Interior point filter line search strategies for nonlinear optimization: practical behavior

M. Fernanda P. Costa†, Edite M. G. P. Fernandes‡, A. Ismael F. Vaz‡

†Department of Mathematics for Science and Technology,  mfc@mct.uminho.pt
‡Department of Production and Systems,  {emgpf;aivaz}@dps.uminho.pt
University of Minho, PORTUGAL

Introduction

The Primal-Dual Barrier Approach
- The 2-D filter methodology in IPOPT
- Our 3-D Filter line search

The Primal-Dual Interior Point Framework

Experiments: P-D barrier vs P-D interior point
- Comparison by performance profiles
- Conclusions
Introduction

The Primal-Dual Barrier Approach

The Primal-Dual Interior Point Framework

Experiments: P-D barrier vs P-D interior point

Outline

1. Introduction

2. The Primal-Dual Barrier Approach
   - The 2-D filter methodology in IPOPT
   - Our 3-D Filter line search

3. The Primal-Dual Interior Point Framework

4. Experiments: P-D barrier vs P-D interior point
   - Comparison by performance profiles
   - Conclusions

M.F.P. Costa, E.M.G.P. Fernandes and A.I.F. Vaz

Interior point filter line search strategies
Motivation

- To enforce progress towards solution, **line search methods** require
  - penalty merit function;
  - filter method.

**Filter method** - alternative to merit function - was proposed by Fletcher & Leyffer (2002):

The concept of nondominance is used to build a filter that accepts iterates that improve either the **objective function** or the **constraints violation** (2-D) instead of a combination of two measures (in a merit function).

It avoids the update of penalty parameters (associated with merit functions).
Review on filter methods in:

Primal-Dual interior point framework

- filter with 2 components (2-D): Ulbrich, Ulbrich & Vicente (2004); Silva, Ulbrich, Ulbrich & Vicente (2008).

Primal-Dual barrier approach

- filter with 2 components (2-D): IPOPT https://projects.coin-or.org/Ipopt
  Wachter & Biegler (2005a); Wachter & Biegler (2005b); Wachter & Biegler (2006).
GOAL: practical behavior of filter line search integrated into

Primal-Dual barrier methodology (P-D barrier) – IPOPT code
- the two-dimensional filter of IPOPT
- with a three-dimensional filter

Primal-Dual interior point methodology (P-D interior point)
- the two-dimensional filter of IPOPT
- with a three-dimensional filter
simulated inside the IPOPT code.
NonLinear Programming Problem

(P1) : \begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad c(x) = 0 \\
& \quad x \geq 0,
\end{align*}

- \( c_i : \mathbb{R}^n \to \mathbb{R} \), \( i = 1, \ldots, m \), \( m \leq n \) and
- \( f : \mathbb{R}^n \to \mathbb{R} \) nonlinear and twice continuously differentiable functions;
- inequality constraints can be rewritten as equality constraints by introducing slack variables.
Outline

1 Introduction

2 The Primal-Dual Barrier Approach
   - The 2-D filter methodology in IPOPT
   - Our 3-D Filter line search

3 The Primal-Dual Interior Point Framework

4 Experiments: P-D barrier vs P-D interior point
   - Comparison by performance profiles
   - Conclusions
The barrier problem

Solution of (P1) $x^*$ is obtained by computing approximate solutions of a sequence of (associated) barrier problems

$$\text{(P2)} : \min_{x \in \mathbb{R}^n} \Phi(x, \mu) \equiv f(x) - \mu \sum_{i=1}^{n} \log(x_i)$$

subject to $c(x) = 0$

for a decreasing sequence of $\mu \downarrow 0$ (maintaining $x > 0$).

- $\Phi(x, \mu)$ is the barrier function
- $\mu$ is a positive barrier parameter

Basic property:

$$x^*(\mu) = \{\arg \min \Phi(x, \mu) : c(x) = 0\} \rightarrow x^* \quad \text{as} \quad \mu \downarrow 0$$
Solving the barrier problem

This is equivalent to solving the perturbed primal-dual system \((S)\):

\[
\nabla f(x) - \nabla c(x) \delta - \lambda = 0
\]

\[
X\Lambda e - \mu e = 0
\]

\[
c(x) = 0
\]

- \(\delta\) is multiplier vector of \(c(x) = 0\) and \(\lambda\) is multiplier vector of \(x \geq 0\)
- \(\nabla f\) is the gradient of \(f\) and \((\nabla c)^T\) is the Jacobian of \(c\)
- \(X = \text{diag}(x_i)\) and \(\Lambda = \text{diag}(\lambda_i)\) are diagonal matrices; \(e \in \mathbb{R}^n\) of ones
NOTE: perturbed primal-dual system \((S)\) with \(\mu = 0\) and \(x \geq 0\) and \(\lambda \geq 0\) define KKT conditions of problem \((P1)\).

Applying Newton’s method and reducing \(\Rightarrow\) linear system to compute search directions \(\Delta x\), \(\Delta \delta\) and \(\Delta \lambda\):

\[
\begin{bmatrix}
- \left( \nabla^2_{xx} \mathcal{L} + X^{-1} \lambda \right) & \nabla c \\
\nabla c^T & 0
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta \delta
\end{bmatrix} =
\begin{bmatrix}
\nabla \Phi(x, \mu) - \nabla c \delta \\
-c(x)
\end{bmatrix}
\]

and

\[
\Delta \lambda = \mu X^{-1} e - \lambda - X^{-1} \lambda \Delta x
\]

\(\mathcal{L}\) is the Lagrangian of problem \((P1)\).
Find, at iteration $k$, $\alpha^{(k)} \in (0, \alpha_{\text{max}}^{(k)}]$ such that 
\[ x^{(k)} + \alpha_{\text{max}}^{(k)} \Delta x^{(k)} \] and 
\[ \lambda^{(k)} + \alpha_{\text{max}}^{(k)} \Delta \lambda^{(k)} \]
are maintained positive;

use backtracking
\[ \alpha^{(k)} \leftarrow \alpha_{\text{max}}^{(k)}, \frac{1}{2} \alpha_{\text{max}}^{(k)}, \frac{1}{4} \alpha_{\text{max}}^{(k)}, \ldots \]
until $x^{(k+1)}$ is acceptable;

obtain
\[ x^{(k+1)} = x^{(k)} + \alpha^{(k)} \Delta x^{(k)} \]
\[ \lambda^{(k+1)} = \lambda^{(k)} + \alpha^{(k)} \Delta \lambda^{(k)} \]
\[ \delta^{(k+1)} = \delta^{(k)} + \alpha^{(k)} \Delta \delta^{(k)}. \]
2-D filter methodology (in IPOPT)

Feasibility measure

\[ \theta_f(x) = \|c(x)\|_1 \]

and

Optimality measure

the barrier function: \( \Phi(x, \mu) \)

**Filter** is a set \( F_k \) that contains pairs \((\theta_f, \Phi)\) that are prohibited for a successful iterate in iteration \( k \).

**Filter is initialized to**

\[ F_0 \subseteq \{(\theta_f, \Phi) \in \mathbb{R}^2 : \theta_f \geq \theta_f^{\text{max}}, \Phi \geq \Phi^{\text{max}}\} \]
"acceptance condition": the iterate $x^{(k+1)}$ is acceptable if a sufficient progress (decrease) in either $\theta_f(x)$ or $\Phi(x, \mu)$, instead of a linear combination, is verified;

if feasibility is small enough and "switching conditions" are verified, then "Armijo condition" on the barrier $\Phi$ is required to accept $x^{(k+1)}$.

The filter is augmented - storing $(\theta_f(x^{(k+1)}), \Phi(x^{(k+1)}, \mu))$ to avoid cycling - if "acceptance condition" is verified.

If during backtracking an acceptable $\alpha^{(k)} \geq \alpha_{\text{min}}^{(k)}$ cannot be find, the algorithm reverts to a feasibility restoration phase.
The P-D barrier filter algorithm

1. initialize $x^{(0)} > 0$, $\mu^{(0)} > 0$, and the filter, set $\ell = 0$, $k = 0$

2. inner iterative process:
   - starting at $x^{(k)}$, compute approximate solution that is acceptable
     $x^{(k+1)} \approx \arg \min x, \mu^{(\ell)}$ subject to $c(x) = 0$
   - $k = k + 1$
   - if "dual infeasibility/primal infeasibility/centrality" > $10\mu^{(\ell)}$, go to 2.

3. decrease $\mu^{(\ell)}$ and reinitialize the filter

4. $\ell = \ell + 1$

5. if "dual infeasibility/primal infeasibility/complementarity" > $\varepsilon_{tol}$ go to 2.
New filter with 3 components

From the perturbed primal-dual system \((S)\), our proposal minimizes:

- **Feasibility measure**
  \[
  \theta_f(x) = \| c(x) \|_1
  \]

- **Centrality measure**
  \[
  \theta_c(x, \lambda, \mu) = \| X\Lambda e - \mu e \|_1
  \]

and

- **Optimality measure**
  \[
  \text{the barrier function: } \Phi(x, \mu)
  \]
A new point \((x^{(k+1)}, \lambda^{(k+1)})\) might be acceptable if sufficient progress is verified

\[
\theta_f(x^{(k+1)}) \leq (1 - \gamma_1) \theta_f(x^{(k)})
\]

or

\[
\theta_c(x^{(k+1)}, \lambda^{(k+1)}, \mu) \leq (1 - \gamma_3) \theta_c(x^{(k)}, \lambda^{(k)}, \mu)
\]

or

\[
\Phi(x^{(k+1)}, \mu) \leq \Phi(x^{(k)}, \mu) - \gamma_2 \theta_f(x^{(k)})
\]
However, if $\theta_f \leq \theta_f^{\text{min}}$ or $\theta_c \leq \theta_c^{\text{min}}$

and

$$\alpha^{(k)} \nabla \Phi(x^{(k)}, \mu)^T \Delta x^{(k)} < 0$$

and

$$\alpha^{(k)} [-\nabla \Phi(x^{(k)}, \mu)^T \Delta x^{(k)}]_{s_0} > \delta [\theta_f(x^{(k)})]^{s_f}$$

and

$$\alpha^{(k)} [-\nabla \Phi(x^{(k)}, \mu)^T \Delta x^{(k)}]_{s_0} > \delta [\theta_c(x^{(k)}, \lambda^{(k)}, \mu)]^{s_c}$$

then new iterate is acceptable if

"Armijo condition"

$$\Phi(x^{(k+1)}, \mu) \leq \Phi(x^{(k)}, \mu) + \eta \alpha^{(k)} \nabla \Phi(x^{(k)}, \mu)^T \Delta x^{(k)}.$$
Outline

1 Introduction

2 The Primal-Dual Barrier Approach
   - The 2-D filter methodology in IPOPT
   - Our 3-D Filter line search

3 The Primal-Dual Interior Point Framework

4 Experiments: P-D barrier vs P-D interior point
   - Comparison by performance profiles
   - Conclusions
Solution of (P1) $x^*$ is obtained by using the KKT conditions:

\[
\nabla f(x) - \nabla c(x) \delta - \lambda = 0
\]

\[
X \Lambda e = 0
\]

\[
c(x) = 0
\]

with $x \geq 0$ and $\lambda \geq 0$.

- $\delta$ is multiplier vector of $c(x) = 0$ and $\lambda$ is multiplier vector of $x \geq 0$. 

The interior point paradigm
Applying Newton’s method and reducing ⇒ linear system to compute search directions $\Delta x$, $\Delta \delta$ and $\Delta \lambda$:

$$
\begin{bmatrix}
-(\nabla^2_{xx}L + X^{-1}\Lambda) & \nabla c \\
\nabla c^T & 0
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta \delta
\end{bmatrix}
= 
\begin{bmatrix}
\nabla \Phi(x, \mu) - \nabla c \delta \\
-c(x)
\end{bmatrix}
$$

and

$$
\Delta \lambda = -\lambda - X^{-1}\Lambda \Delta x
$$

(E)

**NOTE:**

For some $i \in \{1, \ldots, n\}$, if $\lambda_i = 0$ (with $x_i > 0$) ⇒ $\Delta \lambda_i = 0$ and variable is fixed at the boundary along the remaining iterations.
How to overcome this problem?

1. perturb the KKT system, using a positive parameter $\mu$:

\[ \Delta \lambda = \mu X^{-1} e - \lambda - X^{-1} \Lambda \Delta x \]

2. and solve a sequence of systems $(S)$ for a sequence of values of $\mu$ converging to 0.
The P-D interior point filter algorithm

1. Initialize $x^{(0)} > 0$, $\mu^{(0)} > 0$, and the filter, set $k = 0$
2. Starting at $x^{(k)}$, compute approximate solution that is acceptable
   \[ x^{(k+1)} \approx \arg \min \Phi(x, \mu^{(k)}) \text{ subject to } c(x) = 0 \]
3. Decrease $\mu^{(k)}$ and reinitialize the filter
4. $k = k + 1$
5. If "dual infeasibility/primal infeasibility/centrality" $> \varepsilon_{tol}$ go to 2.
Outline

1. Introduction

2. The Primal-Dual Barrier Approach
   - The 2-D filter methodology in IPOPT
   - Our 3-D Filter line search

3. The Primal-Dual Interior Point Framework

4. Experiments: P-D barrier vs P-D interior point
   - Comparison by performance profiles
   - Conclusions
Computing environment

- use IPOPT code (version 3.5.5 with linear solver MA27)
- IPOPT run with default options, namely second order correction and $\varepsilon_{tol} = 10^{-8}$
- 290 small and medium scale constrained problems from CUTEr collection:
  - $n$ from 2 to 300
- results obtained in a computer Core2 Duo T9550 @ 2.66 GHz with 4 GB memory, running Linux Ubuntu 8.10
List of solvers in comparison

P-D barrier (the IPOPT itself)
- 2-dimensional filter – IPOPT
- 3-dimensional filter – integrated into IPOPT

P-D interior point – simulated into IPOPT
- 2-dimensional filter – IPOPT
- 3-dimensional filter – integrated into IPOPT

files in directory "Ipopt/src/Algorithm" were modified

IpBacktrackingLineSearch.hpp /.cpp
IpFilter.hpp/.cpp
IpFilterLSAcceptor.hpp/.cpp
IpIpoptCalculatedQuantities.hpp/.cpp
IpRestoConvCheck.hpp/.cpp
IpRestoFilterConvCheck.hpp/.cpp
Performance profiles

- In the comparison, we use Dolan & Moré (2002) performance profiles.

The performance profile of a solver (for a set of problems) gives the fraction of problems $\rho(\tau)$ with the "ratio($m$)" $\leq \tau$ (for $\tau \in \mathbb{R}$)

- $m$ is the metric used: number of iterations / number of function evaluations.

- The best solver for a particular problem has "ratio($m$)" $= 1$;
- "ratio($m$)" $> 1$ for the other solvers in comparison;
- The higher the $\rho$ the better the solver is.
Number of iterations of P-D barrier vs P-D interior point

⇒ P-D barrier with fewer inner iterations than P-D interior point, for most $\tau$ values
Introduction
The Primal-Dual Barrier Approach
The Primal-Dual Interior Point Framework
Experiments: P-D barrier vs P-D interior point

Functions evaluations of P-D barrier vs P-D interior point

$\Rightarrow$ P-D barrier 3-dimensional with fewer function evaluations than 2-dimensional
Another 3-D filter proposal for a P-D interior point context

Filter with three components not depending on $\mu$. From the system of the KKT conditions of problem (P1), our proposal minimizes:

- **Feasibility measure**
  $\theta_f(x) = \|c(x)\|_1$

- **Complementarity measure**
  $\theta_c(x, \lambda) = \|X\lambda e\|_1$

- **Optimality measure**
  The objective function: $f(x)$
Number of iterations of P-D interior point

Performance profiles on total number of inner iterations

⇒ in P-D interior point context: 3-D (based on KKT) more efficient than 3-D (perturbed primal-dual system)
In P-D interior point context: 3-D (based on KKT) more efficient than 3-D (perturbed primal-dual system)
Conclusions

- We analyzed the practical behavior of two different primal-dual strategies:
  
  1. primal-dual barrier method implemented in IPOPT code
  
  2. primal-dual interior point method simulated inside the IPOPT code

- we also tested and compared two different line search methodologies:
  
  the 2-D filter line search of IPOPT
  
  a 3-D filter line search incorporated into IPOPT
Conclusions

Using a set of small- and medium-scale benchmark problems:
- both (2-D (IPOPT) and proposed 3-D) filter line search methodologies have similar behavior – number of iterations and number of function evaluations;
- in general, both P-D barrier approaches are more efficient than their corresponding P-D interior point framework;
- differences are more significant in the 3-D filter version

1. Under investigation: adapt the restoration phase to new components (3-D filter) to improve efficiency in the P-D barrier approach;
2. exploring further the new 3-D in the P-D interior point context.
3. We aim to contribute with piece of code to IPOPT open source project.
References


Introduction
The Primal-Dual Barrier Approach
The Primal-Dual Interior Point Framework
Experiments: P-D barrier vs P-D interior point

Thanks for your attention

mfc@mct.uminho.pt
emgpf@dps.uminho.pt
aivaz@dps.uminho.pt