The Central Path

John E. Mitchell

Optimality Conditions as a System of Nonlinear Equations

We work with the standard primal-dual pair, with the dual slacks written out explicitly:

\[
\begin{align*}
\min_{x \in \mathbb{R}^n} & \quad c^T x \\
\text{subject to} & \quad Ax = b \quad (P) \\
& \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
\max_{y \in \mathbb{R}^m, s \in \mathbb{R}^n} & \quad b^T y \\
\text{subject to} & \quad A^T y + s = c \quad (D) \\
& \quad s \geq 0
\end{align*}
\]

We define the diagonal matrices \( X \) and \( S \) as:

\[
X := \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \quad \text{and} \quad S := \begin{bmatrix} s_1 & \cdots & s_n \end{bmatrix}.
\]

The optimality conditions for \((P)\) and \((D)\) require \( x \geq 0, s \geq 0, \) and

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S & 0 & X
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
s
\end{bmatrix}
= \begin{bmatrix}
c \\
b \\
0
\end{bmatrix},
\]

a nonlinear system of equations. Let the set of strictly feasible primal-dual solutions be denoted

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

Assume we have a strictly feasible point in \( \mathcal{F}^0 \). We can define a direction \((\Delta x, \Delta y, \Delta s)\) by using a Newton step:

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S & 0 & X
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta s
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
-SX e
\end{bmatrix}.
\]

This gives the primal-dual affine direction.

The Central Path

Adding a barrier function to the objective function in \((P)\) gives the nonlinear program

\[
\begin{align*}
\min_x & \quad c^T x - \tau \sum_{i=1}^n \ln x_i \\
\text{subject to} & \quad Ax = b \quad (P(\tau)) \\
& \quad x \geq 0
\end{align*}
\]

(Note that we’ve used \( \tau \) to denote the positive barrier parameter, as in the text by Wright.)
We assume \((P)\) has an optimal solution. Since \(\tau > 0\), the objective function is strictly convex, so the problem has a unique minimizer, and this minimizer \(x\) must be strictly positive. Thus, we can write the optimality conditions as

\[
\begin{align*}
ATy + s &= c \\
Ax &= b \\
x_is_i &= \tau & \text{for } i = 1, \ldots, n \\
(x, s) &> 0
\end{align*}
\]

For each \(\tau > 0\), the solution to this system defines a point on the central path. We denote the central path by \(C\).

Most primal dual algorithms take Newton steps towards points on \(C\) with \(\tau > 0\), rather than working to solve for a particular \(\tau\).
Algorithmic Framework

The optimality conditions to \((P(\tau))\) can be written equivalently as

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
\frac{1}{2}S & 0 & \frac{1}{2}X
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
s
\end{bmatrix}
= \begin{bmatrix}
c \\
b \\
\tau e
\end{bmatrix},
\]

together with \((x, s) > 0\). Given a strictly feasible point in \(\mathcal{F}^0\), we can define a direction \((\Delta x, \Delta y, \Delta s)\) by using a Newton step:

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S & 0 & X
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta s
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
\tau e - SXe
\end{bmatrix}.
\]

Note that the duality gap is

\[
c^T x - b^T y = (A^T y + s)^T x - (Ax)^T y = (A^T y)^T x + s^T x - (A^T y)^T x = s^T x.
\]

We define the duality measure

\[
\mu := \frac{s^T x}{n} = \frac{1}{n} \sum_{i=1}^{n} s_i x_i,
\]

the duality gap averaged over the components. We also define a centering parameter \(\sigma \in [0, 1]\). The target value for \(\tau\) is then taken to be the product \(\sigma \mu\), giving the generic step equation:

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

Taking \(\sigma = 0\) gives the affine scaling direction. Taking \(\sigma = 1\) gives the centering direction. This is the direction obtained we we try to solve the system of equations for the point on the central path with \(\tau = \mu\).
This leads to an **algorithmic framework**:

- **Given** \((x^0, y^0, s^0) \in \mathcal{F}^0\).
- **For** \(k = 0, 1, 2, \ldots\)
  
  - Solve
    \[
    \begin{bmatrix}
    0 & A^T & I \\
    A & 0 & 0 \\
    S^k & 0 & X^k
    \end{bmatrix}
    \begin{bmatrix}
    \Delta x^k \\
    \Delta y^k \\
    \Delta s^k
    \end{bmatrix}
    =
    \begin{bmatrix}
    0 \\
    0 \\
    \sigma_k \mu_k e - S^k X^k e
    \end{bmatrix}.
    \]

  - **Update**
    \[(x^{k+1}, y^{k+1}, s^{k+1}) \leftarrow (x^k, y^k, s^k) + \alpha_k (\Delta x^k, \Delta y^k, \Delta s^k)\]
    choosing steplength \(\alpha_k\) so that \((x^{k+1}, s^{k+1}) > 0\).

**end for**

### Linear Algebra

**Lemma 1.** If \(A\) has full row rank and if \((x, s) > 0\) then the matrix

\[
M := \begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S & 0 & X
\end{bmatrix}
\]

is square and invertible.

**Proof.** We can block row reduce the system:

\[
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S & 0 & X
\end{bmatrix}
\longrightarrow
\begin{bmatrix}
0 & A^T & I \\
A & 0 & 0 \\
S & 0 & X
\end{bmatrix}
\longrightarrow
\begin{bmatrix}
0 & A^T & I \\
0 & AS^{-1}XAT & 0 \\
S & -XAT & 0
\end{bmatrix}
\]

The matrix \(AS^{-1}XAT\) is square and invertible, since \(A\) has full row rank and the two diagonal matrices have only positive diagonal entries. It follows that \(M\) is invertible. \(\square\)

It follows from the Lemma that the Newton step in the algorithm is well-defined for any \(\sigma \in [0, 1]\), leading to a unique solution \((\Delta x, \Delta y, \Delta s)\).

Note that the matrix product \(S^{-1}X\) is a diagonal matrix with positive diagonal entries. We could define the positive diagonal matrix \(D\) so that \(D^2 = S^{-1}X\). We can then write:

\[
\begin{align*}
\Delta y & = (AD^2A^T)^{-1}(\sigma \mu AS^{-1}e - b) \\
\Delta s & = -A^T \Delta y \\
\Delta x & = \sigma \mu S^{-1}e - x + D^2 A^T \Delta y = -DP_{AD}D \left(s - \sigma \mu X^{-1}e\right)
\end{align*}
\]