Continued Iteration on Predictor Corrector Interior Point Method

Lilian Berti   Aurelio Oliveira   Carla Ghidini

IMECC – UNICAMP

April, 2015

Outline

- Predictor Corrector Interior Point Method
- Continued Iteration
- Numerical Experiments
- Conclusion and Future Works

Linear Programming Problems

Primal Problem

$$\begin{aligned} \text{Min} \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0. \end{aligned}$$

Dual Problem

$$\begin{aligned} \text{Max} \quad & b^T y \\ \text{s.t.} \quad & A^T y + z = c, \\ & z \geq 0. \end{aligned}$$

Predictor Corrector Interior Point Method

The predictor corrector method consists of two directions:

- Predictor direction

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \tilde{d}x \\ \tilde{d}y \\ \tilde{d}z \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_a \end{bmatrix}.$$

- Corrector direction

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \bar{d}x \\ \bar{d}y \\ \bar{d}z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ r_2 \end{bmatrix},$$

where $r_2 = \mu e - (\tilde{D}x\tilde{D}z)e$, $\tilde{D}x = diag(\tilde{d}x)$ and $\tilde{D}z = diag(\tilde{d}z)$.

Predictor Corrector Interior Point Method

Predictor direction

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \tilde{d}x \\ \tilde{d}y \\ \tilde{d}z \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_a \end{bmatrix}.$$

Eliminating $\tilde{d}z$, we get the augmented system:

$$\begin{bmatrix} -D^{-1} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \tilde{d}x \\ \tilde{d}y \end{bmatrix} = \begin{bmatrix} r_d - X^{-1} r_a \\ r_p \end{bmatrix},$$

where $D = Z^{-1}X$.
Eliminating $\tilde{d}x$

Predictor Corrector Interior Point Method

Symetric positive definite normal equations system

$$ADA^T\tilde{d}y = AD(r_d - X^{-1}r_a) + r_p.$$

Other directions:

$$\tilde{d}x = D(A^T\tilde{d}y - r_d + X^{-1}r_a),$$

$$\tilde{d}z = X^{-1}(r_a - Z\tilde{d}x).$$

Predictor Corrector Interior Point Method

The predictor corrector direction is given by $d = \tilde{d} + \bar{d}$ and it can be interpreted as the following linear system solution:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} r_p \\ r_d \\ r_s \end{bmatrix},$$

where $r_s = r_a + r_2$.

Predictor Corrector Interior Point Method

Predictor corrector direction:

- Solve two linear systems at each iteration with the same ADA^T matrix.
- The ADA^T matrix is symetric positive definite.
- Cholesky factorization is used to solve the linear systems.

Continued Iteration

The continued iteration is a approach proposed and incorporated on predictor corrector interior point method in order to:

- reduce the number of iterations;
- reduce the running time.

A new direction $\hat{d}$ is computed: the continued direction.

Continued directions

Motivation

$$r_p^{k+1} = (1 - \alpha_P) r_p^k,$$
$$r_d^{k+1} = (1 - \alpha_D) r_d^k,$$

We propose the continued directions $\hat{d}$ to increase α_P and α_D, reducing the infeasibilities, if $\alpha_P < 1$ and/or $\alpha_D < 1$.

Continued directions

The blocking components in the directions (dx, dz) are given by:

$$i = \arg\min_t \{-\frac{x_t}{dx_t} | x_t + dx_t \leq 0\},$$

$$j = \arg\min_t \{-\frac{z_t}{dz_t} | z_t + dz_t \leq 0\}.$$

- If there is the blocking component i, then $\alpha_P < 1$;
- If there is the blocking component j, then $\alpha_D < 1$;

Continued directions

We compute $\hat{d}$, with:

- Smallest possible predictor corrector direction (d) change.

Allowing:

- Increase α_P and α_D for all direction components that do not block.
- At the same time, keep the maximum value of the stepsize for the blocking components in α_P and α_D. All proposed directions have $\hat{d}x_i = 0$ and/or $\hat{d}z_j = 0$.

Continued directions

Two continued directions are proposed

- Moderated direction;
- Simple direction.

Moderated direction

The moderated direction $\hat{d}$ have to aproximadaly satisfy the predictor corrector direction linear system:

$$\begin{bmatrix} A & 0 & 0 \\ 0 & A^t & I \\ Z & 0 & X \end{bmatrix} \begin{bmatrix} \hat{d}x \\ \hat{d}y \\ \hat{d}z \end{bmatrix} \simeq \begin{bmatrix} r_p \\ r_d \\ r_s \end{bmatrix},$$

where

$$\hat{d}x_i = 0$$

and / or

$$\hat{d}z_j = 0.$$

To determine the moderated direction is necessary to solve an additonal linear system.

Moderated direction

We formulate a subproblem that determines $\hat{d}x$, such that the change from the predictor corrector direction dx is the lowest possible:

$$\begin{array}{ll} \min & \frac{1}{2}\|D^{-\frac{1}{2}}\hat{d}x - D^{-\frac{1}{2}}dx\|^2 \\ \text{s.t} & A\hat{d}x = r_p \\ & \hat{d}x_i = \beta_a \quad \text{and / or} \quad \hat{d}x_j = \beta_b, \end{array}$$

where $D = Z^{-1}X$, $\hat{d}x$ and $dx \in \mathbb{R}^n$, $\beta_a = 0$ and $\beta_b = -x_j + \frac{\mu}{z_j}$.

Moderated direction

Subproblem solution:

$$\hat{dx} = dx - DA^T v - \alpha d_{ii} e_i - \gamma d_{jj} e_j,$$

where

$$v = -\alpha d_{ii} B^{-1} A_i - \gamma d_{jj} B^{-1} A_j,$$

$$\alpha = \frac{(dx_i - \beta_a) + \gamma d_{ii} d_{jj} A_i^T B^{-1} A_j}{d_{ii}(1 - dii A_i^T B^{-1} A_i)},$$

$$\gamma = \frac{(dx_j - \beta_b)(1 - d_{ii} A_i^T B^{-1} A_i) + (dx_i - \beta_a) d_{jj} A_i^T B^{-1} A_j}{d_{jj}\left[(1 - d_{jj} A_j^T B^{-1} A_j)(1 - d_{ii} A_i^T B^{-1} A_i) - d_{ii} d_{jj} (A_i^T B^{-1} A_j)^2\right]},$$

$$B = ADA^T.$$

Moderated direction

Given $\hat{d}x$, we compute the other directions:

$$\hat{d}z = X^{-1}(r_a - Z\hat{d}x).$$
$$A^t\hat{d}y = r_d - \hat{d}z.$$

To take advantage existing Cholesky factorization, consider $D^{\frac{1}{2}}A^t\hat{d}y = D^{\frac{1}{2}}(r_d - \hat{d}z)$. From normal equations system, we get:

$$\hat{d}y = (ADA^t)^{-1}AD(r_d - \hat{d}z).$$

Simple direction

Compute

$$\hat{d}x^k = \begin{cases} dx^k - dx_i^k e_i, & \text{if exists the blocking component } i; \\ dx^k, & \text{otherwise,} \end{cases}$$

$$\hat{d}z^k = \begin{cases} dz^k - dz_j^k e_j, & \text{if exists the blocking component } j; \\ dz^k, & \text{otherwise,} \end{cases}$$

$$\hat{d}y^k = dy^k.$$

where e_i and e_j are canonical vector.

It is not necessary to solve additional linear systems.

Update approaches

The continued direction can be applied in two different approaches:

- Early approach:

$$\hat{x}^{k+1} = x^k + \hat{\alpha}_P(\hat{d}x^k + dx^k),$$
$$(\hat{y}^{k+1}, \hat{z}^{k+1}) = (y^k, z^k) + \hat{\alpha}_D(\hat{d}y^k + dy^k, \hat{d}z^k + dz^k).$$

  Before a complete predictor corrector method iteration.

- Delayed approach

$$\hat{x}^{k+1} = x^{k+1} + \bar{\alpha}_P \hat{d}x^k,$$
$$(\hat{y}^{k+1}, \hat{z}^{k+1}) = (y^{k+1}, z^{k+1}) + \bar{\alpha}_D(\hat{d}y^k, \hat{d}z^k).$$

  After a complete predictor corrector method iteration.

Criteria to applied the continued iteration

The continued iteration is applied if:

1. At least one blocking component exists;

2. $$\left\| \begin{array}{c} \hat{r}_p^{k+1} \\ \hat{r}_d^{k+1} \\ \hat{r}_a^{k+1} \end{array} \right\|_2 < \omega_1 \left\| \begin{array}{c} r_p^{k+1} \\ r_d^{k+1} \\ r_a^{k+1} \end{array} \right\|_2,$$

where $\omega_1 \in (0, 1)$.

Numerical Experiments

The continued iteration in the two approaches are implemented in C language and incorporated into the PCx code.

The numerical experiments are performed in a Intel Core i7 processor, 8 GB RAM and Linux Operating System.

In addition, the continued iteration is not applied in the last iterations, since by the experiments it does not presents good behavior near an optimal solution.

Netlib set of problems

Problem	Dimension	
	Rows	Columns
DFL001	5984	12143
MAROS-R7	2152	7440
PILOT87	1971	6373
STOCFOR3	15362	22228

Kennington set of problems

Problem	Dimension	
	Rows	Columns
CRE-B	5336	36382
CRE-D	4102	28601
KEN11	10085	16740
KEN13	22534	36561
KEN18	78862	128434
OSA-07	1081	25030
OSA-14	2300	54760
OSA-30	4313	104337
OSA-60	10243	243209

Kennington set of problems

Problem	Dimension	
	Rows	Columns
PDS06	9156	28472
PDS10	15648	48780
PDS20	38722	106180
PDS30	47968	156042
PDS40	64276	214385
PDS50	80339	272513
PDS60	96514	332862
PDS70	111896	386238
PDS80	126120	430800
PDS90	139752	471538
PDS100	152300	498530

Qaplib set of problems

Problem	Dimension	
	Rows	Columns
CHR25A	8149	15325
CHR22B	5587	10417
ELS19	4350	13186
KRA30A	18059	85725
KRA30B	18059	85725
SCR15	2234	6210
SCR20	5079	15980
ROU20	7359	37640
STE36A	27683	131076
STE36B	27683	131076
STE36C	27683	131076

Methods comparison

- **EMD:** PCx with early approach and moderated direction.
- **DMD:** PCx with delayed approach and moderated direction.

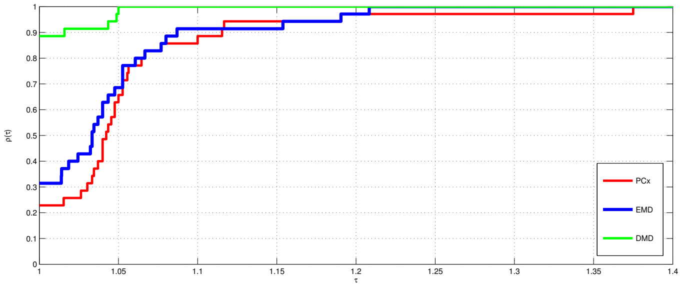
PCx without multiple centrality corrections.

Comparison among PCx, EMD and DMD

Problem	PCx		EMD			DMD		
	k	time(s)	k	ic	time(s)	k	ic	time(s)
DFL001	56	18,34	55	7	18,35	**54**	11	**18.06**
OSA-30	24	**0.96**	24	11	1.17	**23**	17	1.15
OSA-60	33	**4.02**	29	7	4.42	**24**	17	4.04
PDS20	55	149.65	54	7	148.40	**50**	14	**40.09**
PDS30	67	448.41	62	15	419.09	**60**	20	**407.18**
PDS40	66	1223.58	**63**	18	**1175.10**	64	21	1197.03
PDS70	75	6298.36	**71**	21	**5984.94**	**71**	29	5999.03
PDS80	74	8890.05	72	18	8669.56	**71**	26	**8579.01**
PDS90	76	11275.48	73	17	10853.37	**72**	28	**10730.21**
PDS100	78	12841.34	**75**	19	**12377.36**	75	23	12398.67
CHR25A	28	11.49	28	0	11.76	**26**	3	**10.97**
CHR22B	**27**	**4.59**	28	1	4.88	**27**	3	4.72
ELS19	26	33.48	26	0	34.04	**25**	6	**32.57**
KRA30A	26	3704.38	26	1	3719.07	**25**	7	**3585.78**
KRA30B	27	**3841.13**	27	0	3856.48	27	8	3861.80
STE36A	32	**13191.43**	32	0	13212.59	32	3	13278.79
STE36B	31	**12801.62**	31	0	12815.26	31	3	12872.10
STE36C	31	12808.24	31	0	12822.19	**30**	5	**12471.58**
TOTAL	1377	93742.60	1354	330	92360.26	**1315**	500	**91633.48**
REDUCTION			23		1382.34	62		2109.12

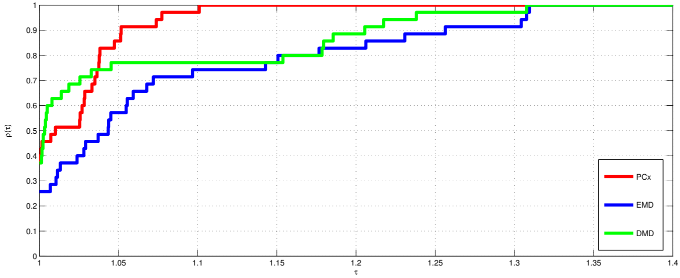
Comparison among PCx, EMD and DMD

Number of iterations



Comparison among PCx, EMD and DMD

Time



Methods comparison

- **ESD:** PCx with early approach and simple direction.
- **DSD:** PCx with delayed approach and simple direction.

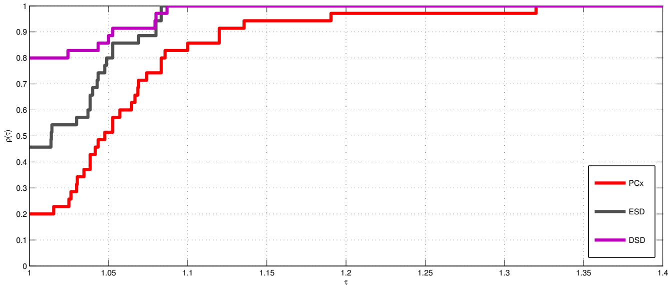
PCx without multiple centrality corrections.

Comparison among PCx, ESD and DSD

Problem	PCx		ESD			DSD		
	k	time(s)	k	ic	time(s)	k	ic	time(s)
DFL001	56	18.34	**50**	25	**16.51**	54	25	17.65
OSA-30	24	**0.96**	**23**	12	1.05	25	19	1.10
OSA-60	33	4.02	27	15	4.02	**25**	19	**3.89**
PDS20	55	149.65	54	19	147.62	**50**	14	**139.50**
PDS30	67	448.41	**59**	17	398.52	**59**	21	**398.01**
PDS40	66	1223.58	**65**	23	1209.27	**65**	22	**1202.23**
PDS70	75	6298.36	73	25	6145.71	**72**	29	**6050.92**
PDS80	74	8890.05	71	22	8550.29	**70**	31	**8425.37**
PDS90	76	11275.48	73	23	**10851.90**	**70**	28	**10408.12**
PDS100	78	12841.34	74	20	12213.38	**73**	24	**12044.59**
CHR25A	28	11.49	26	4	10.76	**25**	10	**10.34**
CHR22B	27	4.59	27	5	4.60	**26**	10	**4.42**
KRA30A	26	3704.38	26	2	3716.87	**24**	10	**3450.69**
KRA30B	27	3841.13	27	0	3855.38	**26**	10	**3719.29**
STE36A	32	**13191.43**	32	0	13251.88	32	12	13244.98
STE36B	31	**12801.62**	31	0	12848.64	31	13	12843.80
STE36C	31	12808.24	31	0	12850.14	**29**	14	**12056.52**
TOTAL	1377	93742.60	1330	437	92220.38	**1315**	590	**90195.69**
REDUCTION			47		1522.22	62		3546.91

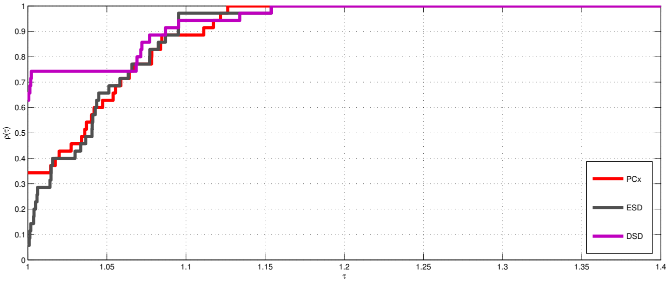
Comparison among PCx, ESD and DSD

Number of iterations



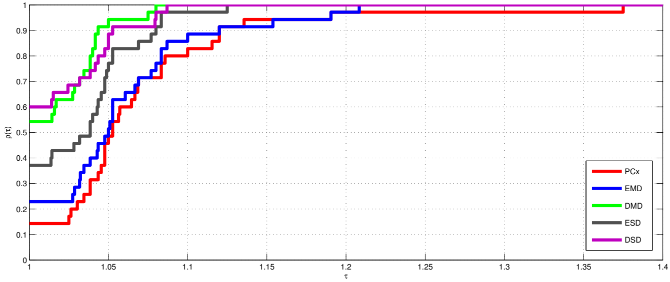
Comparison among PCx, ESD and DSD

Time



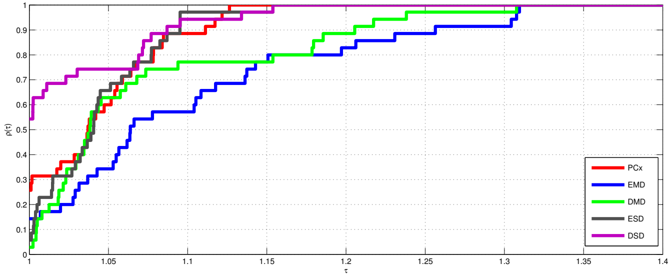
Comparison among PCx, EMD, DMD, ESD and DSD

Number of iterations



Comparison among PCx, EMD, DMD, ESD and DSD

Time



Methods comparison

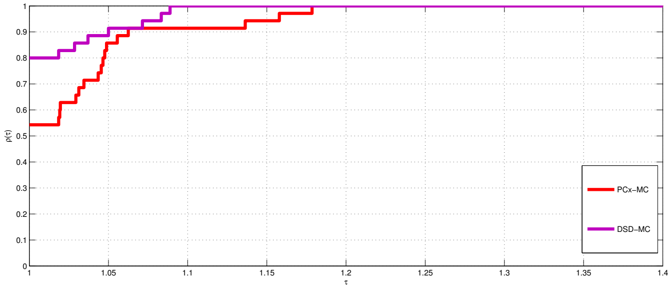
- **PCx-MC:** PCx with multiple centrality corrections.
- **DSD-MC:** PCx-MC with delayed approach and simple direction.

Comparison between PCx-MC and DSD-MC

Problem	PCx-MC			ESD-MC		
	k	mc	time(s)	k	ic	time(s)
DFL001	**45**	3	**15.26**	49	8	16.42
OSA-14	25	0	**0.39**	**22**	8	0.40
OSA-30	24	0	**0.97**	**23**	10	1.02
OSA-60	33	0	4.04	**28**	5	**3.95**
PDS20	43	4	126.81	**41**	3	**121.93**
PDS30	45	5	316.73	**43**	2	**303.24**
PDS40	50	7	961.21	50	4	**951.09**
PDS70	**54**	10	**4648.39**	56	3	4798.71
PDS80	51	10	6283.01	51	1	**6278.06**
PDS90	52	10	7875.59	**51**	1	**7718.6**
PDS100	55	10	9211.17	**54**	4	**9049.76**
CHR25A	23	3	9.85	**22**	1	**9.45**
CHR22B	22	2	3.94	**21**	1	**3.80**
KRA30A	19	10	2770.35	**18**	1	**2630.97**
KRA30B	21	10	3045.86	21	0	**3043.36**
STE36A	23	10	**9647.40**	23	0	9654.27
STE36B	23	10	**9668.41**	23	0	9669.37
STE36C	23	10	**9662.74**	23	2	9664.02
TOTAL	1092	167	69260.08	**1077**	135	**68973.01**
REDUCTION				15		287.07

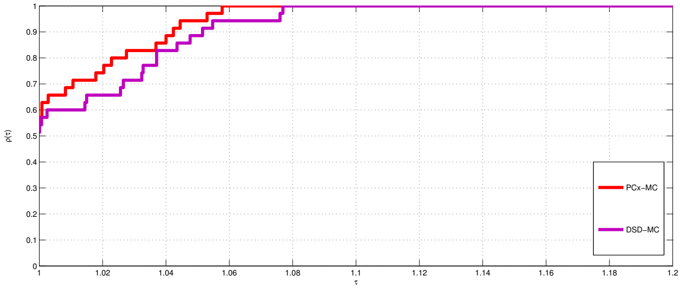
Comparison between PCx-MC and DSD-MC

Number of iterations



Comparison between PCx-MC and DSD-MC

Time



Conclusions

- The continued iteration in the early and delayed approaches are present with moderated and simple directions.
- Numerical results show that:
  - DSD had better performance them EMD, DMD, ESD and PCx.
  - DSD-MC has better performance than PCx-MC on number of iterations, however the total time difference is small.

Future Works

- Search other criteria to apply the continued iteration.
- Study new continued directions.
- More experiments with large-scale problems.