Mehrotra’s Predictor-Corrector Algorithm

John E. Mitchell

Practical refinements

There are several refinements to improve the practical performance:

- It is an infeasible interior point method. At each iteration, the primal residual $r_b := Ax^k - b \in \mathbb{R}^m$ and/or the dual residual $r_c = A^Ty^k + s^k - c \in \mathbb{R}^n$ may be nonzero. Nonetheless, the algorithm maintains $x^k > 0$ and $s^k > 0$. Infeasible interior point methods can still converge in polynomial time.

- It chooses a more aggressive step length for the predictor step.

- It doesn’t make the predictor step, but instead uses it to calculate an updated target value for $\mu$.

- It calculates the corrector step using diagonal matrices based on the iterate immediately before the predictor step.

Each iteration

Each iteration consists of the following:

1. The predictor direction is calculated by solving a modification of the Newton system that includes the residuals:

$$
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S^k & 0 & X^k
\end{bmatrix}
\begin{bmatrix}
\Delta x^a \\
\Delta y^a \\
\Delta s^a
\end{bmatrix}
= 
\begin{bmatrix}
-r_c \\
-r_b \\
-S^k X^k e
\end{bmatrix}
$$

2. Calculate the maximum possible primal and dual step lengths to maintain nonnegativity of $x$ and $s$:

$$
\alpha^P_a = \max \{ \alpha : x + \alpha \Delta x^a \geq 0 \}
$$

$$
\alpha^D_a = \max \{ \alpha : s + \alpha \Delta s^a \geq 0 \}
$$

3. Do not make the affine step. Instead, use it to calculate a new target $\mu$

$$
\mu^a_k = (x^k + \alpha^P_a \Delta x^a)^T (s^k + \alpha^D_a \Delta s^a) / n
$$

Set $\sigma = (\mu^a_k / \mu_k)^3$.

4. Find the corrector direction by solving

$$
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S^k & 0 & X^k
\end{bmatrix}
\begin{bmatrix}
\Delta x^c \\
\Delta y^c \\
\Delta s^c
\end{bmatrix}
= 
\begin{bmatrix}
0 \\
0 \\
\sigma \mu_k e - \Delta S^a \Delta X^a e
\end{bmatrix}
$$
5. Combine the two directions:

\[(\Delta x^k, \Delta y^k, \Delta s^k) = (\Delta x^a, \Delta y^a, \Delta s^a) + (\Delta x^c, \Delta y^c, \Delta s^c).\]

6. Choose primal and dual steplengths, modified to stay strictly feasible:

\[
\begin{align*}
\alpha^P_a &= \min\{1, 0.99 \max\{\alpha : x + \alpha \Delta x^a \geq 0\}\} \\
\alpha^D_a &= \min\{1, 0.99 \max\{\alpha : s + \alpha \Delta s^a \geq 0\}\}
\end{align*}
\]

7. Update iterate:

\[(x^{k+1}, y^{k+1}, s^{k+1}) \leftarrow (x^k + \alpha^P \Delta x^k, y^k + \alpha^D \Delta y^k, s^k + \alpha^D \Delta s^k).\]

Update \(\mu_{k+1} = (x^{k+1})^T s^{k+1}/n.\)

**Trajectories**

The system of nonlinear equalities includes \(XSe = \sigma e.\) We pick a direction

\[\Delta x = \Delta x^a + \Delta x^c, \quad \Delta s = \Delta s^a + \Delta s^c.\]

 Ideally, for each component, we would like the direction to satisfy

\[(x_i + \Delta x_i)(s_i + \Delta s_i) = \sigma \mu,
\]

or equivalently

\[x_i \Delta s_i + s_i \Delta x_i = \sigma \mu - x_i s_i - \Delta x_i \Delta s_i.\]

The quadratic cross term \(\Delta x_i \Delta s_i\) is ignored in the standard Newton system. Because of the way (1) and (2) are constructed, the direction satisfies the system

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S^k & 0 & X^k
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta s
\end{bmatrix}
= 
\begin{bmatrix}
-r_c \\
r_b \\
\sigma \mu e - XSe - \Delta S^a \Delta X^a e
\end{bmatrix}
\]

Thus, the process in effect approximates the quadratic term with the affine value of the term.

We could define a **trajectory** by considering taking a step of infinitesimal length in the primal-dual affine direction, and we can follow this trajectory to the optimal solution. Thus, the affine (or predictor) direction can be regarded as a linearization of this trajectory. The predictor-corrector direction is attempting to follow a quadratic approximation to the trajectory.

Higher order approximations have been investigated.