Convergence of Inexact Primal–Dual Methods for Linear Optimization

Venansius Baryamureeba Makerere University, Uganda

Trond Steihaug University of Bergen

Outline

- LP - for simplicity
- KKT and perturbed KKT
- Review of Inexact Newton
- Inexact or perturbed solution
- Inexact ideas and IP

ISMP 2000

LP for simplicity

$$\text{(P)}\quad \begin{array}{ll} \min & c^Tx \\ \text{s.t.} & Ax = b \\ & x \geq 0 \end{array} \qquad \text{(D)}\quad \begin{array}{ll} \max & b^Ty \\ \text{s.t.} & A^Ty + z = c \\ & z \geq 0 \end{array}$$

The optimality conditions (P) and (D) or
The Karush-Kuhn-Tucker (KKT) conditions:

$$F(x,y,z) \equiv \begin{pmatrix} Ax - b \\ A^Ty + z - c \\ XZe \end{pmatrix} = 0, \quad x \geq 0,\ z \geq 0,$$

where $X = \operatorname{diag}(x)$, $Z = \operatorname{diag}(z)$ and e is the vector of all ones.

The perturbed KKT conditions:

$$F_\mu(x,y,z) \equiv F(x,y,z) - \mu \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix} = 0 \quad x \geq 0,\ z \geq 0.$$

Newton's Method

Consider a function $F : \Re^n \to \Re^n$ so that

- $F(v^*) = 0$
- F is continuously differentiable in a neighborhood of v^*.
- $F'(v^*)$ nonsingular

Algorithm 1
Given initial $v^{(0)}$
For $k = 0$ **step** 1 **until** *convergence* **do**
Solve for $\Delta v^{(k)}$

$$F'(v^{(k)})\Delta v^{(k)} + F(v^{(k)}) = 0$$

Find steplength $\alpha^{(k)}$.
Set $v^{(k+1)} = v^{(k)} + \alpha^{(k)}\Delta v^{(k)}$.

For $\alpha^* \leq \alpha^{(k)} \leq 1$ then $v^{(k)} \to v^*$

Inexact Newton

Algorithm 2
Given initial $v^{(0)}$ *and* $\eta^{(0)} < 1$.
For $k = 0$ **step** 1 **until** *convergence* **do**
Let $\eta^{(k)} \leq \eta^{(0)}$
Find some $\Delta v^{(k)}$ *that satisfies*

$$\|F'(v^{(k)})\Delta v^{(k)} + F(v^{(k)})\| \leq \eta^{(k)}\|F(v^{(k)})\|.$$

Find $\alpha^{(k)}$.
Set $v^{(k+1)} = v^{(k)} + \alpha^{(k)}\Delta v^{(k)}$.

For $\alpha^* \leq \alpha^{(k)} \leq 1$ then $v^{(k)} \to v^*$

Let $r^{(k)} = F'(v^{(k)})\Delta v^{(k)} + F(v^{(k)})$ then

$$\frac{\|r^{(k)}\|}{\|F(v^{(k)})\|} \leq \eta^{(k)}.$$

IP and Inexact Computation

Recall

$$F(x,y,z) \equiv \begin{pmatrix} Ax-b \\ A^Ty+z-c \\ XZe \end{pmatrix} = 0, \quad x \geq 0,\ z \geq 0,$$

$$F_\mu(x,y,z) \equiv F(x,y,z) - \mu \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix} = 0 \quad x \geq 0,\ z \geq 0.$$

Observe $F'(x,y,z) = F'_\mu(x,y,z)$ and Inexact Newton with F_μ with residual r is Inexact Newton with F with residual

$$r_g \equiv \mu \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix} + r$$

Idea:
Work with F_μ but measure progress in F.

IP and Inexact computation

Observe

- $\|F\|_1 \geq x^Tz$ (with $=$ if (x,y,z) is feasible)
- Classical choice $\mu = \beta_1 \frac{(x^k)^T z^k}{n}$
- Choose $\|r\|_1 \leq \eta (x^k)^T z^k$
- $\eta + \beta_1 < 1$

then

$$\frac{\|r_g\|_1}{\|F\|_1} \leq \frac{\|\mu \begin{pmatrix} 0 \\ 0 \\ e \end{pmatrix} + r\|_1}{x^Tz} \leq \beta_1 + \eta < 1.$$

Conjecture:
Under reasonable conditions a PD interior point method with approximate solution of the Inexact Newton equation based on the perturbed KKT condition $\|r\|_1 \leq \eta (x^k)^T z^k$ where $\eta < 1$ will convergence.

Augmented or Normal Equations

$$\begin{pmatrix} \xi^k \\ \zeta^k \\ \phi^k \end{pmatrix} \equiv \begin{pmatrix} b - Ax^k \\ c - A^Ty^k - z^k \\ \mu^k e - X^k Z^k e \end{pmatrix} = -F_{\mu^k}(x^k, y^k, z^k)$$

$$r^k = \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \\ \tilde{r}^k \end{pmatrix}$$

Recall (Inexact) Newton equation

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z^k & 0 & X^k \end{bmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta z^k \end{pmatrix} = \begin{pmatrix} \xi^k \\ \zeta^k \\ \phi^k \end{pmatrix} + \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \\ \tilde{r}^k \end{pmatrix}.$$

- Augmented system: Eliminating Δz^k
$$r^k = \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \\ 0 \end{pmatrix}$$
- Normal equations: Eliminating Δx^k and Δz^k
$$r^k = \begin{pmatrix} \bar{r}^k \\ 0 \\ 0 \end{pmatrix}$$

Review (Local)

- Ito–Kelley–Sachs $\|\bar{r}^k\| \leq \eta^k \|\xi^k\|$ for $\eta^k < 1$
- Portugal–Resende–Veiga–Júdice
$\|\bar{r}^k\| \leq \beta_0 \|\xi^k\|$ $\beta_0 < \beta_1$
- Portugal–Fernandes–Júdice
$$\| \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \end{pmatrix} \| \leq \beta_0 \| \begin{pmatrix} \xi^k \\ \zeta^k \end{pmatrix} \|, \quad \beta_0 < \beta_1$$

Tight error control when the iterates are almost primal and/or dual feasible. However, the error bounds can easily be computed

- Freund–Jarre–Mizuno:
$$\|\bar{r}^k\| \leq \varepsilon_p^k \text{ and } \|\hat{r}^k\| \leq \varepsilon_d^k$$
ε_p^k and ε_d^k are not known a priori and go to zero faster than $\|\xi^k\|$ and $\|\zeta^k\|$ respectively and thus requiring high accuracy close to a solution.
- Mizuno–Jarre control the error $\|(\bar{r}^k, \hat{r}^k)\|$ using a semi norm. This norm is not (easily) computable.
- Bellavia and Bellavia–Macconi $\|r\|_2 \leq \eta^k (x^k)^T z^k$ and $\beta_1 + \eta^k \sqrt{n} < 1$

Issues

- Global convergence require linesearch and the globalization technique may/will restrict the tolerance in the inexact solve.
- The results may be norm dependent. The norm $\|\cdot\|$ can be any l_p norm or any norm in appropriate spaces that satisfies
$$\max\{\|\bar{r}^k\|, \|\hat{r}^k\|\} \leq \|(\bar{r}^k, \hat{r}^k)\|.$$

Globalization

$0 < \gamma < 1$, $\gamma_p > 0$, $\gamma_d > 0$.

$$\begin{aligned}
\mathcal{N} = \{(x, y, z) \in \Re^{n+m+n} &: x > 0, z > 0 \\
x_j z_j &\geq \gamma x^T z/n \text{ for } j = 1, 2, \ldots, n, \\
x^T z &\geq \gamma_p \|Ax - b\| \text{ or } \|Ax - b\| \leq \epsilon_p, \\
x^T z &\geq \gamma_d \|A^T y + z - c\| \text{ or } \|A^T y + z - c\| \leq \epsilon_d\}
\end{aligned}$$

Termination: Either

$$\begin{aligned}
(x^k)^T z^k &\leq \epsilon, \\
\|Ax^k - b\| &\leq \epsilon_p, \qquad \text{or} \qquad (x^k)^T z^k > \omega^* \\
\|A^T y^k + z^k - c\| &\leq \epsilon_d
\end{aligned}$$

where ω^* is large. Define the neighborhood

$$\mathcal{N}^* = \{(x, y, z) \in \mathcal{N} : \epsilon^* \leq x^T z \leq \omega^*\}$$

where $\epsilon^* = \min\{\epsilon, \gamma_p \epsilon_p, \gamma_d \epsilon_d\}$.

Each iterate $(x^k, y^k, z^k) \in \mathcal{N}^*$ or satisfies the *termination criteria*

Kojima–Megiddo–Mizuno Algorithm

Algorithm 3

Step 1. *Set* $k = 0$.
$(x^0, y^0, z^0) \in \mathcal{N}$ *such that* $(x^0)^T z^0 \leq \omega^*$.

Step 2. *Check termination*

Step 3. *Let* $\mu^k = \beta_1 (x^k)^T z^k / n$.

Step 4. *Compute inexact solution* $(\Delta x^k, \Delta y^k, \Delta z^k)$

Step 5. *Let* $0 < \bar{\alpha}^k < 1$ *such that*

$$(x^k, y^k, z^k) + \alpha(\Delta x^k, \Delta y^k, \Delta z^k) \in \mathcal{N}$$
$$(x^k + \alpha \Delta x^k)^T (z^k + \alpha \Delta z^k) \leq (1 - \alpha(1 - \beta_2))(x^k)^T z^k,$$

hold for every $\alpha \in (0, \bar{\alpha}^k]$.

Step 6. *Choose* α_p^k, α_d^k *in* $[\bar{\alpha}^k, 1]$,

$$\begin{pmatrix} x^{k+1} \\ y^{k+1} \\ z^{k+1} \end{pmatrix} = \begin{pmatrix} x^k \\ y^k \\ z^k \end{pmatrix} + \begin{pmatrix} \alpha_p^k \Delta x^k \\ \alpha_d^k \Delta y^k \\ \alpha_d^k \Delta z^k \end{pmatrix}$$

such that

$$\begin{aligned}
(x^{k+1}, y^{k+1}, z^{k+1}) &\in \mathcal{N}, \\
(x^{k+1})^T z^{k+1} &\leq (1 - \bar{\alpha}^k (1 - \beta_3))(x^k)^T z^k.
\end{aligned}$$

Step 7. *Increase k by 1. Go to Step 2.*

Results (1)

Let

$$0 \leq \eta_{\max} < \min\left\{\frac{\beta_1}{\max\{\gamma_p, \gamma_d\}}, 1\right\}.$$

The neighborhood $\mathcal{N}$ is made large by making γ_p and γ_d small which implies (in most cases)

$$0 \leq \eta_{\max} < 1.$$

The inexact Newton direction $(\Delta x^k, \Delta y^k, \Delta z^k)$:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z^k & 0 & X^k \end{bmatrix} \begin{pmatrix} \Delta x^k \\ \Delta y^k \\ \Delta z^k \end{pmatrix} = \begin{pmatrix} \xi^k \\ \zeta^k \\ \phi^k \end{pmatrix} + \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \\ 0 \end{pmatrix}$$

The residual satisfies

$$\|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k (x^k)^T z^k$$

for

$$0 \leq \eta^k \leq \eta_{\max}.$$

Results (2)

Theorem 1 *Define*

$$\psi^k = \max\left\{\|Ax^k - b\|, \|A^T y^k + z^k - c\|, (x^k)^T z^k\right\}.$$

Let the norm of the residual $(\bar{r}^k, \hat{r}^k)$ *satisfy*

$$\|(\bar{r}^k, \hat{r}^k)\| \leq \eta^k (x^k)^T z^k, \text{ for } 0 \leq \eta^k \leq \eta_{\max}$$

Then there exists $\tilde{\alpha} \in (0,1)$ *so that*

$$\psi^{k+1} \leq (1-\tilde{\alpha})\psi^k.$$

For $u \in \Re^m$, $v \in \Re^n$, and $w \in \Re^n$ then

$$|||(u,v,w)||| = \max\{\|u\|, \|v\|, \|w\|_1\}$$

is a norm in $\Re^{m+n+n}$. So

$$|||F(x^{k+1}, y^{k+1}, z^{k+1})||| \leq (1-\tilde{\alpha})|||F(x^k, y^k, z^k)|||$$

Results (3)

Classical stopping criterion

$$\frac{\|Ax^k - b\|}{\max\{1, \|b\|\}} + \frac{\|A^T y^k + z^k - c\|}{\max\{1, \|c\|\}} + \frac{(x^k)^T z^k}{\max\{1, \|b\|\ \|c\|\}} \leq \varepsilon$$

Define the norm in $\Re^{m+n+n}$

$$|||(u,v,w)|||_+ = \lambda_1\|u\| + \lambda_2\|v\| + \lambda_3\|w\|_1.$$

where $\lambda_i > 0$ $\sum \lambda_i = 1$ $u \in \Re^m$, $v \in \Re^n$, and $w \in \Re^n$.

With appropriate choices of λ_i and ε_+ the stopping criterion is equivalent to

$$|||F(x^k, y^k, z^k)|||_+ \leq \varepsilon_+$$

and the contraction is

$$|||F(x^{k+1}, y^{k+1}, z^{k+1})|||_+ \leq (1-\tilde{\alpha})|||F(x^k, y^k, z^k)|||_+$$

Inexact Newton on KKT (1)

Theorem 2 *Choose* $\gamma_p = \gamma_d \,(= \bar{\gamma})$ *and*

$$\eta_{\max} < \min\{\bar{\gamma}, \beta_1, \beta_1/\bar{\gamma}\}.$$

Assume that

$$|||(\bar{r}^k, \hat{r}^k, 0)||| \leq \eta^k\ |||F(x^k, y^k, z^k)||| \text{ where } \eta^k \leq \eta_{\max}$$

then there exists $\tilde{\alpha} \in (0,1)$ *so that*

$$|||F(x^{k+1}, y^{k+1}, z^{k+1})||| \leq (1-\tilde{\alpha})|||F(x^k, y^k, z^k)|||.$$

Still open: In some norm

$$\frac{\|r^k\|}{\|F(x^k, y^k, z^k)\|} \leq \eta^k \leq \eta_{\max} < 1 \Rightarrow \text{ Convergence}$$

or

$$\frac{\|(\bar{r}^k, \hat{r}^k, 0)\|}{\|F(x^k, y^k, z^k)\|} \leq \eta^k \leq \eta_{\max} < 1 \Rightarrow \text{ Convergence}$$

(It will depend on the globalization strategy)

Inexact Newton on KKT (2)

Theorem 3 *Let*

$$\lambda = \frac{\lambda_1}{\gamma_p} + \frac{\lambda_2}{\gamma_d} + \lambda_3.$$

Choose

$$\eta_{\max} < \min\{\beta_1/\max\{\gamma_p, \gamma_d\}, 1\}/\lambda.$$

Assume that

$$|||(\bar{r}^k, \hat{r}^k, 0)|||_+ \leq \eta^k\ |||F(x^k, y^k, z^k)|||_+ \text{ where } \eta^k \leq \eta_{\max}$$

then there exists $\tilde{\alpha} \in (0,1)$ *so that*

$$|||F(x^{k+1}, y^{k+1}, z^{k+1})|||_+ \leq (1-\tilde{\alpha})|||F(x^k, y^k, z^k)|||_+.$$

Extension

- Complementarity Problems
- Quadratic Problems
- Semidefinite
- Combined with other globalization techniques

Appendix (1)

Define $G^k = (Z^k)^{-1} X^k$. Eliminating Δz^k leads to an indefinite augmented system

$$\begin{bmatrix} O & A \\ A^T & -(G^k)^{-1} \end{bmatrix} \begin{pmatrix} \Delta y^k \\ \Delta x^k \end{pmatrix} = \begin{pmatrix} \xi^k \\ \zeta^k - (X^k)^{-1} \phi^k \end{pmatrix}$$

and

$$\Delta z^k = (X^k)^{-1} (\phi^k - Z^k \Delta x^k)$$

An inexact solution satisfies

$$\begin{bmatrix} O & A \\ A^T & -(G^k)^{-1} \end{bmatrix} \begin{pmatrix} \Delta y^k \\ \Delta x^k \end{pmatrix} = \begin{pmatrix} \xi^k \\ \zeta^k + (X^k)^{-1} \phi^k \end{pmatrix} + \begin{pmatrix} \bar{r}^k \\ \hat{r}^k \end{pmatrix}.$$

Update Δz^k

$$\Delta z^k = (X^k)^{-1} (\mu^k e - Z^k \Delta x^k) - z^k.$$

The step $(\Delta y^k, \Delta x^k, \Delta z^k)$ is an inexact Newton step of the perturbed KKT with $r^k = (\bar{r}^k, \hat{r}^k, 0)$.

Appendix (2)

Eliminating Δx^k and Δz^k give

$$AG^k A^T \Delta y^k = AG^k (\zeta^k - (X^k)^{-1} \phi^k) + \xi^k.$$

and

$$\begin{aligned} \Delta x^k &= -G^k (\zeta^k - (X^k)^{-1} \phi^k - A^T \Delta y^k) \\ \Delta z^k &= (X^k)^{-1} (\phi^k - Z^k \Delta x^k) \end{aligned}$$

Compute approximate solution Δy^k

$$AG^k A^T \Delta y^k \ = \ AG^k (\zeta^k - (X^k)^{-1} \phi^k) + \xi^k + \bar{r}^k.$$

Update Δx^k and Δz^k

$$\begin{aligned} \Delta x^k &= -G^k (c - A^T (y^k + \Delta y^k) - \mu^k (X^k)^{-1} e) \\ \Delta z^k &= (X^k)^{-1} (\mu^k e - Z^k \Delta x^k) - z^k. \end{aligned}$$

The normal equation approach is equivalent to the augmented system when $\hat{r}^k = 0$.

The step $(\Delta y^k, \Delta x^k, \Delta z^k)$ is an inexact Newton step of the perturbed KKT with $r^k = (\bar{r}^k, 0, 0)$.

More details

Visit my homepage:

`http://www.ii.uib.no/~trond/`

and follow links to

`~trond/talks`

`~trond/publications/tech_reports/`

Bibliography

[1] V. Baryamureeba, and T. Steihaug, *On the Convergence of an Inexact primal–dual interior point method for linear programming*, Technical report number 188, Department of Informatics, University of Bergen, Norway, 2000.

[2] S. Bellavia, *Inexact interior point method,* Journal of Optimization Theory and Applications, Vol. 96, pp. 109–121, 1998.

[3] S. Bellavia, *Numerical performance of an inexact interior point method,* R. De Leone, A. Murli, P.M. Pardalos, and G. Toraldo (Eds), High Performance Algorithms and Software in Nonlinear Optimization, Kluwer Academic Publishers, 1998, pp. 43–51.

[4] S. Bellavia, and M. Macconi, *An Inexact interior point method for monotone NCP,* Optimization Meth. & Soft, Vol. 11-12, pp. 211–241,1999.

[5] J.F. Bonnans, C. Pola, and R. Rébai, *Perturbed path following predictor–corrector interior point algorithms,* Optimization Meth. & Soft, Vol. 11-12, pp. 183–210,1999.

[6] R.S. Dembo, S.C. Eisenstat, and T. Steihaug, *Inexact Newton Methods,* SIAM J. Numer. Anal., Vol. 19, pp. 400–408, 1982. bitemkn:freund97 R.W. Freund, F. Jarre, and S. Mizuno, *Convergence of a class of inexact interior point algorithms for linear programs,* Mathematics of Operations Research, Vol. 24, pp. 50–71, 1999.

[7] R.W. Freund, F. Jarre, and S. Mizuno, *Convergence of a class of inexact interior point algorithms for linear programs,* Mathematics of Operations Research, Vol. 24, pp. 50–71, 1999.

[8] S. Ito *Inexact implementation of interior point algorithms for optimal control problems*, Lecture Notes in Num. Appl. Anal., Vol 14, pp. 245–248, 1995.

[9] S. Ito, C.T. Kelley, and E.W. Sachs, *Inexact primal-dual interior-point iteration for linear programs in functional spaces*, Computational Optimization and Applications, Vol. 4, pp. 189–201, 1995.

[10] M. Kojima, N. Megiddo, and S. Mizuno, *A primal–dual infeasible–interior point algorithm for linear programming,* Mathematical Programming, Vol. 61, pp. 263–280, 1993.

[11] S. Mizuno and F. Jarre, *Global and polynomial-time convergence of an infeasible–interior–point algorithm using inexact computation*, Mathematical Programming, Vol. 84, pp. 105–122, 1999.

[12] L. Portugal, L. Fernandes, and J. Júdice, *A truncated Newton interior point algorithm for the solution of a multicommodity spatial equilibrium model,* Complementarity and Variational Problems, edited by J. S. Pang and M. Ferris, SIAM, 1997, pp. 315–344.

[13] L.F. Portugal, M.G.C. Resende, G. Veiga, and J. Júdice, *A truncated primal–infeasible dual–feasible interior point network flow method,* Networks, Vol. 35, pp. 91–108, 2000.