# 6. Interior point methods.

## 6.1 Justification and interpretation

⊛ $\boxed{\text{WITH A few exceptions, section 19.}}$

Problem definition with slacks:

$$\min_{x,s} f(x)$$

$$c_E(x) = 0 \qquad (y)$$

$$c_I(x) - s = 0 \qquad (z)$$

$$s \geq 0$$

DEFINE: $S = \mathrm{diag}(s)$

⊛ Optimality conditions

$$\cancel{\nabla_x f + y^T \nabla_x c_E + z^T}$$

Define $A_E(x) = \nabla_x c_E \in \mathbb{R}^{n_E \times n}$

$A_I(x) = \nabla_x c_I \in \mathbb{R}^{n_I \times n}$

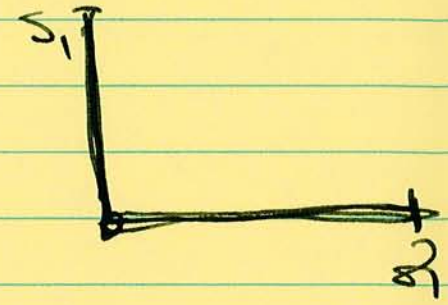
$$\Rightarrow \begin{cases} \nabla_x f - A_E^T \cdot y - A_I^T z = 0 \\ S \cdot z = 0 \\ c_E(x) = 0 \\ c_I(x) - s = 0 \end{cases}$$

complementary condition

$$\boxed{s, z \geq 0}$$

Mihai Anitescu Notes

6.1.2 It is a nonsmooth equation. e.g. one dimensional




How do I regularize it?

$\Rightarrow$ $\boxed{\text{Smooth it}}$

⊗ Approach 1: Homotopy on KKT

$$\begin{cases} \nabla_x f(x(\mu)) - A_E^T(x(\mu)) \cdot y(\mu) - A_I^T(x(\mu))\, z(\mu) = 0 \\ S(\mu) \cdot Z(\mu) = \mu e \\ c_E(x(\mu)) = 0 \\ c_I(x(\mu)) - s(\mu) = 0 \end{cases}$$

If I start positive I do not need to enforce $x > 0$; $s > 0$!!

So; as $\mu \to 0$; $\boxed{x(\mu), s(\mu), y(\mu), z(\mu) \to \text{KKT point}}$

Solve for some $\mu$, update.

Mihai Anitescu Notes

6.1.3 <u>Approach 2</u>: Barrier Methods | primal barrier

$$\min_x \; f(x) - \sum_{i=1}^{n_I} \log(s_i) \cdot \mu$$

$$c_E(x) = 0$$

$$c_I(x) - s = 0$$

→ barrier keeps away from 0

$\mu \downarrow 0$

Write optimality conditions:

$$\nabla_x f - A_E(x)^T y - A_I^T(x) \cdot z = 0$$

$$(\nabla_s) \qquad -\mu S^{-1} e + z = 0$$

$$c_E(x) = 0$$

$$c_I(x) - s = 0$$

ALMOST same as homotopy; however $s$ may go to 0; so inversion is problematic

=> Multiply by $S$ to obtain $\boxed{Sz = \mu e}$

same as before!!

Mihai Anitescu Notes

6.14

$$\begin{tikzcd}
\text{NLP} \arrow[r, leftrightarrow, "\text{barrier}"] \arrow[d, "\text{optimality}"] & \text{barrier prind} \arrow[d, "\text{optimality} + S\cdot(\nabla_S)"] \\
\text{KKT system} \arrow[r, "\text{continuation}"] & \text{homotopy method}
\end{tikzcd}$$

nonsmooth

~~But~~ I prefer to solve nonlinear system attached to homotopy method; since no division by 0!! → But I will use barrier to guide me; for descent

Mihai Anitescu Notes

# 6. Interior point algorithms
## 6.2 Basic concept

- Define $\mathcal{L}(x,y,z) = f(x) - y^T c_E(x) - z^T c_I(x)$

Write Newton's system for the homotopy ~~method~~ method.; BEFORE UPDATING $\mu$.

$$\begin{bmatrix} \nabla^2_{xx}\mathcal{L} & 0 & -A_E^T & -A_I^T \\ 0 & Z & 0 & S \\ A_E & 0 & 0 & 0 \\ A_I & -I & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} p_x \\ p_s \\ p_y \\ p_z \end{bmatrix} = \begin{bmatrix} -\nabla_x \mathcal{L} \\ \mu e - Sz \\ -c_E(x) \\ s - c_I(x) \end{bmatrix}$$

- Update rule

$$x^+ = x + \alpha_1 p_x \qquad y^+ = y + \alpha_2 p_y$$
$$s^+ = s + \alpha_1 p_s \qquad z^+ = z + \alpha_2 p_z$$

- But how large is $s$?

Fraction to the boundary rule:

Mihai Anitescu Notes

6.2.2 $\alpha_s^{max} = \{ \alpha \in (0,1] \mid s + \alpha p_s \geq (1-\tau)\cdot s \} = \alpha_1$

$\alpha_z^{max} = \{ \alpha \in (0,1] \mid z + \alpha p_z \geq (1-\tau)\cdot z \} = \alpha_2$

$\tau \in (0,1)$ ; typically $\tau = 0.995$.

This is the basis of almost all interior point algorithms.

⊗ Define error

$$E(x,s,y,z,\mu) = \left\{ \| \nabla_x \mathcal{L} \|, \| c_E \|, \| c_I - s \|, \| Sz - \mu e \| \right\}$$

⊗ Vanilla algorithm 19.1.

repeat until stop of NLP [e.g. $E(x^k, s^k, y^k, z^k; 0) < \varepsilon$]
  repeat until $E(x^k, s^k, y^k, z^k; \mu^k) \leq \mu^k$
    Solve N; find $(p^x, p^y, p^s, p^z) = p$
    Compute $x^+, y^+, s^+, z^+$
  end $k \leftarrow k+1$ ; $\mu_{k+1} \leftarrow \mu_k$
$\mu_k \leftarrow \sigma \mu_k$ (alternatively, choose some other $\mu^k < \sigma \mu^k$)

6.2.3 Assume (1) repeat exterior ends in a finite # of iterations (2) $f, c$ are twice ctsly diff; (3) any acc point $\hat{x}$ of $x^k$ satisfies LICQ

Then $x^k$ is KKT.

<u>Proof</u> α) $\hat{x}$ is limit point $x^k \mapsto \hat{x}$;

β) $c_I(x^k) - s^k \geq -\mu^k$ (inner loop condition): take limit $\Rightarrow c(\hat{x}) \geq 0$.

γ) $c_E(x^k) = 0 \Rightarrow c_E(\hat{x}) = 0$

δ) $\nabla f(x^k) - A_I(x^k)^T y^k - A_c(x^k)^T \cdot z^k = \rho^k$

$\|\rho^k\| \leq \mu^k$

$$\boxed{\nabla f(\hat{x}) - A_I(\hat{x}) y^k - A_c(\hat{x})^T \cdot z^k = \rho^k +}$$

$\Rightarrow \|y^k; z^k\| \underset{\text{LICQ}}{\leq} \frac{1}{\sigma_0} \cdot \|\rho^k + \nabla f(x^k)\| \Rightarrow$ acc. point

$\hat{z} \geq 0$; $S^k Z^k - \mu^k e + \varepsilon^k \Rightarrow \hat{s} \cdot \hat{z} = 0$

$\Rightarrow$ take limit $\Rightarrow \nabla \mathcal{L}(\hat{x}, \hat{y}, \hat{z}) = 0$ QED.

6.2.4 ⊛ Properties of nonlinear system :

Say $z, s$ satisfies strict complementarity

$\Rightarrow$ $[0, z, 0, s]$ is full row rank (bounded) !! uniformly

matrix is actually invertible

In barrier

$[0, \mu^{-2}, 0, I]$ $\Rightarrow$ goes to $\infty$, AND does not depend on $z$ !!

⊛ Problems with vanilla algorithm.

It is a Newton method, cannot expect to globally converge.

Mihai Anitescu Notes

# 6. Interior point
## 6.3 Algorithmic development

change sign of $p_y$; $p_z$

"primal-dual" system

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

$\Sigma p_s + p_z = \mu S^{-1} e - z$

$\Sigma = S^{-1} Z$ : ~~use symmetric algorithm.~~

~~(GMR; GMRES; LSQR); or Bunch Kaufman; symmetric indefinite factorization~~

$p_s + \Sigma^{-1} p_z = \mu Z^{-1} e - s$

Solving the system: step 1: eliminate 3rd equation

$$\begin{bmatrix} \nabla_{xx}^2 \mathcal{L} & A_E & A_I^T \\ A_E & 0 & 0 \\ A_I & 0 & -\Sigma^{-1} \end{bmatrix} \begin{bmatrix} p_x \\ -p_y \\ -p_z \end{bmatrix} = \begin{bmatrix} -\nabla_x \mathcal{L} \\ -c_E \\ -c_I + \cancel{S(\mu S^{-1} e - z)} + \mu Z^{-1} e - s \end{bmatrix}$$

6.3.2 One more reduction

$$\begin{bmatrix} \nabla_{xx}^2 \mathcal{L} + A_I^T \Sigma A_I & A_E^T \\ A_E & 0 \end{bmatrix} \begin{bmatrix} p_x \\ -p_g \end{bmatrix} = \cdots$$

use QMR; GMRES; LSQR

(1) Update barrier (Fiacco-McCormick)

$\mu_{k+1} = \sigma \mu_k \quad \sigma \in (0,1) \quad (= 0.2 : -1)$

But for superlinear convergence, need $\sigma_k \to 0$.

$\boxed{\text{LOQO.}}$

$$\sigma_k = 0.1 \min\left(0.05 \frac{1-\xi^k}{\xi^k}, 2\right)^3$$

where $\xi_k = \dfrac{\min_i s_i^k \cdot z_i^k}{s^{kT} \cdot z^k / m}$

If close to "balanced complem" aggressively decrease

## 6.3.3 Handling nonconvexity and singularity

Let $\begin{bmatrix} \nabla_{xx}^2 \mathcal{L} & 0 \\ 0 & \Sigma \end{bmatrix}$ be p.d. on $\begin{bmatrix} A_E(x) & 0 \\ A_I(x) & -I \end{bmatrix} \begin{bmatrix} p_x \\ p_z \end{bmatrix} = 0$

$\Rightarrow$ then direction is a descent direction for $E_i$ or $\Phi_\nu$. $\Rightarrow$

~~Theorem (Gould)~~

How do I make it P.D? If I add $\delta I$ on diagonal; then ask $\nabla_{xx}^2 + \delta I$ be pd

$\Rightarrow$ may lose Newton (since $\nabla_{xx}^2 \mathcal{L}$ is not PD)

Mihai Anitescu Notes

6.3.4 Theorem (GOULD). Assume $A$ is full row rank $\Rightarrow$

$$\begin{array}{c} n \\ m \end{array}\begin{bmatrix} Q & A^T \\ A^T & 0 \end{bmatrix}$$ has inertia $(n, m, 0) \Leftrightarrow Q$ is pd on $\ker(A)$.

Therefore; ~~need to make sure~~ safeguarding

$$\begin{bmatrix} \nabla_{xx} + \delta I & 0 & A_E & A_I \\ 0 & \Sigma & 0 & -I \\ A_E & 0 & -\gamma I & 0 \\ A_I & -I & 0 & 0 \end{bmatrix}$$

$\gamma, \delta \to 0$ as $I$ converge to capture Newton, but stay stable

But how do I make sure I have smallest perturbation $\delta$ which satisfies inertia??

Minai Anitescu Notes

6.3.6 ~~[illegible]~~

Bunch-Parlett-Kaufman

$PAP^T = LDL^T$ where:

$L$: unit triangular

$P$: permutation matrix

$D$ is block diagonal either $1 \times 1$ or $2 \times 2$ diagonal blocks

$2 \times 2$ diagonal blocks help with fact that you may not have a nonzero pivot

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

implemented in Matlab by LDL

Theorem (Sylvester): $A$ is symmetric; $W$ is nonsingular $\Rightarrow W^T A W$ has same inertia as $A$

$(n_+(A), n_-(A), n_0(A))$.

Michael Heath Notes

6.3.7 Consequence:

$L^{-1}PA(LP)^{-T} = D \Rightarrow$ D has same inertia as A

inertia of D is easy to compute.

(1) Moreover; once I have $LDL^T$; P linear system is easy to solve:

one backsolve; one forward solve; one block diagonal inversion

$\boxed{\gamma=0;\ \delta=0}$

(2) idea: do $LDL^T = PAP^T$

if inertia not OK: $\lambda_0(A) > 0 \Rightarrow$

add a $\gamma = 10^{-8}\,\eta\,\mu^{\beta}$ ($\eta \simeq 1$; $\beta = 0.5$; $>0$ $\in (0,1)$)

if $\delta_{old} = 0$

$\delta \leftarrow 10^{-4}$

else $\delta = \delta_{old}/2$

6.3.8. repeat

factorize

if not correct inertia

$\delta = \delta \cdot 10$

$\delta_{old} = \delta$

~~end~~

else

[ compute PD step w. LDL\; step ]

~~while~~ until success

Mihai Anitescu Notes

# 6. Interior point
## 6.4 Merit function and line search

$\cancel{\Phi_\nu(x}$

Use $\Phi_\nu(x,s) = f(x) - \mu \sum_{i=1}^{n_I} \log(s_i) + \nu \|c_E(x)\|_1 + \nu \cdot \|c_I(x) - s\|_1$

Result: if $\nu \geq \|\lambda^*\|_\infty \Rightarrow$ minimum point is unconstrained minimum of this function

Idea: choose $\boxed{\nu > \|\lambda^k\|_\infty + \eta}$ safety factor (say 10)

if you converge; pick up right point

$\nu^{k+1} = \max\{\nu^k; \|\lambda^k\|_\infty + \eta\}$ $\eta$ safety factor

At one iteration; choose $\alpha_s$; s.t

$$\min_{\alpha} \Phi_\nu(x + \alpha p_x, s + \alpha p_s)$$
$$\alpha \in (0, \alpha_s^{max}]$$

$\alpha_z$ by fraction to boundary

Mihai Anitescu Notes

# 6 Interior points
## 6.4 Line search algorithm

6.4.2

Solve $\min_{y,z} \|\nabla f - A_E^T(x) y - A_I^T z\|$ ;

$z = \max\{z, 0\} \Rightarrow y^0; z^0 + \mu e = z^0$

Choose $\mu^0 > 0$ (10) ; $x^0; s^0 > 0$ ($s^0 = \|c_I^+\|_{1,r}$)
$s = s^0 + \mu e$)

$y^0; z^0; \eta; \sigma \in (0,1)$ ; $\varepsilon_{TOL}$ ; $\varepsilon_\mu$ ($= \mu$?)

repeat until $E(x^k, s^k, y^k, z^k; 0) \le \varepsilon_{TOL}$

repeat until $E(x^k, s^k, y^k, z^k, \mu^k) \le \varepsilon^k$

compute $p = (p_x, p_s, p_y, p_z)$ with modified method

Compute $\alpha_s^{max}$ ; $\alpha_z^{max}$ ; $p_w^* = (p_x^*, p_s)$

Compute $\alpha_s; \alpha_z$ satisfying FBR

$\Phi_\nu(x^k + \alpha_s p_x^k, s^k + \alpha_s p_s^k) \le \Phi_\nu(x^k, s^k) + \eta \alpha_s D\Phi_\nu(x_k, s_k; p_w)$

Compute $(x^{k+1}, s^{k+1}, y^{k+1}, z^{k+1})$

$\boxed{\text{Update } \nu}$

$k \leftarrow k+1$

end

$\mu^{k+1} = \sigma \mu^k$

end.

6.43 Compute directional derivative by divided differences :-) (cause it is lot smooth)

takes layer to explain)

Mihai Anitescu Notes