The Zoom strategy for accelerating and warm-starting interior methods

PhD Thesis

IMPA, Rio de Janeiro, July 21, 2008

Leo Tenenblat

IMPA
Rio de Janeiro
Brazil
leo@sferi.com

Acknowledgements

Thanks to:
Alfredo, Michael, the defense committee, IMPA, CAPES,
Keti, Moyses, and extended family

This thesis is dedicated to:
Ingrid & Milla

Abstract

Interior methods using iterative solvers for each search direction can require drastically increasing work per iteration as higher accuracy is sought.

The Zoom strategy solves first to low accuracy, and then solves for a correction to both primal and dual variables, again to low accuracy. We "zoom in" on the correction by scaling it up, thus permitting a cold start for the correction.

The same strategy applies to warm-starting in general.

Outline

The Problems

CP

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \phi(x) \\ \text{subject to} & c_i(x) \geq 0, i = 1, \cdots, m \end{array}$$

$\phi(x), c_i \in C^2$, convex

NP

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \phi(x) \\ \text{subject to} & Ax = b, \quad \ell \leq x \leq u \end{array}$$

$\phi(x)$ convex, separable

LP-Primal

$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^T x \\ \text{subject to} & Ax = b, \quad x \geq 0 \end{array}$$

$A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$

Basic Elements (cont.)

Duality

LP-Dual	$$\begin{aligned} \max_{y \in \mathbb{R}^m} \quad & b^T y \\ \text{subject to} \quad & A^T y \leq c. \end{aligned}$$

Weak duality:

$$c^T x \geq b^T y$$

Strong duality:

$$c^T x = b^T y \iff x \text{ and } y \text{ are optimal}$$

Basic Elements (cont.)

KKT

x^* is a KKT point for (CP) if there exists an m-vector λ^*, such that

(i) $c_i(x^*) \geq 0$ (feasibility)

(ii) $g(x^*) = J(x^*)^T \lambda^*$ (optimality)

(iii) $\lambda^* \geq 0$

(iv) $c_i(x^*)\lambda_i^* = 0, \ i = 1, \ldots, m$ (complementarity)

- No duality gap $\longrightarrow$ Lagrange multipliers are the dual solutions
- Non-zero duality gap $\longrightarrow$ no Lagrange multipliers exist

Basic Elements (cont.)

Constraint Qualifications

Constraint qualifications ensure linear approximation at a point captures the essential geometric information of the true feasible set in a neighborhood

CQs are therefore required to ensure a KKT point is a local minimizer

- LICQ: linear problem or $J_A(\bar{x})$ has full row rank
- MFCQ: $c_i(\bar{x}) > 0, \ \ \forall\ i$ or there exists $p; J_A(\bar{x})p > 0$
- SLCQ: $c_i(\bar{x}) > 0, \ \ \forall\ i$

Basic Elements (cont.)

Necessary Optimality Conditions

First order
If x^* is a local minimizer of (CP) at which the MFCQ holds then x^* must be a KKT point

Second order
If x^* is a local constrained minimizer of (CP) at which the LICQ holds. Then there exists $\lambda^*; \lambda^* \geq 0,\ c^{*T}\lambda^* = 0$, $g^* = J^{*T}\lambda^*$, and

$$p^T H(x^*, \lambda^*) p \geq 0 \text{ for all } p \text{ satisfying } J_A^* p = 0$$

Basic Elements (cont.)

Sufficient Optimality Conditions

Sufficient conditions for an isolated constrained minimizer to (CP)

(i) x^* is a KKT point, i.e. $c^* \geq 0$ and there exists a nonempty set M_λ of multipliers λ satisfying $\lambda \geq 0, c^{*T}\lambda = 0$, and $g^* = J^{*T}p$;

(ii) the MFCQ holds at x^*

(iii) for all $\lambda \in M_\lambda$ and all nonzero p satisfying $g^{*T}p = 0$ and $J_A^* p \geq 0$, there exists $\omega > 0$ such that $p^T H(x^*, \lambda) p \geq \omega \|p\|^2$

IPM strategies

Strategy categories

- Algorithm type: affine-scaling, potential-reduction, and path-following
- Iterate space: primal, dual, primal-dual
- Iterate type: feasible and infeasible (equality constraints)
- Step type: short-step, long-step

Affine Scaling

LP'	$\underset{x}{\text{minimize}}$	$c^T x$	
	subject to	$Ax = b,$	$\|X^{k^{-1}} x - e\| \leq 1,$

Algorithm scheme is $x^{k+1} = x^k + \alpha \Delta x^k$ $(0 < \alpha \leq 1)$, where

$$\Delta x^k = -\frac{X^k P_{AX^k} X^k c}{\|P_{AX^k} X^k c\|},$$

and $P_{AX^k} = I - (AX^k)(AX^k(AX^k)^T)^{-1}(AX^k)^T$ is the projection matrix onto the null space of AX^k

Convergence proof exists when $\alpha = 1/8$, but no complexity results

Potential Reduction

PRP

$$\begin{aligned} \underset{x,y,z}{\text{minimize}} \quad & P(x,z) = q\log(x^T z) + I(x,z) \\ \text{subject to} \quad & Ax = b \\ & x > 0, \\ & A^T y + z = c \\ & z \geq 0, \end{aligned}$$

q is a potential function parameter (Karmarkar used $q = m$)

$I(x)$ is usually logarithmic interior function or Tanabe-Todd-Ye:

$$I(x) = -\sum_{i=1}^{m} \ln(x_j z_j)$$

Assuming min δ reduction per iteration yields max N iterations:

$$N = \frac{q}{\delta}(\log(\frac{x^T z}{\epsilon}) + C_2)$$

Path-following

$$
\text{LP}(\mu) \qquad \underset{x}{\text{minimize}} \quad B(x,\mu) = \phi(x) + \mu I(x) \\
\text{subject to} \quad c_i(x) \geq 0,
$$

where $I(x)$ satisfies:

- $I(x)$ depends only on $x \geq 0$
- $I(x)$ preserves continuity
- For any feasible sequence converging to boundary, $I(x) \to \infty$

Two frequently used interior functions:

$$
I(x) = \sum_{i=1}^{m} \frac{1}{x_j}, \quad I(x) = -\sum_{i=1}^{m} \ln(x_j)
$$

Formulating the Primal-Dual Equation

$$F^\mu(x,y) = \begin{pmatrix} g(x) - J(x)^T y \\ C(x)y - \mu e \end{pmatrix} = 0$$

Newton step:

$$F^\mu(x,y)'(\Delta x, \Delta y) = -F^\mu(x,y)$$

Collecting terms on the right-hand side:

$$\begin{pmatrix} H(x,y) & -J(x)^T \\ YJ(x) & C(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = -\begin{pmatrix} g(x) - J(x)^T y \\ C(x)(y - \pi(x,\mu)) \end{pmatrix},$$

$\pi_i = \mu / c_i(x)$ for $i = 1, \ldots, m$, and we assume $C \succ 0, Y \succ 0$

Efficiently Solving the Newton Equatio

One option: Pre-conditioned CG on P-D Newton Equations

Block elimination allows for LL^T Cholesky factorization to

$$(H(x,y) + J(x)^T Y^{-1} C(x) J(x))\Delta x = -(g(x) - J(x)^T \pi(x,\mu))$$

Pre-multiplying by $Y^{-1/2}$ allows for $PMP^T = LDL^T$ Cholesky factorization to

$$\begin{pmatrix} H(x,y) & J(x)^T Y^{1/2} \\ Y^{1/2} J(x) & -C(x) \end{pmatrix} \begin{pmatrix} \Delta x \\ Y^{-1/2}\Delta y \end{pmatrix} = -\begin{pmatrix} g(x) - yJ(x)^T \\ Y^{-1/2} C(x)(y-\pi) \end{pmatrix}$$

Line Search

Exact: find argmin of merit function along Newton step direction
Inexact: find step size reducing merit function by a threshold

Backtracking is most popular inexact line search:

Given direction Δx and $\alpha \in (0, 0.5), \beta \in (0, 1), t^0 = 1$, $t^{k+1} = \beta t^k$
while

$$f(x + t\Delta x) > f(x) + \alpha t \nabla f(x)^T \Delta x$$

Consider residual based on actual newton step p_k

$$r^k = \nabla^2 f(x^k) p_k + \nabla f(x^k)$$

Inexact methods are locally convergent if

$$\|r^k\| / \|\nabla f(x^k)\| \leq \eta_k$$

and $\{\eta_k\}$ is uniformly less than 1

Handling μ

- Too large μ yields small step sizes (slow)
- Too small μ yields steps far from central path

Adaptive updates are typically used:

$$\mu^k = \sigma^k \frac{(x^k)^T z}{n}$$

n is the dimension of x and $\sigma^k \in (0, 1)$, since iterates are assumed not to be on central path

Outline

1 Background on IPMs

2 PDCO

3 Zoom

4 Conclusions and Next Steps

PDCO Solver

MATLAB primal-dual interior method

http://www.stanford.edu/group/SOL/software.html

Nominal problem:

$$\text{NP} \qquad \underset{x}{\text{minimize}} \quad \phi(x) \qquad \text{subject to} \quad Ax = b, \quad \ell \le x \le u$$

$\phi(x)$ convex, separable

Regularized problem:

$$\text{NP}(D_1, D_2) \qquad \underset{x,\, r}{\text{minimize}} \quad \phi(x) + \tfrac{1}{2}\|D_1 x\|^2 + \tfrac{1}{2}\|r\|^2 \qquad \text{subject to} \quad Ax + D_2 r = b, \quad \ell \le x \le u$$

PDCO Primal-Dual Equations

Introduce slack variables and replace nonnegativity constraints by the log barrier:

$$
\text{NP}(\mu) \qquad \begin{array}{lrcll}
\underset{x,r,x_1,x_2}{\text{minimize}} & \multicolumn{3}{l}{\phi(x) + \frac{1}{2}\|D_1 x\|^2 + \frac{1}{2}\|r\|^2 - \mu \sum_j \ln([x_1]_j [x_2]_j)} & \\
\text{subject to} & Ax + D_2 r & = & b & : y \\
 & x - x_1 & = & \ell & : z_1 \\
 & -x - x_2 & = & -u, & : z_2
\end{array}
$$

PDCO Primal-Dual Equations (cont.)

Eliminate $r = D_2 y$ and apply Newton's method:

$$\begin{pmatrix} \Delta x - \Delta x_1 \\ -\Delta x - \Delta x_2 \end{pmatrix} = \begin{pmatrix} r_\ell \\ r_u \end{pmatrix} \equiv \begin{pmatrix} \ell - x + x_1 \\ -u + x + x_2 \end{pmatrix},$$

$$\begin{pmatrix} X_1 \Delta z_1 + Z_1 \Delta x_1 \\ X_2 \Delta z_2 + Z_2 \Delta x_2 \end{pmatrix} = \begin{pmatrix} c_\ell \\ c_u \end{pmatrix} \equiv \begin{pmatrix} \mu e - X_1 z_1 \\ \mu e - X_2 z_2 \end{pmatrix},$$

$$\begin{pmatrix} A \Delta x + D_2^2 \Delta y \\ -H_1 \Delta x + A^T \Delta y + \Delta z_1 - \Delta z_2 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \equiv \begin{pmatrix} b - Ax - D_2^2 y \\ g + D_1^2 x - A^T y - z_1 + z_2 \end{pmatrix},$$

where $H_1 = H + D_1^2$.

PDCO Primal-Dual Equations (cont.)

Substitute the first 2 equations into the third:

$$\begin{pmatrix} -H_2 & A^T \\ A & D_2^2 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} w \\ r_1 \end{pmatrix},$$

where

$$\begin{aligned} H_2 &\equiv H + D_1^2 + X_1^{-1} Z_1 + X_2^{-1} Z_2 \\ w &\equiv r_2 - X_1^{-1}(c_\ell + Z_1 r_\ell) + X_2^{-1}(c_u + Z_2 r_u) \end{aligned}$$

PDCO search directions

3 methods for computing Δy:

- Cholesky on $(AD^2A^T + D_2^2 I)\Delta y = AD^2 w + D_2 r_1$
- Sparse QR on $\min \left\| \begin{pmatrix} DA^T \\ D_2 I \end{pmatrix} \Delta y - \begin{pmatrix} Dw \\ r_1 \end{pmatrix} \right\|$
- LSQR on same LS problem (iterative solver)

Must use LSQR when A is an operator

Scaling inside PDCO

PDCO allows inputs β and ζ to scale problem data

Guiding principle:

$$\begin{aligned}\beta &= \text{input estimate of} \|x\|_\infty \\ \zeta &= \text{input estimate of} \|z\|_\infty\end{aligned}$$

Typical choices: $\beta = \|b\|_\infty$ and $\zeta = \|\tilde{c}\|_\infty = \|\beta c\|_\infty$

Final scaling becomes:

$$\begin{array}{lll} \bar{A} = A, & \bar{b} = b/\beta, & \bar{c} = \beta c/\zeta, \\ \bar{l} = l/\beta, & \bar{u} = u/\beta, & \bar{x} = x/\beta, \\ \bar{y} = \beta y/\zeta, & \bar{z}_1 = \beta z_1/\zeta, & \bar{z}_2 = \beta z_2/\zeta, \\ \bar{D}_1 = \beta D_1/\sqrt{\zeta}, & \bar{D}_2 = \sqrt{\zeta} D_2/\beta, & \bar{r} = r/\sqrt{\zeta}. \end{array}$$

Outline

1 Background on IPMs
2 PDCO
3 Zoom
4 Conclusions and Next Steps

Motivation

The problem that started it all

Image reconstruction

Nagy and Strakoš 2000 Byunggyoo Kim thesis, 2002

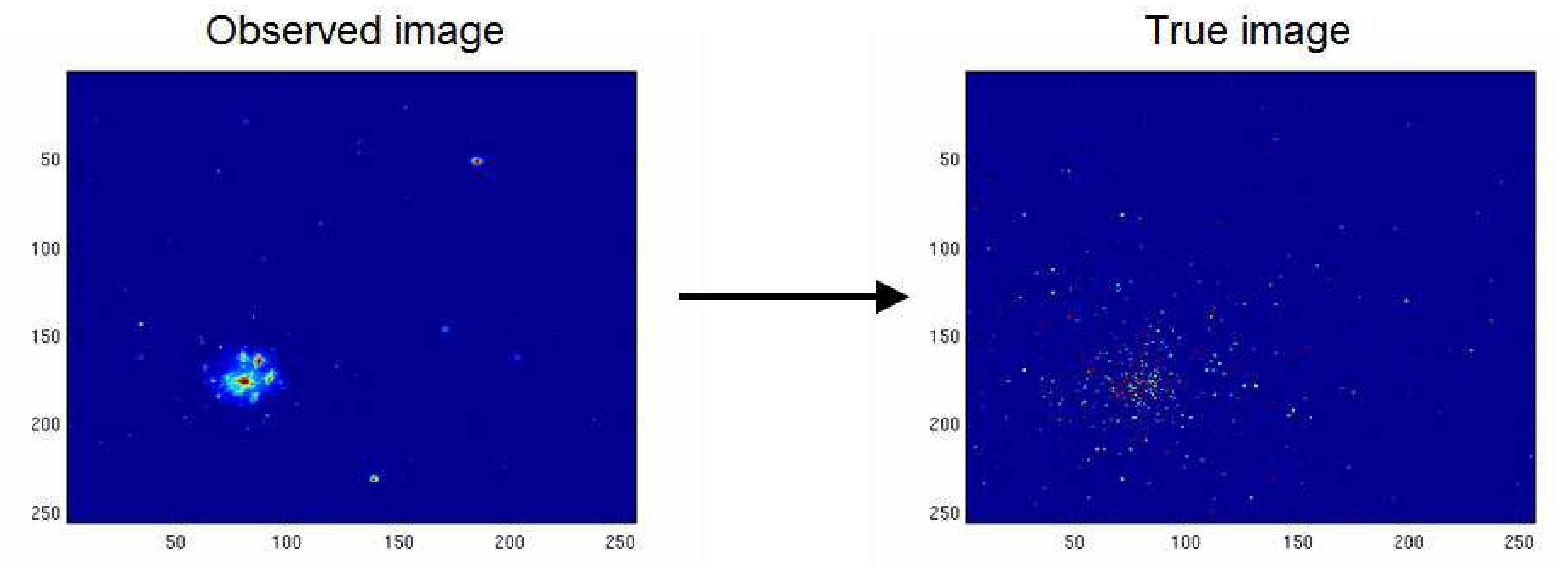


Image Reconstruction

$$\min \quad \lambda e^T x + \tfrac{1}{2}\|r\|^2$$
$$\text{st} \quad Ax + r = b, \quad x \geq 0$$

NNLS: Non-negative least squares $\lambda = 10^{-4}$

A is an expensive operator 2-D DFT

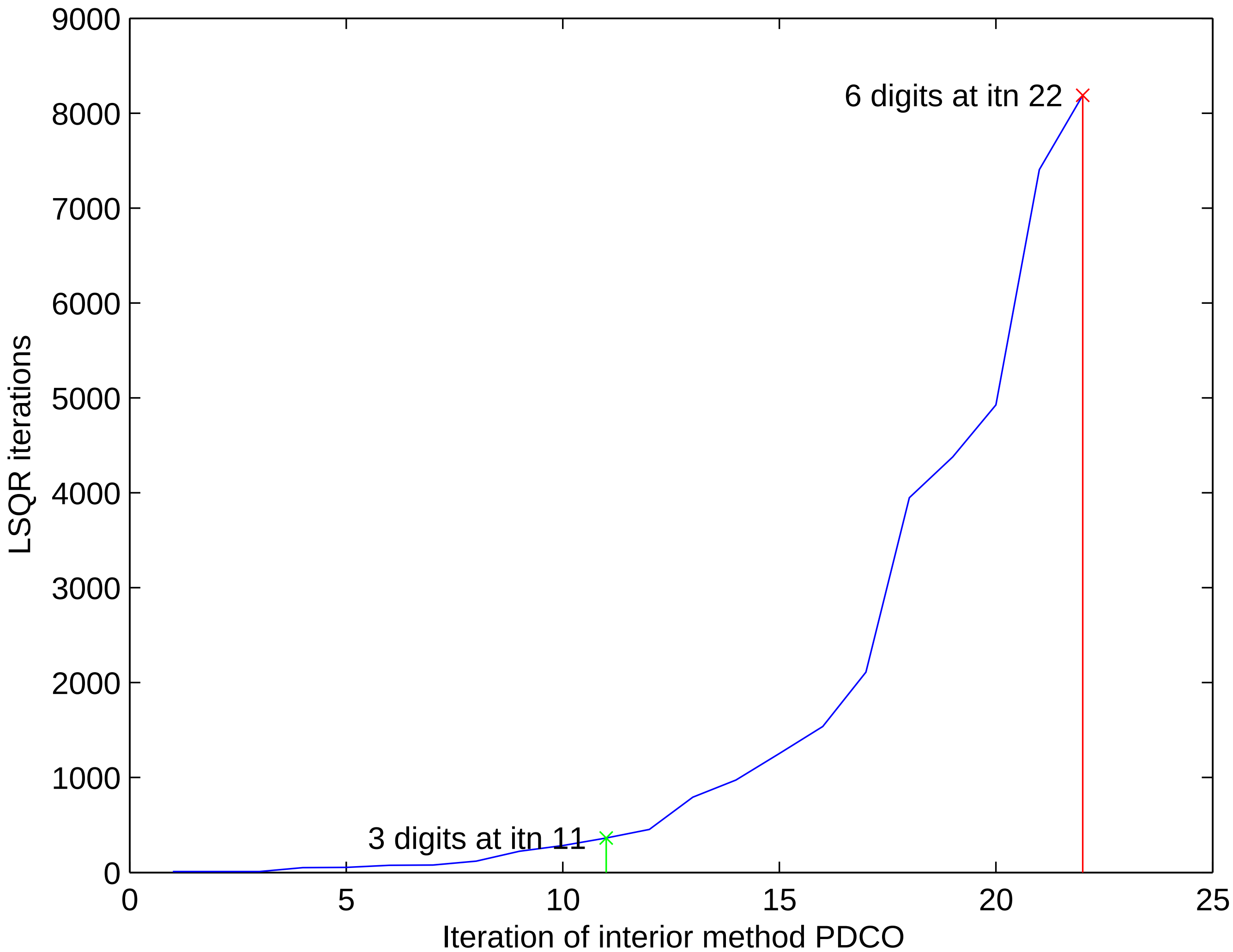
65K $\times$ 65K

PDCO uses LSQR for each dual search direction Δy:

$$\min \left\| \begin{pmatrix} DA^T \\ I \end{pmatrix} \Delta y - \begin{pmatrix} Dw \\ r_1 \end{pmatrix} \right\|$$

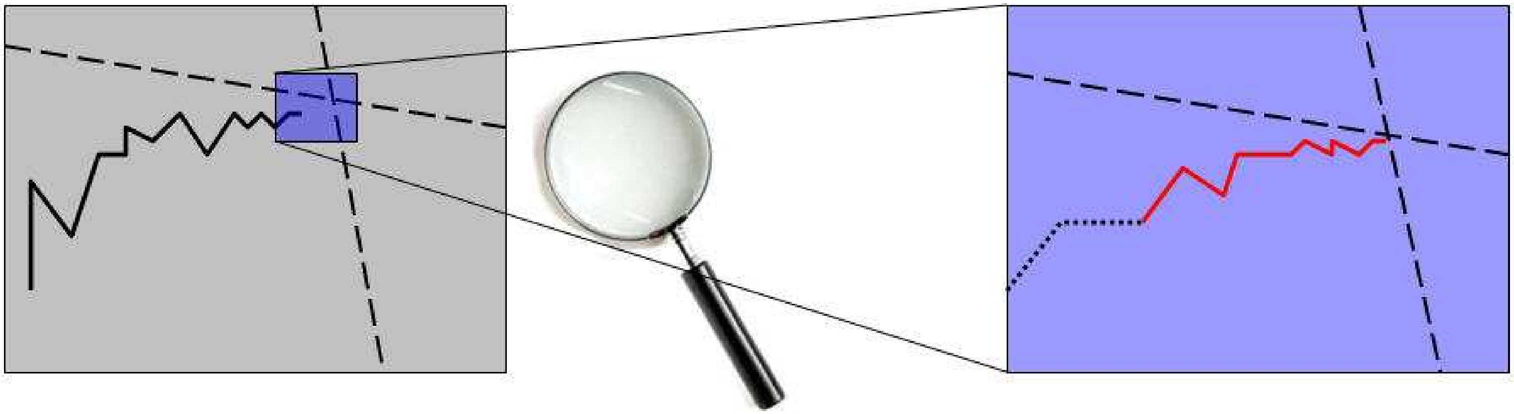
Motivation

LSQR iterations increase exponentially with requested accuracy



Zoom strategy: Accelerating IPMs

- Solve to 3 digits: cheap approximation to x, y, z
- Define new problem for correction dx, dy, dz
- Zoom in (scale up correction)
- Solve to 3 digits: cheap approximation to dx, dy, dz



Cold start for both solves

Zoom theory

Regularized LP:

$$
\begin{array}{lll}
\text{RLP} & \underset{x,r,x_1,x_2}{\text{minimize}} & c^Tx + \frac{1}{2}\|D_1x\|^2 + d^Tr + \frac{1}{2}\|r\|^2 + c_1^Tx_1 + c_2^Tx_2 \\
 & \text{subject to} & Ax + D_2r = b \quad : y \\
 & & x - x_1 = \ell \quad : z_1 \\
 & & -x - x_2 = -u \quad : z_2 \\
 & & x_1,\ x_2 \geq 0
\end{array}
$$

Suppose $(\widetilde{x}, \widetilde{y}, \widetilde{z}_1, \widetilde{z}_2, \widetilde{x}_1, \widetilde{x}_2, \widetilde{r})$ is an approximate solution

Redefine problem with

$$
\begin{aligned}
x &= \widetilde{x} + dx \\
r &= \widetilde{r} + dr
\end{aligned}
$$

Zoom theory (cont.)

RLP′

$$\begin{array}{lll} \underset{dx,dr,x_1,x_2}{\text{minimize}} & c^T dx + \frac{1}{2}\|D_1 dx\|^2 + \cdots & \\ \text{subject to} & A dx + D_2 dr = \widetilde{b} & : y \\ & dx - x_1 = \tilde{\ell} & : z_1 \\ & -dx - x_2 = -\tilde{u} & : z_2 \\ & x_1,\ x_2 \geq 0 & \end{array}$$

where

$$\begin{aligned} \widetilde{b} &= b - A\widetilde{x} - \delta\widetilde{r} \\ \tilde{\ell} &= \ell - \widetilde{x} \\ \tilde{u} &= u - \widetilde{x} \end{aligned}$$

Zoom theory (cont.)

Add Lagrangian terms

$$\tilde{y}^T(\tilde{b} - A\,dx - D_2\,dr) \qquad \tilde{z}_1^T(\tilde{\ell} - dx + x_1) \qquad \tilde{z}_2^T(-\tilde{u} + dx + x_2)$$

to objective:

RLP''

$$\begin{array}{lll} \underset{dx,dr,x_1,x_2}{\text{minimize}} & \tilde{c}^T dx + \frac{1}{2}\|D_1 dx\|^2 + \tilde{d}^T dr + \frac{1}{2}\|dr\|^2 + \tilde{c}_1^T x_1 + \tilde{c}_2^T x_2 & \\ \text{subject to} & A\,dx + D_2\,dr = \tilde{b} & : dy \\ & dx - x_1 = \tilde{\ell} & : dz_1 \\ & -dx - x_2 = -\tilde{u} & : dz_2 \\ & x_1,\ x_2 \geq 0 & \end{array}$$

Same form as original RLP

Primal **and** dual variables are **small**

Hence, scale up and use cold start

Revisions to PDCO

Define $\bar{z}_1 = z_1 + c_1, \bar{z}_1 = z_1 + c_1$

Eliminate $r = D_2 y - d$ and apply Newton's method:

$$\begin{pmatrix} \Delta x - \Delta x_1 \\ -\Delta x - \Delta x_2 \end{pmatrix} = \begin{pmatrix} r_\ell \\ r_u \end{pmatrix} \equiv \begin{pmatrix} \ell - x + x_1 \\ -u + x + x_2 \end{pmatrix},$$

$$\begin{pmatrix} X_1 \Delta z_1 + \bar{Z}_1 \Delta x_1 \\ X_2 \Delta z_2 + \bar{Z}_2 \Delta x_2 \end{pmatrix} = \begin{pmatrix} c_\ell \\ c_u \end{pmatrix} \equiv \begin{pmatrix} \mu e - X_1 \bar{z} \\ \mu e - X_2 \bar{z}_2 \end{pmatrix},$$

$$\begin{pmatrix} A\Delta x + D_2^2 \Delta y \\ -H_1 \Delta x + A^T \Delta y + \Delta z_1 - \Delta z_2 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} \equiv \begin{pmatrix} b - Ax - D_2^2 y + D_2 d \\ g + D_1^2 x - A^T y - z_1 + z_2 \end{pmatrix},$$

where $H_1 = H + D_1^2$

Revisions to PDCO (cont.)

Substitute the first 2 equations into the third:

$$\begin{pmatrix} -H_2 & A^T \\ A & D_2^2 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} w \\ r_1 \end{pmatrix}$$

as before, where

$$\begin{aligned} H_2 &\equiv H + D_1^2 + X_1^{-1}\bar{Z}_1 + X_2^{-1}\bar{Z}_2 \\ w &\equiv r_2 - X_1^{-1}(c_\ell + \bar{Z}_1 r_\ell) + X_2^{-1}(c_u + \bar{Z}_2 r_u) \end{aligned}$$

Scaling for additional terms:

$$\bar{c}_1 = \beta c_1/\zeta, \qquad \bar{c}_2 = \beta c_2/\zeta,$$
$$\bar{d} = d/\sqrt{\zeta}, \qquad \bar{\kappa} = \kappa/\zeta$$

Scaling Outside PDCO

A well-scaled A improves numerical properties inside PDCO.

Find diagonal matrices R, C such that $\hat{A} = R^{-1}AC^{-1}$ has entries close to 1.

Adjust other terms for problem consistency:

$$
\begin{aligned}
\hat{A} &= R^{-1}AC^{-1}, & \hat{b} &= R^{-1}b, & \hat{c} &= C^{-1}c, \\
\hat{c}_1 &= C^{-1}c_1, & \hat{c}_2 &= C^{-1}c_2, & \hat{l} &= Cl, \\
\hat{u} &= Cu, & \hat{x} &= Cx, & \hat{y} &= Ry, \\
\hat{z}_1 &= C^{-1}z_1, & \hat{z}_2 &= C^{-1}z_2, & \hat{D}_1 &= C^{-1}D_1, \\
\hat{D}_2 &= R^{-1}D_2,
\end{aligned}
$$

while d, r, κ remain unchanged.

Zooming is the choice of appropriate β, ζ for PDCO

Convergence

Theorem:
Using a strictly self-concordant barrier, if the sublevel set $S = \{x; f(x) \leq f(x^0)\}$ is closed and f is bounded below, then there exists $\eta, \gamma > 0$, with $0 < \eta \leq 1/4$, dependent only on the line search parameters such that:

- If $\lambda(x^k) > \eta$, then $f(x^{k+1}) - f(x^k) \leq -\gamma$
- If $\lambda(x^k) \leq \eta$, then the line search selects $t = 1$ and

$$2\lambda(x^{k+1}) \leq (2\lambda(x^k))^2.$$

In the damped phase the objective decreases monotonically by at least γ every iteration, so convergence is guaranteed (with the assumption about f on the sublevel set S being bounded below). Convergence is quadratic in the pure Newton phase

Complexity

Lemma (outer iterations):
Given μ^0 and updates of $\mu^{k+1} = \sigma\mu^k$, with $0 < \sigma < 1$, then after at most

$$\frac{1}{1-\sigma}\log\left(\frac{n\mu^0}{\varepsilon}\right)$$

iterations we have $n\mu \leq \varepsilon$

Lemma (inner iterations):
For given $\sigma, 0 < \sigma < 1$, the number of iterations between μ updates is not larger than

$$2\left(1 + \sqrt{\frac{(1-\sigma)\sqrt{n}}{\sigma}}\right)^4$$

Complexity (cont.)

Theorem:

Upper bound for total iterations of 2-phase zoom and refine technique:

$$O\left(n \log\left(\frac{n^2 \mu_1^0 \mu_2^0}{\varepsilon}\right)\right),$$

where μ_1^0 and μ_2^0 are the starting duality gap values for phase 1 and phase 2, respectively

A Closer Look at the Inner Workings

The typical Newton system for an IPM applied to an LP is of the form

$$\begin{pmatrix} -X^{-1}Z & A^T \\ A & \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} r_4 \\ r_1 \end{pmatrix}$$

Wright shows that in the degenerate case,

$$\mathrm{cond}(M) \approx \frac{1}{\mu}$$

for iterates near the central path and μ sufficiently small

Zoom Inner Workings (cont.)

At the end of the first stage with a target precision of 10^{-4} we have

$$\text{cond}(M_1) \approx \frac{1}{\mu} \simeq 10^4,$$

where M_1 represents the Jacobian at the intermediate solution

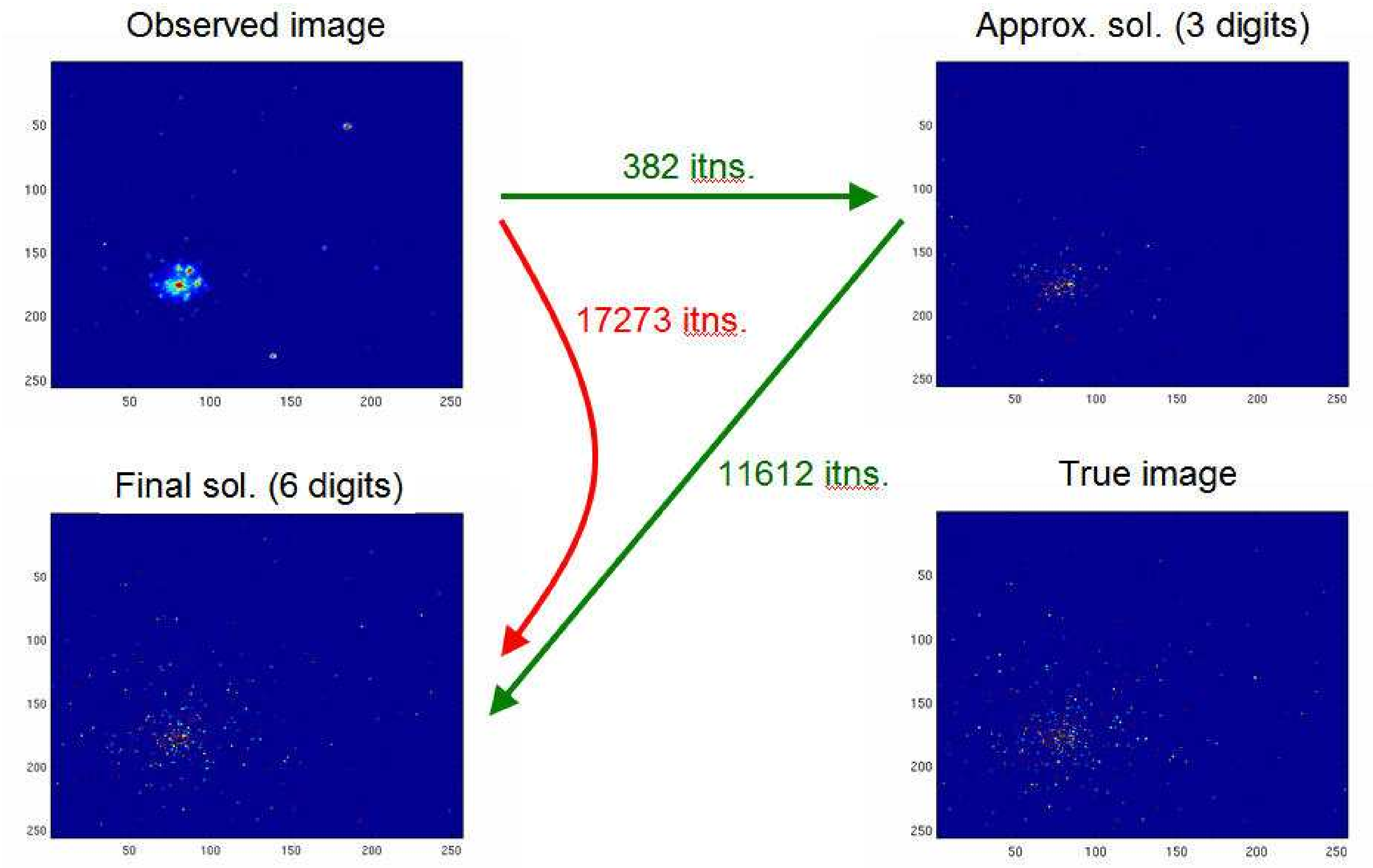
At the next stage, by design $1 \simeq \mu_2 >> \mu_1$ and hence $\text{cond}(M_2) \simeq 1$, so

$$\text{cond}(M_2) << \text{cond}(M_1),$$

where M_2 represents the Jacobian at the starting point for the scaled new problem (at the start of second stage)

Results: Accelerating IPMs

LSQR iterations inside PDCO



Netlib Results: Accelerating IPMs

LSQR iterations inside PDCO

Zoom strategy applied to Netlib LP problems:

criteria (based on single solve)	# problems	avg Zoom benefit
# LSQR itns < 5,000	16	-1.49%
5,000 $\leq$ # LSQR itns < 15,000	12	1.37%
15,000 $\leq$ # LSQR itns	12	28.14%

Motivation

Solving several related problems is common in industry

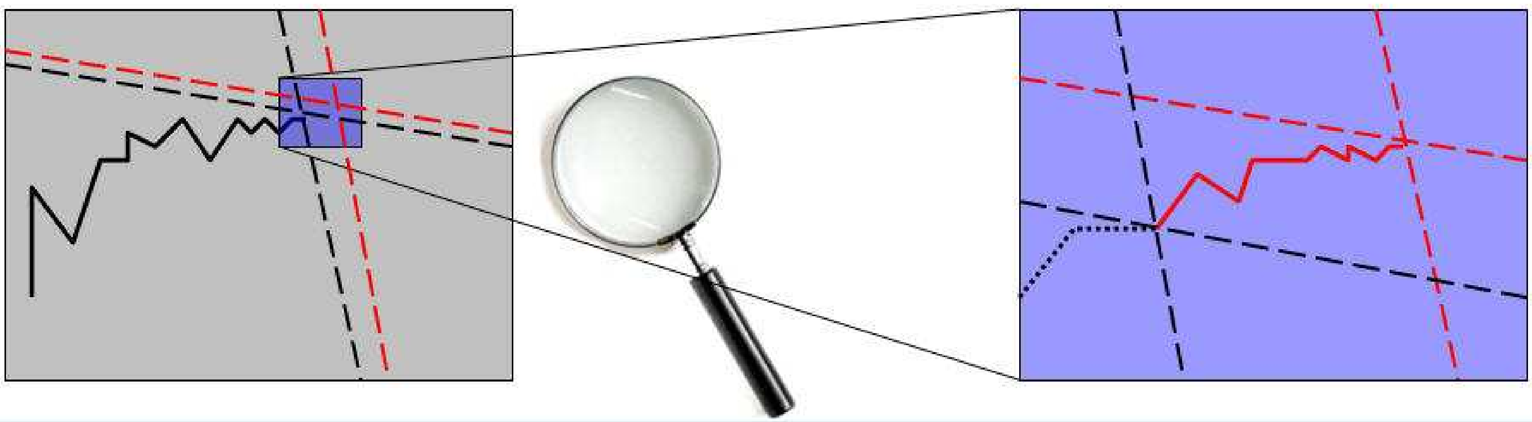
Reusing prior solutions in IPMs (warmstarting) is usually not efficient

- Solutions are close to each other but proximity to central path varies drastically between problems
- Newton steps end up being greatly shortened and backtracking often occurs
- Coldstarts typically take less time than warmstarting

Much renewed interest in research: Benson & Shanno (2005), Gondzio & Grothey (2006), Roos (2006)

Zoom strategy: Warm-starting IPMs

- Set solution to original LP as current approximation
- Define new problem for correction dx, dy, dz
- Zoom in (scale up correction)
- Solve loosely: cheap approximation to dx, dy, dz



Cold start for loose solve

Warm-starting IPMs

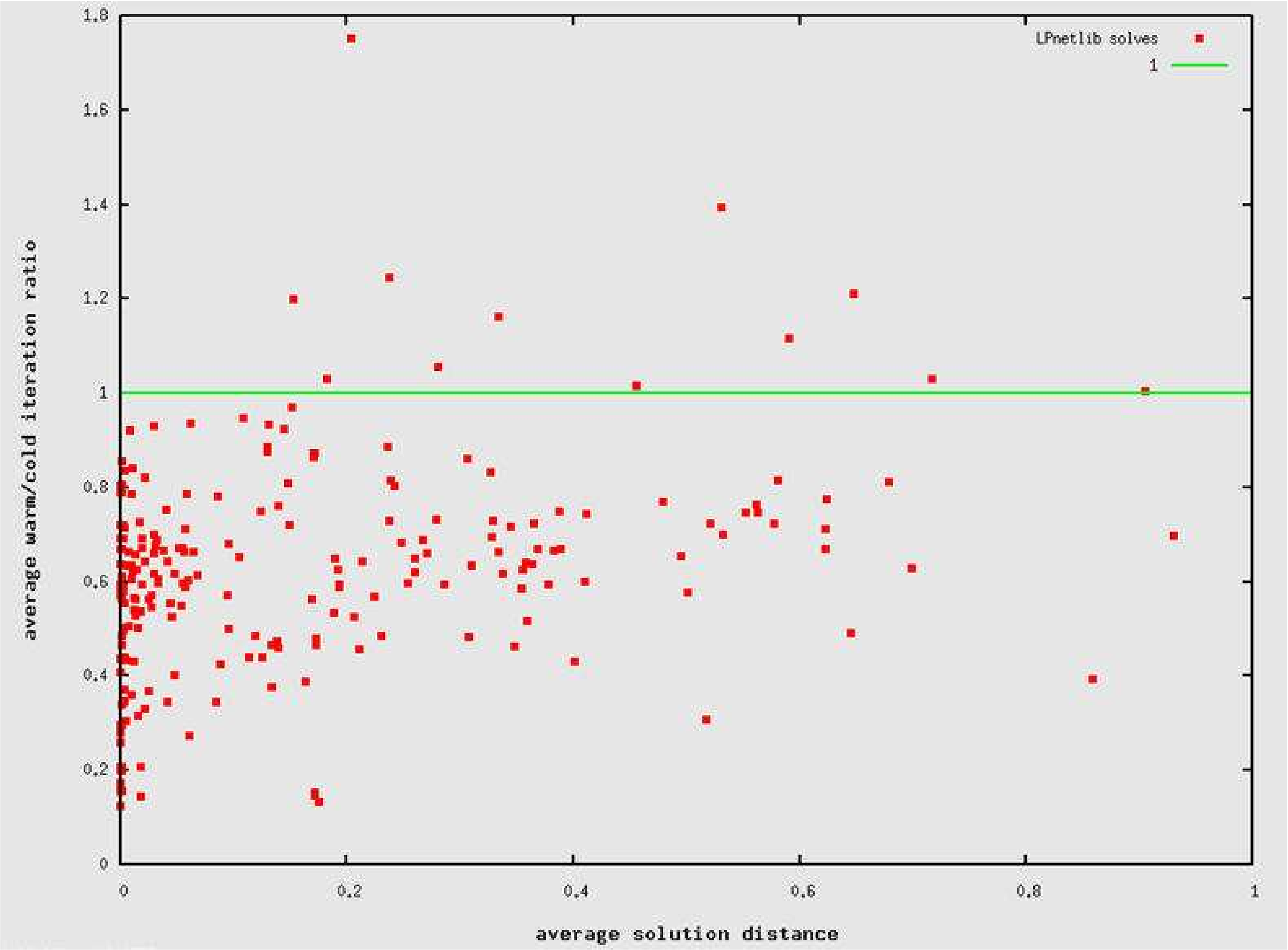
$$\text{LP}(\gamma,\delta) \qquad \underset{x,r}{\text{minimize}} \quad c^T x + \tfrac{1}{2}\|\gamma x\|^2 + \tfrac{1}{2}\|r\|^2$$
$$\text{subject to} \quad Ax + \delta r = b, \qquad \ell \le x \le u$$

Regularized LP $\quad \gamma = \delta = 10^{-3}$

PDCO with Cholesky on $AD^2A^T + \delta^2 I$

- LPnetlib problems with 5 random perturbations to A, b, or c
  (cf. Benson and Shanno 2005)
- Smaller problems ($<$ 100KB): 45 runs for each problem
- Compare Zoom to single solve

Results: Warm-starting IPMs



PDCO iterations (warm/cold) vs. perturbation to x, y

Next steps

- Multiple Zooms?
- Adaptive Zooms?
- How much is attributable to Zoom, to scaling?
- Explain outliers
  (e.g. Check size of residuals to decide Zoom scaling)

Conclusions

- Minor changes to existing primal-dual algorithms
- Zoom time reduced 30–60%