$, $\hat{z} > 0$, but not $A\hat{x} + \hat{s} = b$ or $A^T\hat{z} + c = 0$

Path-following algorithm

choose starting points $\hat{s}$, $\hat{x}$, $\hat{z}$ with $\hat{s} > 0$, $\hat{z} > 0$

1. **compute residuals and evaluate stopping criteria**

$$r_{\mathrm{p}} = A\hat{x} + \hat{s} - b, \qquad r_{\mathrm{d}} = A^T\hat{z} + c$$

   terminate if r_{p}, r_{d}, and $\hat{s}^T\hat{z}$ are small

2. **compute affine scaling direction:** solve the linear equation

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Z \end{bmatrix} \begin{bmatrix} \Delta z_{\mathrm{a}} \\ \Delta x_{\mathrm{a}} \\ \Delta s_{\mathrm{a}} \end{bmatrix} = \begin{bmatrix} -r_{\mathrm{p}} \\ -r_{\mathrm{d}} \\ -\hat{s} \circ \hat{z} \end{bmatrix}$$

3. **select barrier parameter:** find

$$\begin{array}{rcl}\alpha_{\mathrm{p}} & = & \max\{\alpha \in [0,1] \mid \hat{s} + \alpha \Delta s_{\mathrm{a}} \geq 0\} \\ \alpha_{\mathrm{d}} & = & \max\{\alpha \in [0,1] \mid \hat{z} + \alpha \Delta z_{\mathrm{a}} \geq 0\}\end{array}$$

and take

$$\sigma = \left(\frac{(\hat{s} + \alpha_{\mathrm{p}} \Delta s_{\mathrm{a}})^T (\hat{z} + \alpha_{\mathrm{d}} \Delta z_{\mathrm{a}})}{\hat{s}^T \hat{z}} \right)^{\delta}$$

δ is an algorithm parameter (a typical value is $\delta = 3$)

4. **compute search direction:** solve the linear equation

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Z \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_{\mathrm{p}} \\ -r_{\mathrm{d}} \\ \sigma(\hat{s}^T \hat{z}/m)\mathbf{1} - \hat{s} \circ \hat{z} \end{bmatrix}$$

5. **update iterates:** find maximum steps to the boundary

$$\begin{aligned}\alpha_{\mathrm{p}} &= \max\{\alpha \geq 0 \mid \hat{s} + \alpha \Delta s \geq 0\} \\ \alpha_{\mathrm{d}} &= \max\{\alpha \geq 0 \mid \hat{z} + \alpha \Delta z \geq 0\}\end{aligned}$$

and take

$$\begin{aligned}(\hat{x}, \hat{s}) &:= (\hat{x}, \hat{s}) + \min\{1, 0.99\alpha_{\mathrm{p}}\}(\Delta x, \Delta s) \\ \hat{z} &:= \hat{z} + \min\{1, 0.99\alpha_{\mathrm{d}}\}\Delta z\end{aligned}$$

return to step 1

Example stopping criteria

use tolerances $\epsilon_{\rm feas}$, $\epsilon_{\rm abs}$, $\epsilon_{\rm rel}$ to limit primal, dual residuals and duality gap

primal and dual feasibility: check that iterates satisfy

$$\|r_{\rm p}\| \leq \epsilon_{\rm feas} \max\{1, \|b\|\} \quad \text{and} \quad \|r_{\rm d}\| \leq \epsilon_{\rm feas} \max\{1, \|c\|\}$$

duality gap: check that condition 1 or 2 is satisfied

1. small absolute duality gap: $\hat{s}^T \hat{z} \leq \epsilon_{\rm abs}$
2. small relative duality gap

$$\left(c^T \hat{x} < 0 \quad \text{and} \quad \frac{\hat{s}^T \hat{z}}{-c^T \hat{x}} \leq \epsilon_{\rm rel}\right) \quad \text{or} \quad \left(-b^T \hat{z} > 0 \quad \text{and} \quad \frac{\hat{s}^T \hat{z}}{-b^T \hat{z}} \leq \epsilon_{\rm rel}\right)$$

Interpretation of search directions

affine scaling direction (step 2)

- $(\Delta s_{\rm a}, \Delta x_{\rm a}, \Delta z_{\rm a})$ solves linearized central path equation with $\sigma = 0$
- this is also the solution of the linearized optimality conditions

selection of barrier parameter (step 3)

- take σ small if step in affine scaling direction gives a large gap reduction
- a heuristic, using an estimate of how good the affine scaling direction is

combined search direction (step 4)

- linear equation has same coefficient matrix as equation in step 2
- we can reuse the factorization; hence, extra cost is negligible

Mehrotra correction

replace equation in step 4 by

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Z \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta x \\ \Delta x \end{bmatrix} = \begin{bmatrix} -r_{\mathrm{p}} \\ -r_{\mathrm{d}} \\ \sigma(\hat{s}^T\hat{z}/m)\mathbf{1} - \hat{s}\circ\hat{z} - \Delta s_{\mathrm{a}}\circ\Delta z_{\mathrm{a}} \end{bmatrix}$$

- extra term $\Delta s_{\mathrm{a}} \circ \Delta z_{\mathrm{a}}$ is approximation of the second-order term in

$$(\hat{s} + \Delta s) \circ (\hat{z} + \Delta z) = \sigma\mu\mathbf{1}$$

- adding the correction typically saves a few iterations

Search equations

step 2 and step 4 involve equations of the form

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Z \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} b_z \\ b_x \\ b_s \end{bmatrix}$$

- eliminating $\Delta s = Z^{-1}(b_s - S\Delta z)$ gives

$$\begin{bmatrix} -SZ^{-1} & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} \Delta z \\ \Delta x \end{bmatrix} = \begin{bmatrix} b_z - Z^{-1}b_s \\ b_x \end{bmatrix}$$

- usually solved by eliminating $\Delta z = S^{-1}ZA\Delta x - S^{-1}Zb_z + S^{-1}b_s$

$$A^T S^{-1} Z A \, \Delta x = b_x + A^T S^{-1} Z b_z - A^T S^{-1} b_s$$

Cholesky factorization

definition: every symmetric positive definite B can be factored as

$$B = LL^T$$

- Cholesky factor L is lower triangular with positive diagonal entries
- cost is $n^3/3$ floating-point operations (flops) if B is dense

linear equation with positive definite coefficient

$$Bx = d$$

- factor B as $B = LL^T$ ($n^3/3$)
- solve $Ly = d$ by forward substitution (n^2 flops)
- solve $L^T x = y$ by backward substitution (n^2 flops)

Sparse positive definite equation

algorithm

1. reorder rows and columns of B symmetrically to increase sparsity of L

$$(PBP^T)(Px) = Pd \qquad P \text{ a permutation matrix}$$

2. symbolic factorization: find sparsity pattern of L (from pattern of B)
3. numerical factorization: $PBP^T = LL^T$ (from values of entries of B)
4. use forward and backward substitution to solve $LL^TPx = Pd$

complexity

- most expensive steps are 2 and 3
- only steps 3, 4 depend on numerical values of B
- only step 4 depends on right-hand side d

Linear equations in interior-point method

the algorithm on page 15–8 requires two linear equations with coefficient

$$B = A^T S^{-1} Z A$$

- A is typically large and sparse
- $S^{-1}Z$ is positive diagonal, different at each iteration
- B is positive definite if $\mathrm{rank}(A) = n$
- sparsity pattern of B is pattern of $A^T A$ (independent of $S^{-1}Z$)

solution via sparse Cholesky factorization

- steps 1, 2 (reordering, symbolic factorization) are needed only once
- step 3 (numerical factorization) is needed once per iteration
- step 4 (forward/backward substitution) is repeated twice per iteration