Introduction

The predictor-corrector method of Mizuno, Todd, and Ye first takes a “predictor” step, which is in a primal-dual affine direction that ignores centering. The step length is selected so that the iterates stay somewhat centered. A “corrector” step is then taken, which is in a pure centering direction and brings the iterate back towards the central trajectory. The algorithm enjoys quadratic convergence to optimality. We work with the standard primal-dual pair with the dual slacks written out explicitly:

\[
\min_{x \in \mathbb{R}^n} c^T x \quad \text{subject to} \quad Ax = b \quad (P) \quad \max_{y \in \mathbb{R}^m, s \in \mathbb{R}^n} b^T y \quad \text{subject to} \quad A^T y + s = c \quad (D) \quad x \geq 0 \quad s \geq 0
\]

The matrix \( A \in \mathbb{R}^{m \times n} \) and all vectors are dimensioned appropriately. Finding the Newton direction requires solving the system

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S & 0 & X
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta s
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
\sigma \mu e - SXe
\end{bmatrix}, \tag{1}
\]

where \( \mu = x^T s/n \). Taking \( \sigma = 0 \) gives the primal-dual affine scaling direction. Taking \( \sigma = 1 \) gives the centering direction. Let the set of strictly feasible primal-dual solutions be denoted

\[ \mathcal{F}^0 := \{ (x, y, s) : Ax = b, x > 0, A^T y + s = c, s > 0 \} \]

The \( \mathcal{N}_2(\theta) \) neighborhood of the central path is

\[ \mathcal{N}_2(\theta) := \{ (x, y, s) \text{ feasible} : \| XSe - \mu e \|_2 \leq \theta \mu \} \]

for \( 0 \leq \theta \leq 1 \).

The predictor step

Assume we have a point \((x^k, y^k, s^k) \in \mathcal{F}_0\). Set \( \mu_k = (x^k)^T s^k/n \). Let \( \beta \in (0, 0.5) \). Assume further that \((x^k, y^k, s^k) \in \mathcal{N}_2(\beta) \). Find the primal-dual affine direction by solving (1) with \( \sigma = 0 \). Choose a steplength \( \theta_k \) and update to

\[ (\hat{x}^k, \hat{y}^k, \hat{s}^k) \leftarrow (x^k, y^k, s^k) + \theta_k (\Delta x^k, \Delta y^k, \Delta s^k), \]

with the step length chosen to ensure that \((\hat{x}^k, \hat{y}^k, \hat{s}^k) \in \mathcal{N}_2(2\beta)\), a slightly larger neighborhood. We get \( \hat{\mu}_k = (\hat{x}^k)^T s^k/n = (1 - \theta_k) \mu_k \) (exercise).
The corrector step

We now solve (1) with the updated iterate \((\hat{x}^k, \hat{y}^k, \hat{s}^k)\) and with \(\sigma = 1\). This is a pure centering step, with direction \((\Delta x^k, \Delta y^k, \Delta s^k)\). We update using a step of length 1, so the new iterate is

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

We get \(\mu_{k+1} = (x^{k+1})^T s^{k+1}/n = \hat{\mu}_k = (1 - \theta_k)\mu_k\) (exercise).

Quadratic convergence

Note that we have

\[\frac{\mu_{k+1}}{\mu_k} = (1 - \theta_k).\]

This is the same expression as we saw in the short step path following algorithm SPF. One major difference is that the value \(\theta\) is fixed in SPF, which leads to linear convergence. In the predictor-corrector, \(\theta\) is chosen using a line search.

- If \(1 - \theta_k \to 0\) then we have superlinear convergence.

- If there exists a positive constant \(c\) so that \(1 - \theta_k \leq c\mu_k\) for all \(k > k'\) for some finite iteration count \(k'\) then we have quadratic convergence.

It is shown in Chapter 7 of Wright (pages 135-136) that the predictor-corrector method is quadratically convergent. Further, \(\theta_k\) is bounded below by \(0.4/\sqrt{n}\) when \(\beta = 0.25\), so the algorithm is globally convergent in \(O(\sqrt{n}L)\) iterations.