Postponing the choice of parameters in interior-point methods for Linear Programming

Luiz-Rafael Santos[1]
Aurelio R. L. Oliveira   Fernando Villas-Bôas   Clovis Perin

Department of Applied Mathematics
Institute of Mathematics, Statistics and Scientific Computation
University of Campinas – Brazil

LPOO/Campinas-SP
07 Dec 2012

[1] lrsantos@ime.unicamp.br

Outline

Outline

Some issues in Interior-point methods

- How to combine predictor, corrector and other directions to generate a better direction?
  - Different types of directions need to be combined in an efficient way, however it seems to be no magical formula valid for all problems
- How to keep interactions within “good conditions”?
  - Iterates have to be kept within some predefined conditions (neighborhoods of the central path, heuristics) that are successful in practice.

Some background

- [Colombo and Gondzio, 2008]: conditions that iterates should meet for good practical performance
- [Jarre and Wechs, 1999]: solve a small LP (simplex) to combine directions
- [Villas-Bôas and Perin, 2003]: Postpone the choice of the barrier parameter solving a polynomial optimization subproblem in auto-dual framework

Outline of our method

Develop and implement a method for Linear Programming problems that considers the points above, but based on – and extensively using – additional tools:

- A polynomial *predictive* merit function that allow us to know in advance some properties of the next iterate.
- Use of [Colombo and Gondzio, 2008]'s symmetric neighborhood, that defines conditions for the iterates, in terms of polynomial constraints
  - Keystone for good performance of any of their implementations

- Polynomials depend on the following parameters/variables (α, μ, σ)
  - α is the step length,
  - μ is the parameter for a more general central path,
  - σ models the weight of the the corrector direction (predictor-corrector method)
- The merit function constrained to the symmetric neighbourhood defines a polynomial optimizations subproblem, whose solutions give us the next iterate in a *optimal* way

Problem Formulation

- The standard linear programming primal and dual problems are

$$\begin{array}{ll} \min\limits_{x} & c^T x \\ \text{s.t.} & \begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{array} \qquad \text{(Primal)} \qquad\qquad \begin{array}{ll} \max\limits_{(y,z)} & b^T y \\ \text{s.t.} & \begin{cases} A^T y + z = b \\ z \geq 0, \ y \text{ free} \end{cases} \end{array} \qquad \text{(Dual)}$$

  where $A \in \mathbb{R}^{m \times n}$, $m \leq n$ is a full rank matrix, $c, x, z \in \mathbb{R}^n$ and $y, b \in \mathbb{R}^m$.

- The first order optimality conditions for this problem, the so called KKT conditions, can be written as

$$\begin{cases} Ax = b, & \text{(1a)} \\ A^T y + z = c, & \text{(1b)} \\ XZe = 0, & \text{(1c)} \\ (x, z) \geq 0. & \text{(1d)} \end{cases}$$

  where $X = \operatorname{diag}(x)$, $Z = \operatorname{diag}(z)$ and $e = (1, \ldots, 1)^T$.

Outline

Residuals

We define, for any (x, y, z), the vectors of residuals of (1), r_P, r_D and r_C, as

$$r_P = Ax - b, \tag{2a}$$

$$r_D = A^T y + z - c, \tag{2b}$$

$$r_C = XZe. \tag{2c}$$

Let (x^0, y^0, z^0) be an initial point such as $(x^0, z^0) > 0$. Then

$$\begin{aligned} r_P^0 &= Ax^0 - b, \\ r_D^0 &= A^T y^0 + z^0 - c, \\ r_C^0 &= X^0 Z^0 e > 0. \end{aligned}$$

Sign of KKT

To ensure that both r_P^0 and r_D^0 are non negative, we define diagonal matrices H_P and H_D, such as each entry of its diagonal is formed by

$$(H_P)_i = \begin{cases} 1, & \text{if } (r_P^0)_i \geq 0 \\ -1, & \text{if } (r_P^0)_i < 0 \end{cases}, \qquad (H_D)_j = \begin{cases} 1, & \text{if } (r_D^0)_j \geq 0 \\ -1, & \text{if } (r_D^0)_j < 0 \end{cases},$$

for $i = 1, \ldots, m$ and for $j = 1, \ldots, n$.

- $H_P r_P^0 \geq 0$ and $H_C r_C^0 \geq 0$
- The solution of the KKT system (1) and of

$$\begin{cases} H_P(Ax - b) = 0, & \text{(3a)} \\ H_D(A^T y + z - c) = 0, & \text{(3b)} \\ XZe = 0, & \text{(3c)} \\ (x, z) \geq 0, & \text{(3d)} \end{cases}$$

  are the same.
- We are only multiplying by a scalar each row of the usual KKT system.

Homotopy continuation method

We suggest a homotopy continuation method to solve the scaled KKT system (3) by approximately solving at each iteration, for any point (x, y, z) interior and any $\mu > 0$ the system

$$
\begin{cases}
H_P(Ax - b) = 0, & \text{(4a)} \\
H_D(A^T y + z - c) = 0, & \text{(4b)} \\
XZe = \mu e, & \text{(4c)} \\
(x, z) > 0 & \text{(4d)}
\end{cases}
$$

Affine-Scale Directions

- The affine-scaling (Newton) direction to approximately solve system (1) is found when one solves the following nonlinear system

$$
\begin{cases}
A\Delta x^{\text{af}} + r_P = 0 & \text{(5a)} \\
A^T \Delta y^{\text{af}} + \Delta z^{\text{af}} + r_D = 0 & \text{(5b)} \\
Z\Delta x^{\text{af}} + X\Delta z^{\text{af}} + r_C = 0 & \text{(5c)}
\end{cases}
$$

- Solve through Normal Equations (direct method)
- Involves one Cholesky factorization and one backsolve.

Ideal direction to homotopy

- Given (x, y, z) and $\mu > 0$
- How to find and ideal single step $\Delta w = (\Delta x, \Delta y, \Delta z)$, such as

$$\hat{w} = w + \Delta w,$$

that is solution of

$$\begin{cases} A\hat{x} - b = 0 \\ A^T\hat{y} + \hat{z} - c = 0 \\ \hat{X}\hat{Z}e = \mu e \end{cases} . \tag{6}$$

Predictor-Corrector directions to homotopy

- Define $\Delta w = \Delta w^{\text{af}} + \Delta w^{\text{c}}$, where Δw^{af} is the affine-scaling direction and Δw^{c} é the ideal corrector direction.
- Using some simplifications we obtain the Nonlinear system

$$\begin{cases} A\Delta x^c = 0 \\ A^T\Delta y^c + \Delta z^c = 0 \\ X\Delta z^c + Z\Delta x^c + \Delta X\Delta z = \mu e \end{cases} \tag{7}$$

Vector $\Delta X\Delta z$ is a second order direction similar to the ones used on [Mehrotra, 1992, Gondzio, 1996] works.

Our contribution

- For some scalar $\sigma > 0$ bounded, we are regarding the approximation

$$\Delta X\Delta z \approx \sigma\Delta X^{\text{af}}\Delta z^{\text{af}}$$

as acceptable.

A σ-weighted correction

- This approximation transforms the nonlinear system (8) into the linear system

$$\begin{cases} A\Delta x^c = 0 \\ A^T\Delta y^c + \Delta z^c = 0 \\ X\Delta z^c + Z\Delta x^c + \sigma\Delta X^{\text{af}}\Delta z^{\text{af}} = \mu e \end{cases} \tag{8}$$

- If $\sigma = 1$ and $\mu = (x^{\text{af}})^T(z^{\text{af}})/n)^3/(x^T z/n)$ we have Mehrotra's method
- In Gondzio's method, μ is chosen as in Mehrotra's, however $\Delta X \Delta z$ is multiple times approached by directions that are projections component wise, the complementarity onto the neighbourhood $N_s(\gamma)$.

μ and σ directions

- We can split the corrector direction as

$$\Delta w^{\mathrm{c}} = \mu \Delta w^{\mu} + \sigma \Delta w^{\sigma}, \tag{9}$$

μ and σ directions

- We can split the corrector direction as

$$\Delta w^{\mathsf{c}} = \mu \Delta w^{\mu} + \sigma \Delta w^{\sigma}, \tag{9}$$

- Why? We can write the system as

$$\nabla F(w) \Delta w^{c} = \begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \Delta x^c \\ \Delta y^c \\ \Delta z^c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \mu e - \sigma \Delta X^{\mathsf{af}} \Delta z^{\mathsf{af}} \end{bmatrix}, \tag{10}$$

providing the equality holds for every (μ, σ).

- Find vectors Δw^{μ} e Δw^{σ} when we solve systems $\nabla F(w) \Delta w^{\mu} = (0, 0, e)$ e $\nabla F(w) \Delta w^{\sigma} = (0, 0, -\Delta X^{\mathsf{af}} \Delta z^{\mathsf{af}})$ and the same Choleksy factorization that was used on the affine-scaling direction.

Purpose of these transformations

- The next point for each variable would be

$$\hat{x} = x + \alpha \Delta x \tag{11a}$$

$$\hat{y} = y + \alpha \Delta y \tag{11b}$$

$$\hat{z} = z + \alpha \Delta z \tag{11c}$$

Purpose of these transformations

- The next point for each variable would be

$$\hat{x} = x + \alpha(\Delta x^{\text{af}} + \mu\Delta x^{\mu} + \sigma\Delta x^{\sigma}) \tag{11a}$$

$$\hat{y} = y + \alpha(\Delta y^{\text{af}} + \mu\Delta y^{\mu} + \sigma\Delta y^{\sigma}) \tag{11b}$$

$$\hat{z} = z + \alpha(\Delta z^{\text{af}} + \mu\Delta z^{\mu} + \sigma\Delta z^{\sigma}) \tag{11c}$$

Purpose of these transformations

- The next point for each variable would be

$$\hat{x} = x + \alpha(\Delta x^{\text{af}} + \mu\Delta x^{\mu} + \sigma\Delta x^{\sigma}) \tag{11a}$$

$$\hat{y} = y + \alpha(\Delta y^{\text{af}} + \mu\Delta y^{\mu} + \sigma\Delta y^{\sigma}) \tag{11b}$$

$$\hat{z} = z + \alpha(\Delta z^{\text{af}} + \mu\Delta z^{\mu} + \sigma\Delta z^{\sigma}) \tag{11c}$$

- Same Cholesky factorization to find the 3 components of Δ.
- Up to three backsolves, but depending on “good direction” found
- To be chosen:
  - (μ, σ) that defines such new directions
  - the step length α.
- Expressed with the variables (α, μ, σ) any “educated guess” of an ideal straight direction from a point (x, y, z) to the μ-homotopy,

Outline

The Residual for the Signed KKT System

Definition

We define ρ, the *vector of residuals of the Signed KKT system* (3) for a point (x, y, z) as

$$\rho(x,y,z) = \begin{cases} \rho_P(x,y,z) = H_P(Ax - b) \\ \rho_D(x,y,z) = H_D(A^T y + z - c) \\ \rho_C(x,y,z) = XZe \end{cases} \tag{12}$$

Let $\rho_L = (\rho_P, \rho_D)^T \in \mathbb{R}^{m+n}$ be the linear residual of the Scaled KKT system. We define de vectors of residuals at iteration k as ρ^k. By construction $\rho^0 > 0$ for (x^0, y^0, z^0). We also define the next (predictive) residual at iteration k as

$$\hat{\rho} = \rho(x^{k+1}, y^{k+1}, z^{k+1}).$$

Polynomial Merit Function

Definition (Merit Function)

We define the *merit function* of a point (x, y, z) as

$$\varphi(x, y, z) = \frac{1}{m+n} \|\rho_L\|_1 + \frac{x^T z}{n} \tag{13}$$

Polynomial Merit Function

Definition (Merit Function)

We define the *merit function* of a point (x, y, z) as

$$\varphi(x, y, z) = \frac{1}{m+n} \|\rho_L\|_1 + \frac{x^T z}{n} \tag{13}$$

or

$$\varphi(x, y, z) = \frac{1}{m+n} \sum_{i=1}^{m+n} (\rho_L)_i + \frac{1}{n} \sum_{j=1}^{n} (\rho_C)_j \tag{14}$$

where ρ_L and ρ_C are the residuals of the Signed KKT System given by equation (12) at point (x, y, z).

Next residual

Proposition

The next residual for the KKT system (3) *is expressed as*

$$\hat{\rho}(\alpha,\mu,\sigma) = \begin{cases} (\hat{\rho}_L)_\ell = (1-\alpha)(\rho_L)_\ell, \\ \text{for } \ell = 1,\ldots,n+m. \\ (\hat{\rho}_C)_j = (1-\alpha)(\rho_C)_j + \alpha\mu + \alpha(\alpha-\sigma)(L_{0,0})_j + \alpha^2 \Lambda(\mu,\sigma)_j, \\ \text{for } j = 1,\ldots,n. \end{cases} \tag{15}$$

where

$$\Lambda(\mu,\sigma) = \left(\mu^2 L_{2,0} + \mu L_{1,0} + \mu\sigma L_{1,1} + \sigma^2 L_{0,2} + \sigma L_{0,1}\right)$$

and

$$\begin{aligned} L_{0,0} &= \Delta x^{af} \Delta z^{af} & L_{1,1} &= \Delta x^{\mu} \Delta z^{\sigma} + \Delta x^{\sigma} \Delta z^{\mu} \\ L_{1,0} &= \Delta x^{af} \Delta z^{\mu} + \Delta z^{af} \Delta x^{\mu} & L_{0,1} &= \Delta x^{af} \Delta z^{\sigma} + \Delta z^{af} \Delta x^{\sigma} \\ L_{2,0} &= \Delta x^{\mu} \Delta z^{\mu} & L_{0,2} &= \Delta x^{\sigma} \Delta z^{\sigma} \end{aligned}$$

Corollary: In each iteration, if $\alpha \in (0,1]$ and $(\mu,\sigma) > 0$, then $\rho(\alpha,\mu,\sigma) \geq 0$.

Polynomial Merit Function

Notation:

For any vector $v \in \mathbb{R}^p$, we define $\overline{v} = \dfrac{1}{p}\sum_{i=1}^{p} v_i$ (The arithmetic mean of v).

Polynomial Merit Function

Notation:

For any vector $v \in \mathbb{R}^p$, we define $\overline{v} = \dfrac{1}{p}\sum_{i=1}^{p} v_i$ (The arithmetic mean of v).

Remarks

- If (x^*, y^*, z^*) is solution if (3), then $\varphi(x^*, y^*, z^*) = 0$
- Matrices H_P e H_D and the corollary above ensure that given (x, y, z) interior calculated by our method, then $\rho_L(x, y, z) \geq 0$
- Equation (14) becomes:

$$\varphi(x, y, z) = \overline{\rho_L} + \overline{\rho_C} \tag{16}$$

- Notice that

$$\overline{\rho_L} = \frac{\|\rho_L\|_1}{m+n} \quad \text{and} \quad \overline{\rho_C} = \frac{x^T z}{n}$$

Predicting the Next Merit

- How can one predict the Merit Function value for the next point $(\hat{x}, \hat{y}, \hat{z})$?

Predicting the Next Merit

- How can one predict the Merit Function value for the next point $(\hat{x}, \hat{y}, \hat{z})$?

Definition (Next Merit)

The *next merit function* at iteration k is

$$\hat{\varphi}(x^k, y^k, z^k) = \overline{\hat{\rho}_L}(x^k, y^k, z^k) + \overline{\hat{\rho}_C}(x^k, y^k, z^k).$$

Because of equation (15) we can write

$$\hat{\varphi}(\alpha, \mu, \sigma) = \overline{\hat{\rho}_L}(\alpha, \mu, \sigma) + \overline{\hat{\rho}_C}(\alpha, \mu, \sigma) \quad (17)$$

Predicting the Next Merit

- How can one predict the Merit Function value for the next point $(\hat{x}, \hat{y}, \hat{z})$?

Definition (Next Merit)

The *next merit function* at iteration k is

$$\hat{\varphi}(x^k, y^k, z^k) = \overline{\hat{\rho}_L}(x^k, y^k, z^k) + \overline{\hat{\rho}_C}(x^k, y^k, z^k).$$

Because of equation (15) we can write

$$\hat{\varphi}(\alpha, \mu, \sigma) = \overline{\hat{\rho}_L}(\alpha, \mu, \sigma) + \overline{\hat{\rho}_C}(\alpha, \mu, \sigma) \tag{17}$$

Proposition (Predictive Merit Function)

Using Equations (15) *the predictive polynomial merit function can be expressed as the following polynomial on variables* (α, μ, σ)*.*

$$\hat{\varphi}(\alpha, \mu, \sigma) = (1-\alpha)(\overline{\rho_L}^k + \overline{\rho_C}^k) + \alpha\mu + \alpha(\alpha - \sigma)\overline{L_{0,0}} + \alpha^2\overline{\Lambda(\mu, \sigma)}$$

Predicting the Next Merit

- How can one predict the Merit Function value for the next point $(\hat{x}, \hat{y}, \hat{z})$?

Definition (Next Merit)

The *next merit function* at iteration k is

$$\hat{\varphi}(x^k, y^k, z^k) = \overline{\hat{\rho}_L}(x^k, y^k, z^k) + \overline{\hat{\rho}_C}(x^k, y^k, z^k).$$

Because of equation (15) we can write

$$\hat{\varphi}(\alpha, \mu, \sigma) = \overline{\hat{\rho}_L}(\alpha, \mu, \sigma) + \overline{\hat{\rho}_C}(\alpha, \mu, \sigma) \tag{17}$$

Proposition (Predictive Merit Function)

Using Equations (15) *the predictive polynomial merit function can be expressed as the following polynomial on variables* (α, μ, σ).

$$\hat{\varphi}(\alpha, \mu, \sigma) = (1 - \alpha)(\overline{\rho_L}^k + \overline{\rho_C}^k) + \alpha\mu + \alpha(\alpha - \sigma)\overline{L_{0,0}} + \alpha^2\overline{\Lambda(\mu, \sigma)}$$

$$\hat{\varphi}(\alpha, \mu, \sigma) = \sum_{i=0}^{2}\sum_{j=0}^{2}\sum_{\ell=0}^{2} a_{i,j,\ell}\alpha^i\mu^j\sigma^\ell$$

Generalized symmetric neighborhood

- [Colombo and Gondzio, 2008] proposed a neighborhood $\mathcal{N}_s$ that the iterates should comply in order to be "good". For $\gamma \in (0,1)$ and $\beta > 1$ and (x,y,z) infeasible they define

$$\mathcal{N}_s(\gamma,\beta) = \left\{(x,y,z) \in \mathcal{Q}^+ : \frac{\|\rho_L\|}{\tau} \leq \beta \frac{\|\rho_L^0\|}{\tau_0}, \gamma\tau \leq x_i z_i \leq \frac{1}{\gamma}\tau, \forall i = 1,\ldots,n\right\}. \tag{18}$$

Generalized symmetric neighborhood

- [Colombo and Gondzio, 2008] proposed a neighborhood $\mathcal{N}_s$ that the iterates should comply in order to be "good". For $\gamma \in (0,1)$ and $\beta > 1$ and (x,y,z) infeasible they define

$$\mathcal{N}_s(\gamma,\beta) = \left\{ (x,y,z) \in \mathcal{Q}^+ : \frac{\|\rho_L\|}{\tau} \leq \beta \frac{\|\rho_L^0\|}{\tau_0}, \gamma\tau \leq x_i z_i \leq \frac{1}{\gamma}\tau, \forall i = 1,\ldots,n \right\}. \tag{18}$$

Theorem

A point (x,y,z) *interior is on* $\mathcal{N}_{gs}(\gamma,\beta)$ *if the following inequalities hold*

$$\overline{\rho_L}(x,y,z) \leq \beta_L \overline{\rho_C}(x,y,z), \tag{19a}$$

$$\gamma \overline{\rho_C}(x,y,z) \leq (\rho_C)_i(x,y,z) \leq \frac{1}{\gamma}\overline{\rho_C}(x,y,z), \tag{19b}$$

for $i = 1,\ldots,n$ *and*

$$\beta_L = \frac{\beta}{n} \frac{\|\rho_L^0\|}{\overline{\rho_C}_0}.$$

PO Subproblem

- Find (α, μ, σ) such that it maximizes globally the *next merit function* constrained to constrained to $\mathcal{N}_{gs}$, i.e.,

$$\min_{(\alpha,\mu,\sigma)} \hat{\varphi}(\alpha, \mu, \sigma)$$

$$\text{s.t.} \quad (\hat{x}, \hat{y}, \hat{z}) \in \mathcal{N}_{gs} \text{ and the ratio test.}$$

PO Subproblem

- Find (α, μ, σ) such that it maximizes globally the *next merit function* constrained to constrained to $\mathcal{N}_{gs}$, i.e.,

$$\begin{array}{ll} \min\limits_{(\alpha,\mu,\sigma)} & \hat{\varphi}(\alpha, \mu, \sigma) \\ \text{s.t.} & \begin{cases} \psi(\alpha, \mu, \sigma) \geq 0 \\ l \leq (\alpha, \mu, \sigma) \leq u \end{cases} \end{array} \qquad \text{(PO Subproblem)}$$

PO Subproblem

- Find (α, μ, σ) such that it maximizes globally the *next merit function* constrained to constrained to $\mathcal{N}_{gs}$, i.e.,

$$\min_{(\alpha,\mu,\sigma)} \quad \hat{\varphi}(\alpha, \mu, \sigma)$$
$$\text{s.t.} \quad \begin{cases} \psi(\alpha, \mu, \sigma) \geq 0 \\ l \leq (\alpha, \mu, \sigma) \leq u \end{cases} \qquad \text{(PO Subproblem)}$$

- Global optimization of a polynomial constrained to a set of $2n + 1$ polynomials and a box
- $\hat{\varphi}$ and ψ are a 2nd degree, 3 variable polynomials on variables (α, μ, σ).
  - $\varphi(\alpha, \mu, \sigma)$, under a variable transformation, can be seen as a Quadratic function. There are cases where the Hessian is indefinite (NP-Hard problem)

Global optimization approach

Smart Grid using Cubic Splines

- Order of computation: $(\mu \to \sigma \to \alpha)$, since μ is often zero
- POP Subproblem can be rewritten as

$$\begin{array}{ll} \min\limits_{(\mu,\sigma)} & \hat{\varphi}_R(\mu,\sigma) \\ \text{s.t.} & \begin{cases} l_\mu \le \mu \le u_\mu \\ l_\sigma \le \sigma \le u_\sigma \end{cases} \end{array} \quad \text{where } \hat{\varphi}_R(\mu,\sigma) = \begin{array}{ll} \min\limits_{\alpha} & \hat{\varphi}(\alpha,\mu,\sigma) \\ \text{s.t.} & \begin{cases} \psi(\alpha,\mu,\sigma) \ge 0 \\ l_\alpha \le \alpha \le u_\alpha \end{cases} \end{array} \tag{20}$$

- PO subproblem can be approximately solve in a competitive time (IMA).
- Most of constrains ψ are not active.
  - Preprocessing using Ranges and Quadratic Programming.

Global optimization approach

Smart Grid using Cubic Splines

- Order of computation: $(\mu \to \sigma \to \alpha)$, since μ is often zero
- POP Subproblem can be rewritten as

$$\begin{array}{ll} \min\limits_{(\mu,\sigma)} & \hat{\varphi}_R(\mu,\sigma) \\ \text{s.t.} & \begin{cases} l_\mu \le \mu \le u_\mu \\ l_\sigma \le \sigma \le u_\sigma \end{cases} \end{array} \quad \text{where } \hat{\varphi}_R(\mu,\sigma) = \begin{array}{ll} \min\limits_{\alpha} & \hat{\varphi}(\alpha,\mu,\sigma) \\ \text{s.t.} & \begin{cases} \psi(\alpha,\mu,\sigma) \ge 0 \\ l_\alpha \le \alpha \le u_\alpha \end{cases} \end{array} \tag{20}$$

- PO subproblem can be approximately solve in a competitive time (IMA).
- Most of constrains ψ are not active.
  - Preprocessing using Ranges and Quadratic Programming.

Proposition

Let $(\alpha^*, \mu^*, \sigma^*)$ be a global solution of the PO subproblem and $(\bar{\mu}, \bar{\sigma})$ be a global solution of (20)*. Then,*

$$\hat{\varphi}(\alpha^*, \mu^*, \sigma^*) = \varphi_R(\bar{\mu}, \bar{\sigma})$$

Outline

Numerical Highlights

1. Complete computation requires up to 3 backsolves at each iterates, but 1 or 2 can be saved if we have a “good enough” direction.
2. Preliminary implemented method had optimal $\mu = 0$ in more than 80% of iterations (without centralization)
   - Only two backsolves are needed in these cases
3. So far, best way to evaluate the global optimization in the merit function performed is using Smart Grid and reprocessing constraints
4. Method is competitive with PCx: Iteration count and time ranging from 80% to 130% of PCx performance.
   - Converges in $\approx 70\%$ of NETLIB problems

Furtherwork

- Complexity and convergence of method is being proved.

  Roadmap: For $\sigma = 0$ and $\mu = \eta x^T z/n$ find $\alpha > 0$ such that

  $$\varphi^{k+1} < (1 - \theta(\alpha))\varphi^k.$$

  - Following the approach of [Zhang, 1994] for infeasible method.
  - In practice, much better improvement in each iteration.
- There were problems where no further improvement in optimality merit function could be obtained
  - Matrices H_P and H_D can be used as scaling factors that guarantee if $\varphi < \varepsilon$, then stop criteria of `PCx` is achieved.
- Compare and test other global optimization methods for merit (polynomial) function, which is the core of our method.
- More robust implementations is being performed.

Outline

References

M. Colombo and J. Gondzio. Further development of multiple centrality correctors for interior point methods.
Computational Optimization and Applications, **41**(3), 277-305 (2008).

J. Gondzio.
Multiple centrality corrections in a primal-dual method for linear programming.
Computational Optimization and Applications, 6(2):137–156, 1996.

F. Jarre and M. Wechs.
Extending Mehrotra's corrector for linear programs.
Advanced Modeling and Optimization, **1**(2), 38-60 (1999).

S. Mehrotra.
On the Implementation of a Primal-Dual Interior Point Method.
SIAM Journal on Optimization, 2(4):575–601, 1992.

F. Villas-Bôas and C. Perin.
Postponing the choice of penalty parameter and step length.
Computational Optimization and Applications, **24**(1), 63-81 (2003).

Y. Zhang
On the convergence of a class of infeasible interior-point methods for the horizontal linear complementarity problem
SIAM Journal of Optimization, **4**(1), 208-227 (1994).