NOTE ON THE FORMULATION OF A SHIFTED PRIMAL-DUAL PENALTY-BARRIER METHOD FOR NONLINEAR OPTIMIZATION

Philip E. Gill∗ Vyacheslav Kungurtsev† Daniel P. Robinson‡

UCSD Center for Computational Mathematics
Technical Report CCoM-19-4
March 1, 2019

Abstract
The modified Newton equations for the shifted primal-dual penalty-barrier method are derived for a nonlinearly constrained problems with various constraint types.

Key words. Nonlinear programming, nonlinear constraints, augmented Lagrangian methods, interior methods, path-following methods, regularized methods, primal-dual methods.

AMS subject classifications. 49J20, 49J15, 49M37, 49D37, 65F05, 65K05, 90C30
1. Introduction

This note derives the shifted primal-dual penalty-barrier merit functions and associated path-following equations for an optimization problem with constraints written in eight different ways:

\[
\begin{align*}
\text{minimize} & \quad f(x) \quad \text{subject to} \quad Ax = b, \quad \ell \leq x \leq u, \\
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad s \geq 0, \\
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad L_X s = h_X, \quad L_F s \geq 0, \\
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad Ax = b, \quad x \geq 0, \quad s \geq 0, \\
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad L_X s = h_X, \quad \ell \leq L_F s \leq u, \\
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad s \geq 0, \quad E_X x = b_X, \quad \ell \leq E_F x \leq u, \\
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad \ell_X \leq x \leq u_X, \quad \ell^s \leq s \leq u^s. \\
\end{align*}
\]

The equations and merit function for this general problem are given in Section 10.6.

Convergence results for Problem (1.2) are given by Gill, Kungurtsev and Robinson [1].
Notation. Given vectors $x$ and $y$, the vector consisting of $x$ augmented by $y$ is denoted by $(x, y)$. The subscript $i$ is appended to vectors to denote the $i$th component of that vector, whereas the subscript $k$ is appended to a vector to denote its value during the $k$th iteration of an algorithm, e.g., $x_k$ represents the value for $x$ during the $k$th iteration, whereas $[x_k]_i$ denotes the $i$th component of the vector $x_k$. Given vectors $a$ and $b$ with the same dimension, the vector with $i$th component $a_i b_i$ is denoted by $a \cdot b$. Similarly, $\min(a, b)$ is a vector with components $\min(a_i, b_i)$. The vector $e$ denotes the column vector of ones, and $I$ denotes the identity matrix. The dimensions of $e$ and $I$ are defined by the context. The vector two-norm or its induced matrix norm are denoted by $\| \cdot \|$. The vector $g(x)$ is used to denote $\nabla f(x)$, the gradient of $f(x)$. The vector $g(x)$ is used to denote $\nabla f(x)$, the gradient of $f(x)$. The matrix $J(x)$ denotes the $m \times n$ constraint Jacobian, which has $i$th row $\nabla c_i(x)^T$. Given a Lagrangian function $L(x, y) = f(x) - c(x)^T y$ with $y$ a $m$-vector of dual variables, the Hessian of the Lagrangian with respect to $x$ is denoted by $H(x, y) = \nabla^2 f(x) - \sum_{i=1}^m y_i \nabla^2 c_i(x)$. 

1. Introduction
2. Linear Equality Constraints and Upper and Lower Bounds on the Variables

Next we consider methods for an optimization problem with linear equality constraints and upper and lower bounds on the variables.

2.1. Problem statement and optimality conditions

The problem has the form

\[
\begin{align*}
\text{minimize } f(x) & \quad \text{subject to } Ax = b, \quad \ell \leq x \leq u, \\
& \quad x^* - \ell \geq 0, \quad u - x^* \geq 0,
\end{align*}
\]

(2.1)

where \( f : \mathbb{R}^n \mapsto \mathbb{R} \) is twice-continuously differentiable. The first-order KKT conditions for this problem are

\[
\begin{align*}
Ax^* &= b, \\
x^* - \ell &= 0, \\
u - x^* &= 0, \\
g(x^*) - z_1^* + z_2^* - A^Tv^* &= 0, \\
z_1^* &\geq 0, \\
z_2^* &\geq 0, \\
z_1^* \cdot (x^* - \ell) &= 0, \\
z_2^* \cdot (u - x^*) &= 0.
\end{align*}
\]

(2.2)

The \( n \)-vectors \( z_1^* \) and \( z_2^* \) may be interpreted as Lagrange multipliers for the inequality constraints \( x - \ell \geq 0 \) and \( u - x \geq 0 \), respectively. The vector \( v^* \) is the multiplier vector for the linear equality constraints.

2.2. The path-following equations

Let \( z_1^E \) and \( z_2^E \) denote \( n \)-vectors of nonnegative estimates of the Lagrange multipliers for the inequality constraints \( x - \ell \geq 0 \) and \( u - x \geq 0 \), respectively. Let \( v^E \) denote an estimate of \( v^* \). Given a small positive scalars \( \mu^A \) and \( \mu^B \), consider the perturbed optimality conditions

\[
\begin{align*}
Ax - b &= \mu^A(v^E - v), \\
x - \ell &= 0, \\
u - x &= 0, \\
g(x) - z_1 + z_2 - A^Tv &= 0, \\
z_1 &\geq 0, \\
z_2 &\geq 0, \\
z_1 \cdot (x - \ell) &= \mu^B(z_1^E - z_1), \\
z_2 \cdot (u - x) &= \mu^B(z_2^E - z_2).
\end{align*}
\]

(2.3)

Consider the primal-dual path parameterized by \( \mu^a \) consisting of points \((x, z_1, z_2)\) such that \( F(x, v, z_1, z_2; \mu^a, v^E, z_1^E, z_2^E) = 0 \), where

\[
F(x, v, z_1, z_2; \mu^a, v^E, z_1^E, z_2^E) = \begin{pmatrix}
g(x) - z_1 + z_2 - A^Tv \\
Ax - b + \mu^a(v - v^E) \\
z_1 \cdot (x - \ell) + \mu^a(z_1 - z_1^E) \\
z_2 \cdot (u - x) + \mu^a(z_2 - z_2^E)
\end{pmatrix}.
\]

(2.4)

Any zero \((x, v, z_1, z_2)\) of \( F \) that satisfies \( z_1 > 0 \) and \( z_2 > 0 \) approximates a point satisfying the optimality conditions (2.2), with the approximation becoming increasingly accurate as \( \mu^a(v - v^E) \to 0 \), \( \mu^a(z_1 - z_1^E) \to 0 \) and \( \mu^a(z_2 - z_2^E) \to 0 \). For any sequence
of \( v^e, z_1^e \) and \( z_2^e \) such that \( v^e \to v^*, z_1^e \to z_1^* \) and \( z_2^e \to z_2^* \), and it must hold that solutions \((x, v, z_1, z_2)\) of (2.3) must satisfy \( Ax - b \to 0, z_1 \cdot (x - \ell) \to 0 \) and \( z_2 \cdot (u - x) \to 0 \). This implies that a solution \((x, v, z_1, z_2)\) of (2.3) will approximate a solution of (2.2) independently of the values of \( \mu^e \) and \( \mu^a \) (i.e., it is not necessary that \( \mu^a, \mu^a \to 0 \)).

If \((x, v, z_1, z_2)\) is a given approximate zero of \( F \) such that \( x - \ell + \mu^a e > 0, u - x + \mu^a e > 0, z_1 > 0 \) and \( z_2 > 0 \), the Newton equations for the change in variables \((\Delta x, \Delta v, \Delta z_1, \Delta z_2)\) are given by

\[
\begin{pmatrix}
H & -A^T & -I & I \\
A & -\mu^a I & 0 & 0 \\
Z_1 & 0 & X_1^\mu & 0 \\
-Z_2 & 0 & 0 & X_2^\mu
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta v \\
\Delta z_1 \\
\Delta z_2
\end{pmatrix}
= -
\begin{pmatrix}
g - A^T v - z_1 + z_2 \\
Ax - b - \mu^a(v - v^e) \\
z_1 \cdot (x - \ell) + \mu^a(z_1 - z_1^*) \\
\sum_{j=1}^n \{ \mu^a[z_1^j]_j \ln([x_1^j]_j + \mu^a) + \mu^a[z_1^j]_j \ln([z_1]_j([x_1]_j + \mu^a)) - [z_1]_j([x_1]_j + \mu^a) \}
\end{pmatrix},
\]

(2.5)

where \( X_1^\mu = \text{diag}(x_j - \ell_j + \mu^a), X_2^\mu = \text{diag}(u_j - x_j + \mu^a) \), \( Z_1 = \text{diag}([z_1]_j) \), and \( Z_2 = \text{diag}([z_2]_j) \).

### 2.3. A shifted primal-dual penalty-barrier function

Problem (2.1) is equivalent to

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad Ax = b, \\
& \quad x - x_1 = \ell, \quad x_1 \geq 0, \\
& \quad x + x_2 = u, \quad x_2 \geq 0.
\end{align*}
\]

Consider the shifted primal-dual penalty-barrier problem

\[
\begin{align*}
\text{minimize} & \quad M(x, x_1, x_2, v, z_1, z_2; \mu^a, z_1^e, z_2^e) \\
\text{subject to} & \quad x - x_1 = \ell, \quad x_1 + \mu^a e > 0, \quad z_1 > 0, \\
& \quad x + x_2 = u, \quad x_2 + \mu^a e > 0, \quad z_2 > 0,
\end{align*}
\]

(2.6)

where \( M(x, x_1, x_2, v, z_1, z_2; \mu^a, \mu^a, v^e, z_1^e, z_2^e) \) is the penalty-barrier function

\[
\begin{align*}
f(x) - (Ax - b)^Tv^e + \frac{1}{2\mu^a}||Ax - b||^2 + \frac{1}{2\mu^a}||Ax - b + \mu^a(v - v^e)||^2 \\
- \sum_{j=1}^n \{ \mu^a[z_1^j]_j \ln([x_1]_j + \mu^a) + \mu^a[z_1^j]_j \ln([z_1]_j([x_1]_j + \mu^a)) - [z_1]_j([x_1]_j + \mu^a) \}
\end{align*}
\]

- \sum_{j=1}^n \{ \mu^a[z_2^j]_j \ln([x_2]_j + \mu^a) + \mu^a[z_2^j]_j \ln([z_2]_j([x_2]_j + \mu^a)) - [z_2]_j([x_2]_j + \mu^a) \}.
\]

(2.7)
Differentiating $M(x, x_1, x_2, v, z_1, z_2, w)$ with respect to $x, x_1, x_2, v, z_1,$ and $z_2$ gives

$$
\nabla M(x, x_1, x_2, v, z_1, z_2) = \begin{pmatrix}
    g - A^T v^e + \frac{1}{\mu^e} A^T (Ax - b) + \frac{1}{\mu^v} A^T (Ax - b + \mu^v (v - v^e)) \\
    z_1 - 2 \mu^e (X_1^e)^{-1} z_1^e \\
    z_2 - 2 \mu^e (X_2^e)^{-1} z_2^e \\
    Ax - b + \mu^v (v - v^e) \\
    x_1 + \mu^e z_1 \\
    x_2 + \mu^e z_2
\end{pmatrix},
$$

where $X_1^e = \text{diag}(x_1 + \mu^e)$ and $X_2^e = \text{diag}(x_2 + \mu^e)$.

Similarly,

$$
x_1 + \mu^e z_1 = Z_1^{-1} (Z_1 (x_1 + \mu^e) - \mu^e z_1^e) = Z_1^{-1} (Z_1 x_1 + \mu^e z_1 - \mu^e z_1^e) = Z_1^{-1} (z_1 \cdot (x - \ell) + \mu^e (z_1 - z_1^e)).
$$

(2.8)

Similarly,

$$
x_1 + \mu^e - \mu^e Z_1^{-1} z_1^e = Z_1^{-1} (Z_1 (x_1 + \mu^e) - \mu^e z_1^e) = Z_1^{-1} (X_1^e z_1 - \mu^e z_1^e) = Z_1^{-1} X_1^e (z_1 - \mu^e (X_1^e)^{-1} z_1^e) = D_1^e (z_1 - \pi_1^e),
$$

(2.9)

where $D_1^e = X_1^e Z_1^{-1} = Z_1^{-1} X_1^e$ and $\pi_1^e = \mu^e (X_1^e)^{-1} z_1^e$. Analogous identities hold for $x_2 + \mu^e e - \mu^e Z_2^{-1} z_2^e$.

The identities above imply that the gradient may be written in several equivalent forms

$$
\nabla M(x, x_1, x_2, z_1, z_2) = \begin{pmatrix}
    g - A^T v^e + \frac{1}{\mu^e} A^T (Ax - b) + \frac{1}{\mu^v} A^T (Ax - b + \mu^v (v - v^e)) \\
    z_1 - 2 \mu^e (X_1^e)^{-1} z_1^e \\
    z_2 - 2 \mu^e (X_2^e)^{-1} z_2^e \\
    Ax - b + \mu^v (v - v^e) \\
    x_1 + \mu^e e - \mu^e Z_1^{-1} z_1^e \\
    x_2 + \mu^e e - \mu^e Z_2^{-1} z_2^e
\end{pmatrix} = \begin{pmatrix}
    g - A^T (2v^e + \frac{1}{\mu^e} (Ax - b) - v) \\
    (X_1^e)^{-1} (z_1 \cdot x_1 + \mu^e z_1^e + \mu^e (z_1 - z_1^e)) \\
    (X_2^e)^{-1} (z_2 \cdot x_2 + \mu^e z_2^e + \mu^e (z_2 - z_2^e)) \\
    -\mu^v (v^e - \frac{1}{\mu^e} (Ax - b) - v) \\
    Z_1^{-1} (z_1 \cdot x_1 + \mu^e (z_1 - z_1^e)) \\
    Z_2^{-1} (z_2 \cdot x_2 + \mu^e (z_2 - z_2^e))
\end{pmatrix} = \begin{pmatrix}
    g - A^T (2\pi^e - v) \\
    -2\pi_1^e - z_1 \\
    -2\pi_2^e - z_2 \\
    -D_1^e (\pi^e - v) \\
    -D_1^e (\pi_1^e - z_1) \\
    -D_2^e (\pi_2^e - z_2)
\end{pmatrix},
$$

where $g = A^T (2v^e - v)$. The identities above imply that the gradient may be written in several equivalent forms.
where

\[ D_A = \mu^A I, \quad \pi^A = v^e - \frac{1}{\mu^A} (Ax - b), \]  

\[ D_z^1 = X^1_z Z_1^{-1}, \quad \pi_z^1 = \mu^B (X^1_z)^{-1} z_1^e, \]  

\[ D_z^2 = X^2_z Z_2^{-1}, \quad \pi_z^2 = \mu^B (X^2_z)^{-1} z_2^e. \]  

(2.10a) (2.10b) (2.10c)

Similarly, the Hessian of \( M(x, x_1, x_2, v, z_1, z_2) \) is given by

\[
\begin{pmatrix}
H + \frac{2}{\mu^A} A^T A & 0 & 0 & 0 & 0 & 0 \\
0 & 2\mu^B (X^1_1)^{-2} Z_1^e & 0 & 0 & I & 0 \\
0 & 0 & 2\mu^B (X^2_2)^{-2} Z_2^e & 0 & 0 & I \\
0 & 0 & 0 & D_A & 0 & 0 \\
0 & I & 0 & 0 & \mu^B Z^{-2}_1 Z_1^e & 0 \\
0 & 0 & I & 0 & 0 & \mu^B Z^{-2}_2 Z_2^e \\
\end{pmatrix},
\]

where \( H = \nabla^2 f \). Substituting \( \mu^B Z_1^e = X^1_z \Pi_1^z \) and \( \mu^B Z_2^e = X^2_z \Pi_2^z \) from (2.10) gives the Hessian

\[
\begin{pmatrix}
H + \frac{2}{\mu^A} A^T A & 0 & 0 & A^T & 0 & 0 \\
0 & 2(X^1_1)^{-1} \Pi_1^z & 0 & 0 & I & 0 \\
0 & 0 & 2(X^2_2)^{-1} \Pi_2^z & 0 & 0 & I \\
A & 0 & 0 & D_A & 0 & 0 \\
0 & I & 0 & 0 & X^1_z Z_1^{-2} \Pi_1^z & 0 \\
0 & 0 & I & 0 & 0 & X^2_z Z_2^{-2} \Pi_2^z \\
\end{pmatrix}.
\]

2.4. Derivation of the shifted primal-dual penalty-barrier direction

The primal-dual penalty-barrier problem may be written in the form

\[
\text{minimize } M(p) \text{ subject to } Cp = b_c,
\]

(2.11)

where

\[ I = \{ p : p = (x, x_1, x_2, v, z_1, z_2), \text{ with } x_1 + \mu^e e > 0, x_2 + \mu^e e > 0, z_1 > 0, z_2 > 0 \}, \]

and

\[ C = \begin{pmatrix} I & -I & 0 & 0 & 0 & 0 \\ I & 0 & I & 0 & 0 & 0 \end{pmatrix}, \quad \text{and } \quad b_c = \begin{pmatrix} \ell \\ u \end{pmatrix}. \]
Let \( p \in \mathcal{I} \) be given. As in the bounded slack case, assume that \( p \) is not necessarily feasible for the linear constraints, i.e., it may not hold that \( x - x_1 = \ell \) and \( x + x_2 = u \), in which case \( b_C - C p \) may not be zero. The Newton direction \( \Delta p \) is given by the solution of the subproblem

\[
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T \nabla^2 M(p) \Delta p \quad \text{subject to} \quad C \Delta p = b_C - C p. \quad (2.12)
\]

However, instead of solving (2.12), we define a linearly constrained modified Newton method by approximating the Hessian \( \nabla^2 M(x, x_1, x_2, v, z_1, z_2) \) by a matrix \( B(x, x_1, x_2, v, z_1, z_2) \). Consider the matrix defined by replacing \( \pi Z_1 \) by \( z_1 \) and \( \pi Z_2 \) by \( z_2 \) everywhere in the matrix \( \nabla^2 M(x, x_1, x_2, v, z_1, z_2) \). This gives an approximate Hessian

\[
B(x, x_1, x_2, v, z_1, z_2) = \begin{pmatrix}
H + \frac{\mu}{2} A^T A & 0 & 0 & A^T & 0 & 0 \\
0 & 2(X_1^\mu)^{-1} Z_1 & 0 & 0 & I & 0 \\
0 & 0 & 2(X_2^\mu)^{-1} Z_2 & 0 & 0 & I \\
A & 0 & 0 & D_A & 0 & 0 \\
0 & I & 0 & 0 & X_1^\mu Z_1^{-1} & 0 \\
0 & 0 & I & 0 & 0 & X_2^\mu Z_2^{-1}
\end{pmatrix}.
\]

The definitions of \( D_A, D_1^\mu \) and \( D_2^\mu \) (2.10) may be used to write \( B(x, x_1, x_2, v, z_1, z_2) \) in the form

\[
\begin{pmatrix}
H + 2A^T D_A^{-1} A & 0 & 0 & A^T & 0 & 0 \\
0 & 2(D_1^\mu)^{-1} & 0 & 0 & I & 0 \\
0 & 0 & 2(D_2^\mu)^{-1} & 0 & 0 & I \\
A & 0 & 0 & D_A & 0 & 0 \\
0 & I & 0 & 0 & D_1^\mu & 0 \\
0 & 0 & I & 0 & 0 & D_2^\mu
\end{pmatrix}.
\]

Given \( B(p) = B(x, x_1, x_2, v, z_1, z_2) \), a modified Newton direction is given by the solution of the QP subproblem

\[
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T B(p) \Delta p \quad \text{subject to} \quad C \Delta p = b_C - C p. \quad (2.13)
\]

Let \( N \) denote a matrix whose columns form a basis for \( \text{null}(C) \), i.e., the columns of \( N \) are linearly independent and \( CN = 0 \). The vector

\[
\Delta p_0 = \begin{pmatrix}
0 \\
-(\ell - x + x_1) \\
(u - x - x_2) \\
0 \\
0 \\
0
\end{pmatrix} \triangleq \begin{pmatrix}
0 \\
-r_L \\
r_U \\
0 \\
0 \\
0
\end{pmatrix} \quad (2.14)
\]
satisfies $C\Delta p_0 = b_c - Cp$, and every feasible $\Delta p$ may be written in the form
$$\Delta p = \Delta p_0 + Nd.$$  

This implies that $d$ satisfies the reduced equations
$$N^T B(p) Nd = -N^T (\nabla M(p) + B(p) \Delta p_0).$$

Consider the null-space basis defined from the columns of
\begin{equation}
N = \begin{pmatrix}
I & 0 & 0 & 0 \\
I & 0 & 0 & 0 \\
-I & 0 & 0 & 0 \\
0 & I & 0 & 0 \\
0 & 0 & I & 0 \\
0 & 0 & 0 & I \\
\end{pmatrix}.
\end{equation}

The definition of $N$ of (2.15) gives the reduced Hessian
$$N^T B(p) N = \begin{pmatrix}
H + 2A^T D_A^{-1} A + 2((D_1^z)^{-1} + (D_2^z)^{-1}) & A^T & I & -I \\
A & D_A & 0 & 0 \\
I & 0 & D_1^z & 0 \\
-I & 0 & 0 & D_2^z \\
\end{pmatrix}.$$

Similarly,
$$N^T \nabla M(p) = N^T \begin{pmatrix}
g - A^T (2\pi^A - v) \\
-(2\pi_1^z - z_1) \\
-(2\pi_2^z - z_2) \\
-D_A(\pi^A - v) \\
-D_1^z(\pi_1^z - z_1) \\
-D_2^z(\pi_2^z - z_2) \\
\end{pmatrix} = \begin{pmatrix}
g - A^T (2\pi^A - v) - (2\pi_1^z - z_1) + (2\pi_2^z - z_2) \\
-D_A(\pi^A - v) \\
-D_1^z(\pi_1^z - z_1) \\
-D_2^z(\pi_2^z - z_2) \\
\end{pmatrix},$$

and
$$N^T B(p) \Delta p_0 = N^T \begin{pmatrix}
0 \\
-2(D_1^z)^{-1} r_L \\
2(D_2^z)^{-1} r_U \\
0 \\
-r_L \\
r_L \\
\end{pmatrix} = \begin{pmatrix}
-2((D_1^z)^{-1} r_L + (D_2^z)^{-1} r_U) \\
0 \\
-r_L \\
r_U \\
\end{pmatrix},$$
where $r_L = \ell - x + x_1$ and $r_U = u - x - x_2$. This gives the reduced gradient

\[
N^T(\nabla M(p) + B(p)\Delta p_0) = \begin{bmatrix} g - A^T(2\pi^2 - v) - (2\pi_1^2 - z_1) + (2\pi_2^2 - z_2) - 2((D_1^2)^{-1}r_L + (D_2^2)^{-1}r_U) \\ -D_\lambda(\pi^2 - v) \\ -D_1^2(\pi_1^2 - z_1) - r_L \\ -D_2^2(\pi_2^2 - z_2) + r_U \end{bmatrix}.
\]

The reduced modified Newton equations $N^T B(p)Nd = -N^T(\nabla M(p) + B(p)\Delta p_0)$ are then

\[
\begin{pmatrix} H + 2A^TD^{-1}A + 2((D_1^2)^{-1} + (D_2^2)^{-1}) & A^T & I & -I \\ A & D_\lambda & 0 & 0 \\ I & 0 & D_1^2 & 0 \\ -I & 0 & 0 & D_2^2 \end{pmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{bmatrix} = \begin{bmatrix} g - A^T(2\pi^2 - v) - (2\pi_1^2 - z_1) + (2\pi_2^2 - z_2) - 2((D_1^2)^{-1}r_L + (D_2^2)^{-1}r_U) \\ -D_\lambda(\pi^2 - v) \\ -D_1^2(\pi_1^2 - z_1) - r_L \\ -D_2^2(\pi_2^2 - z_2) + r_U \end{bmatrix}.
\]

Given any nonsingular matrix $R$, the direction $d$ satisfies

\[
RN^T B(p)Nd = -RN^T(\nabla M(p) + B(p)\Delta p_0).
\]

In particular, as $Z_1$ and $Z_2$ are positive definite, the block upper-triangular matrix

\[
R = \begin{pmatrix} I & -2A^TD^{-1} & -2(\pi_1^2)^{-1} & 2(\pi_2^2)^{-1} \\ I & 0 & 0 & Z_1 \\ 0 & 0 & Z_2 \end{pmatrix}
\]

is nonsingular, with

\[
RN^T B(p)N = \begin{pmatrix} H & -A^T & -I & I \\ A & D_\lambda & 0 & 0 \\ Z_1 & 0 & Z_1^2 & 0 \\ -Z_2 & 0 & 0 & Z_2^2 \end{pmatrix} = \begin{pmatrix} H & -A^T & -I & I \\ A & D_\lambda & 0 & 0 \\ Z_1 & 0 & X_1^a & 0 \\ -Z_2 & 0 & 0 & X_2^a \end{pmatrix}.
\]
Also, \(RNT(\nabla M(p) + B(p)\Delta p_0)\) is given by

\[
\begin{pmatrix}
    I & -2ATD^{-1} & -2(D_1^{-1}) & 2(D_2^{-1})
    \\
    0 & 0 & 0 & 0
    \\
    Z_1 & 0 & 0 & 0
    \\
    Z_2 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
g - A^T(2\pi^a - v) - (2\pi^z - z_1) + (2\pi^z - z_2) - 2((D_1)^{-1}r + (D_2)^{-1}r_v)
    \\
-D_1(\pi^a - v)
    \\
-D_1^2(\pi_1^z - z_1) - r_L
    \\
-D_2^2(\pi_2^z - z_2) + r_u
\end{pmatrix}
\]

This gives the following unsymmetric equations for \(d\)

\[
\begin{pmatrix}
    H & -A^T & -I & I
    \\
    A & D \pi & 0 & 0
    \\
    Z_1 & 0 & X_1^\mu & 0
    \\
    -Z_2 & 0 & 0 & X_2^\mu
\end{pmatrix}
\begin{pmatrix}
d_1
    \\
d_2
    \\
d_3
    \\
d_4
\end{pmatrix}
= -
\begin{pmatrix}
g - A^Tv - z_1 + z_2
    \\
Ax - b - \mu^a(v - v^a)
    \\
\pi_1 \cdot (x - \ell) + \mu^z(z_1 - z_1^z)
    \\
\pi_2 \cdot (u - x) + \mu^z(z_2 - z_2^z)
\end{pmatrix}.
\]

Then, \(2.14\) implies that

\[
\begin{pmatrix}
\Delta x
    \\
\Delta x_1
    \\
\Delta x_2
    \\
\Delta u
    \\
\Delta z_1
    \\
\Delta z_2
\end{pmatrix}
= \Delta p = \Delta p_0 + N d
= \begin{pmatrix}
d_1
    \\
(d_1 - r_L)
    \\
d_2
    \\
d_3
    \\
d_4
\end{pmatrix}
.
\]

These identities allow us to write equations \(2.16\) in the form

\[
\begin{pmatrix}
    H & -A^T & -I & I
    \\
    A & D \pi & 0 & 0
    \\
    Z_1 & 0 & X_1^\mu & 0
    \\
    -Z_2 & 0 & 0 & X_2^\mu
\end{pmatrix}
\begin{pmatrix}
\Delta x
    \\
\Delta v
    \\
\Delta z_1
    \\
\Delta z_2
\end{pmatrix}
= -
\begin{pmatrix}
g - A^Tv - z_1 + z_2
    \\
Ax - b - \mu^a(v - v^a)
    \\
\pi_1 \cdot (x - \ell) + \mu^z(z_1 - z_1^z)
    \\
\pi_2 \cdot (u - x) + \mu^z(z_2 - z_2^z)
\end{pmatrix},
\]

from which we can compute \(\Delta x_1 = \Delta x - (\ell - x + x_1)\) and \(\Delta x_2 = -\Delta x + (u - x - x_2)\). If \(x_1\) and \(x_2\) satisfy \(x - x_1 = \ell\) and \(x + x_2 = u\) (i.e., they are feasible for \(2.6\)), then \(\Delta x_1 = \Delta x\) and \(\Delta x_2 = -\Delta x\). This assumption is made for the remainder of this section. Under this feasibility assumption, if \(X_1\) and \(X_2\) are written in terms of \(x\), i.e., \(X_1 = \text{diag}(x - \ell_j)\) and \(X_2 = \text{diag}(u_j - x_j)\),
respectively, then equations (2.5) are the Newton path-following equations (2.18) for a solution of the perturbed optimality conditions (2.3). The variables \( x_1 \) and \( x_2 \) may be computed implicitly for the line search, in which case the appropriate merit function is
\[
f(x) - (Ax - b)^T v^e + \frac{1}{2\mu_a} \| Ax - b \|^2 + \frac{1}{2\mu_a} \| Ax - b + \mu^a (v - v^e) \|^2
\]
\[
- \sum_{j=1}^{n} \left\{ \mu^a [z_1^{p}]_j \ln (x_j - \ell_j + \mu^a) + \mu^a [z_2^{p}]_j \ln ([z_1]_j (x_j - \ell_j + \mu^a)) - [z_1]_j (x_j - \ell_j + \mu^a) \right\}
\]
\[
- \sum_{j=1}^{n} \left\{ \mu^a [z_2^{p}]_j \ln (u_j - x_j + \mu^a) + \mu^a [z_2^{p}]_j \ln ([z_2]_j (u_j - x_j + \mu^a)) - [z_2]_j (u_j - x_j + \mu^a) \right\}.
\]

2.5. Computation of the shifted primal-dual penalty-barrier direction

Next we consider the solution of the path-following modified Newton equations (2.5), which will be written in the form
\[
\begin{pmatrix}
H & A^T & -I & I \\
A & -D_A & 0 & 0 \\
Z_1 & 0 & Z_1 D_1^T & 0 \\
-Z_2 & 0 & 0 & Z_2 D_2^T
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta v \\
\Delta z_1 \\
\Delta z_2
\end{pmatrix}
= - \begin{pmatrix}
g - A^T v - z_1 + z_2 \\
D_A (v - \pi^A) \\
Z_1 D_1^T (z_1 - \pi_1^2) \\
Z_2 D_2^T (z_2 - \pi_2^2)
\end{pmatrix}.
\]

which may be row-scaled to give
\[
\begin{pmatrix}
H & A^T & -I & I \\
A & -D_A & 0 & 0 \\
I & 0 & D_1^T & 0 \\
-I & 0 & 0 & D_2^T
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta v \\
\Delta z_1 \\
\Delta z_2
\end{pmatrix}
= - \begin{pmatrix}
g - A^T v - z_1 + z_2 \\
D_A (v - \pi^A) \\
D_1^T (z_1 - \pi_1^2) \\
D_2^T (z_2 - \pi_2^2)
\end{pmatrix}.
\]

The equations and variables can be rescaled and reordered to give
\[
\begin{pmatrix}
I & 0 & 0 & (D_1^T)^{-1} \\
0 & I & 0 & -(D_2^T)^{-1} \\
0 & 0 & -I & D_A^{-1} A \\
-I & I & A^T & H
\end{pmatrix}
\begin{pmatrix}
\Delta z_1 \\
\Delta z_2 \\
\Delta v \\
\Delta x
\end{pmatrix}
= - \begin{pmatrix}
z_1 - \pi_1^2 \\
z_2 - \pi_2^2 \\
v - \pi^A \\
g - A^T v - z_1 + z_2
\end{pmatrix}.
\]

Applying the nonsingular matrix
\[
\begin{pmatrix}
I & 0 & 0 & I \\
0 & I & 0 & I \\
0 & 0 & I & 0 \\
-I & I & A^T & I
\end{pmatrix}
\]
to both sides of (2.19) gives the block upper-triangular system

\[
\begin{pmatrix}
I & 0 & 0 & (D_1^z)^{-1} \\
I & 0 & - (D_2^z)^{-1} \\
I & D_1^{-1}A & H + A^TD_1^{-1}A + D_2^{-1} \\

\end{pmatrix}
\begin{pmatrix}
\Delta z_1 \\
\Delta z_2 \\
\Delta v \\
\Delta x
\end{pmatrix}
= -
\begin{pmatrix}
z_1 - \pi^x_1 \\
z_2 - \pi^x_2 \\
v - \pi^x \\
g - A^T\pi^x - \pi^x
\end{pmatrix},
\]

where \( \pi^x = \pi^x_1 - \pi^x_2 \), and \( D_2^{-1} = (D_1^z)^{-1} + (D_2^z)^{-1} \), for which \( D \) is \( ((D_1^z)^{-1} + (D_2^z)^{-1})^{-1} \). It follows that the solution of the path-following equations is given by

\[
\begin{align*}
\Delta v &= v^e - \frac{1}{\mu^x}(A(x + \Delta x) - b) - v, \\
\Delta z_1 &= -(X_1^b)^{-1}(z_1 \cdot (x + \Delta x - \ell + \mu^b e) - \mu^b z^x_1), \\
\Delta z_2 &= -(X_2^b)^{-1}(z_2 \cdot (u - x - \Delta x + \mu^b e) - \mu^b z^x_2),
\end{align*}
\]

where \( \Delta x \) satisfies \((H + A^TD_1^{-1}A + D_2^{-1})\Delta x = -(g - A^T\pi^x - \pi^x)\).
2. Linear Equality Constraints and Upper and Lower Bounds on the Variables

2.6. Summary: linear equalities with upper and lower bounds

Define the quantities

\[ D_A = \mu^A \mathbf{I}, \quad \pi^* = v^e - \frac{1}{\mu^A} (A\mathbf{x} - b), \]
\[ D_1^Z = X_1^e \mathbf{Z}_1^{-1}, \quad \pi_1^* = \mu^e (X_1^e)^{-1} z_1^e, \]
\[ D_2^Z = X_2^e \mathbf{Z}_2^{-1}, \quad \pi_2^* = \mu^e (X_2^e)^{-1} z_2^e, \]
\[ D_Z = ((D_1^Z)^{-1} + (D_2^Z)^{-1})^{-1}, \quad \pi^* = \pi_2^* - \pi_1^*, \]

then \( \Delta \mathbf{v}, \Delta z_1 \) and \( \Delta z_2 \) are given by

\[ \hat{x} = \mathbf{x} + \Delta \mathbf{x}, \quad \Delta z_1 = -(X_1^e)^{-1} (z_1 \cdot (\hat{x} - \ell + \mu^e e) - \mu^e z_1^e), \]
\[ \Delta z_2 = -(X_2^e)^{-1} (z_2 \cdot (u - \hat{x} + \mu^e e) - \mu^e z_2^e), \]

\[ \hat{\pi}^A = v^e - \frac{1}{\mu^A} (A\hat{x} - b), \quad \Delta \mathbf{v} = \hat{\pi}^A - v, \]

where \( \Delta \mathbf{x} \) is the solution of the equations

\[ (H + A^T D_A^{-1} A + D_Z^{-1}) \Delta \mathbf{x} = -(g - A^T \pi^A - \pi^*). \]

The line-search merit function is

\[ f(\mathbf{x}) - (A\mathbf{x} - b)^T v^e + \frac{1}{2\mu^A} \|A\mathbf{x} - b\|^2 + \frac{1}{2\mu^A} \|A\mathbf{x} - b + \mu^A (v - v^e)\|^2 \]
\[ - \sum_{j=1}^{n} \{ \mu^e [z_1^e]_j \ln ([z_1]_j (x_j - \ell_j + \mu^e e)) - [z_1]_j (x_j - \ell_j + \mu^e e) \} \]
\[ - \sum_{j=1}^{n} \{ \mu^e [z_2^e]_j \ln ([z_2]_j (u_j - x_j + \mu^e e)) - [z_2]_j (u_j - x_j + \mu^e e) \}. \quad (2.20) \]
3. Nonnegativity Constraints on the Slacks

We start by considering methods for an optimization problem with nonlinear equality constraints and non-negativity constraints on the slack variables only.

3.1. Problem statement and optimality conditions

The problem has the form

\[
\begin{align*}
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad s \geq 0, \\
\end{align*}
\]

where \( c : \mathbb{R}^n \to \mathbb{R}^m \) and \( f : \mathbb{R}^n \to \mathbb{R} \) are twice-continuously differentiable. The first-order KKT conditions for this problem are

\[
\begin{align*}
g(x^*) - J(x^*)^T y^* &= 0, \\
y^* - w^* &= 0, \quad w^* \geq 0, \\
c(x^*) - s^* &= 0, \quad s^* \geq 0, \\
w^* \cdot s^* &= 0.
\end{align*}
\]

3.2. The path-following equations

Let \( y^e \) denote an estimate of the Lagrange multipliers \( y^* \) associated with the equality constraints \( c(x) - s = 0 \). Similarly, let \( w^e \) denote a nonnegative estimate of the multipliers for the inequality constraints \( s \geq 0 \). Given small positive scalars \( \mu^p \) and \( \mu^b \), consider the perturbed optimality conditions

\[
\begin{align*}
g(x) - J(x)^T y &= 0, \\
y - w &= 0, \quad w \geq 0, \\
c(x) - s &= \mu^p(y^e - y), \quad s \geq 0, \\
w \cdot s &= \mu^b(w^e - w).
\end{align*}
\]

Consider the following primal-dual path following equations given by

\[
F(x, s, y, w; \mu^p, \mu^b, y^e, w^e) = 0,
\]

with

\[
F(x, s, y, w; \mu^p, \mu^b, y^e, w^e) = \begin{pmatrix}
g(x) - J(x)^T y \\
y - w \\
c(x) - s + \mu^p(y^e - y) \\
w \cdot s + \mu^b(w^e - w)
\end{pmatrix}.
\]

Any zero \((x, s, y, w)\) of \( F \) that satisfies \( s > 0 \) and \( w > 0 \) approximates a solution to problem (3.1), with the approximation becoming increasingly accurate as \( \mu^p(y^e - y^*) \to 0 \) and \( \mu^b(w - w^e) \to 0 \). For any sequence of \( y^e \) and \( w^e \) such that \( y^e \to y^* \) and
$w^e \to w^*$, and it must hold that solutions $(s, w)$ of (3.4) must satisfy $s \cdot w \to 0$. This implies that a solution $(x, s, y, w)$ of (3.2) will approximate a solution of (3.4) independently of the values of $\mu^e$ and $\mu^o$ (i.e., it is not necessary that $\mu \to 0$).

If $(x, s, y, w)$ is a given approximate zero of $F$ such that $s + \mu^o e > 0$ and $w > 0$, the Newton equations for the change in variables $(\Delta x, \Delta s, \Delta y, \Delta w)$ are given by

$$
\begin{pmatrix}
H & 0 & -J^T & 0 \\
-J & I & 0 & 0 \\
0 & 0 & I & -I \\
0 & -J & 0 & I
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta y \\
\Delta w
\end{pmatrix}
= -
\begin{pmatrix}
g - J^Ty \\
y - w \\
c - s + \mu^o(y - y^e) \\
w \cdot s + \mu^o(w - w^e)
\end{pmatrix},
$$

(3.5)

where $S^u = \text{diag}(s_i + \mu^e)$, $W = \text{diag}(w_i)$, $c = c(x)$, $g = g(x)$, $J = J(x)$, and $H = H(x, y)$.

3.3. A shifted primal-dual penalty-barrier function

The shifted primal-dual problem associated with problem (3.1) is obtained by including the constraints $c(x) - s = 0$ with the objective using a shifted primal-dual augmented Lagrangian term, and using a shifted primal-dual penalty-barrier term for the simple bounds. This gives the problem

$$
\begin{align*}
\text{minimize} \quad & M(x, s, y, w; \mu^e, \mu^o, y^e, w^e) \\
\text{subject to} \quad & s + \mu^o e > 0, \quad w > 0,
\end{align*}
$$

(3.6)

where $M(x, s, y, w; \mu^e, \mu^o, y^e, w^e)$ is the shifted primal-dual penalty-barrier function

$$
f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^e} ||c(x) - s||^2 + \frac{1}{2\mu^o} ||c(x) - s + \mu^o(y - y^e)||^2
$$

$$
- \sum_{i=1}^{m} \left\{ \mu^e w_i^e \ln(s_i + \mu^o) + \mu^o w_i^e \ln(w_i(s_i + \mu^o)) + \mu^o(w_i^e - w_i) - w_i s_i \right\},
$$

which is well defined for all $s$ such that $s + \mu^o e > 0$. This function has the same gradient as

$$
M(x, s, y, w; \mu^e, \mu^o, y^e, w^e) = f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^e} ||c(x) - s||^2 + \frac{1}{2\mu^o} ||c(x) - s + \mu^o(y - y^e)||^2
$$

$$
- \sum_{i=1}^{m} \left\{ \mu^e w_i^e \ln(s_i + \mu^o) + \mu^o w_i^e \ln(w_i(s_i + \mu^o)) - w_i s_i \right\}.
$$

(3.7)

Let $c$, $g$ and $J$ denote the quantities $c(x)$, $g(x)$ and $J(x)$. For clarity, the dependence of $M$ on the parameters $\mu^e$, $\mu^o$, $y^e$ and $w^e$, will be suppressed when appropriate, with $M(x, s, y, w)$ being used to denote $M(x, s, y, w; \mu^e, \mu^o, y^e, w^e)$. This function
may be written in the form:

\[
 f - (c - s)^T y^e + \frac{1}{2\mu^e}||c - s||^2 + \frac{1}{2\mu^e}||c - s + \mu^e(y - y^e)||^2 - \sum_{i=1}^m \{\mu^e w^e_i \ln (w_i(s_i + \mu^e)^2) - w_i(s_i + \mu^e)\}.
\]

(3.8)

Differentiating \(M(x, s, y, w)\) with respect to \(x, s, y\) and \(w\) gives:

\[
 \nabla M(x, s, y, w) = \begin{pmatrix}
 g - JT \begin{pmatrix} 2(y^e - \frac{1}{\mu^e}(c - s)) - y \\
 2(y^e - \frac{1}{\mu^e}(c - s) - y - (2\mu^e(S^\mu)^{-1}w^e - w) \\
 c - s + \mu^e(y - y^e) \\
 s + \mu^e - \mu^e W^{-1} w^e \\
 \end{pmatrix}
 \end{pmatrix},
\]

with \(S = \text{diag}(s_1, s_2, \ldots, s_m)\) and \(W = \text{diag}(w_1, w_2, \ldots, w_m)\). The gradient may be written in several equivalent forms:

\[
 \nabla M(x, s, y, w) = \begin{pmatrix}
 g - JT \begin{pmatrix} 2(y^e - \frac{1}{\mu^e}(c - s)) - y \\
 2(y^e - \frac{1}{\mu^e}(c - s) - y - (2\mu^e(S^\mu)^{-1}w^e - w) \\
 \frac{1}{2\mu^e}(c - s) + y - y^e \\
 W^{-1}(w \cdot s + \mu^e(w - w^e)) \\
 \end{pmatrix}
 \end{pmatrix} = \begin{pmatrix}
 \frac{1}{\mu^e} \begin{pmatrix} \pi^e + (\pi^e - y) \\
 \pi^e + (\pi^e - y) - (\pi^w + (\pi^w - w)) \\
 -D_v (\pi^e - y) \\
 -D_w (\pi^w - w) \\
 \end{pmatrix},
 \end{pmatrix},
\]

where

\[
 D_v = \mu^e I, \quad \pi^e = y^e - \frac{1}{\mu^e}(c - s),
\]

(3.9a)

\[
 D_w = S^\mu W^{-1}, \quad \pi^w = \mu^e(S^\mu)^{-1}w^e.
\]

(3.9b)

Similarly, the Hessian of \(M(x, s, y, w)\) is given by

\[
 \begin{pmatrix}
 H_1 & -\frac{1}{\mu^e} J^T \\
 -\frac{1}{\mu^e} J & 2(D_v^{-1} + \mu^e(S^\mu)^{-2}W^e) \\
 J & -I \\
 -I & \mu^e I \\
 0 & 0 \mu^e W^{-2} W^e \\
 \end{pmatrix},
\]
3. Nonnegativity Constraints on the Slacks

where $H_1 = H(x, 2\pi^y - y) + \frac{2}{\mu^r} J^T J$. Substituting $\mu^W \Pi^w = S^\mu \Pi^w$ from (3.9) gives the Hessian

$$
\begin{pmatrix}
H_1 & -\frac{2}{\mu^r} J^T D^{-1}_v + (S^\mu)^{-1} \Pi^w & J^T & 0 \\
-\frac{2}{\mu^r} J & 2(D^{-1}_v + (S^\mu)^{-1} \Pi^w) & -I & I \\
J & -I & \mu^r I & 0 \\
0 & I & 0 & W^{-2} \Pi^w S^\mu
\end{pmatrix},
$$

where $H_1 = H(x, 2\pi^y - y) + \frac{2}{\mu^r} J^T J$.

3.4. Derivation of the shifted primal-dual penalty-barrier direction

The primal-dual penalty-barrier problem (3.6) may be written in the form

$$
\min_{p \in \mathcal{I}} M(p)
$$

where

$$
\mathcal{I} = \{ p : p = (x, s, y, w), \text{ with } s + \mu^r e > 0, \ w > 0 \}.
$$

Let $p \in \mathcal{I}$ be given. The Newton direction $\Delta p$ is given by the solution of the subproblem

$$
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T \nabla^2 M(p) \Delta p.
$$

However, instead of solving (3.11), we define a modified Newton method by approximating the Hessian $\nabla^2 M(x, s, y, w)$ by a matrix $B(x, s, y, w)$. Consider the matrix defined by replacing $\pi^y$ by $y$ and $\pi^w$ by $w$ everywhere in the matrix $\nabla^2 M(x, s, y, w)$. This gives an approximate Hessian $B(x, s, y, w)$ of the form

$$
\begin{pmatrix}
\hat{H}_1 & -\frac{2}{\mu^r} J^T D^{-1}_v + (S^\mu)^{-1} \Pi^w & J^T & 0 \\
-\frac{2}{\mu^r} J & 2(D^{-1}_v + (S^\mu)^{-1} \Pi^w) & -I & I \\
J & -I & \mu^r I & 0 \\
0 & I & 0 & S^\mu W^{-1}
\end{pmatrix},
$$

where $\hat{H}_1 = H(x, y) + \frac{2}{\mu^r} J^T J$. The definitions of $D_v$ and $D_w$ may be used to write $B(x, s, y, w)$ in the form

$$
\begin{pmatrix}
H + 2J^T D^{-1}_v J & -2J^T D^{-1}_v D_v & J^T & 0 \\
-2D^{-1}_v J & 2(D^{-1}_v + D_w^{-1}) & -I & I \\
J & -I & D_v & 0 \\
0 & I & 0 & D_w
\end{pmatrix},
$$
3. Nonnegativity Constraints on the Slacks

where \( H = H(x,y) \) and \( D_w = D_w \). Given \( B(p) = B(x,s,y,w) \), a modified Newton direction is given by the solution of the subproblem

\[
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T B(p) \Delta p.
\]

Given \( p \), the modified Newton equations for this problem are given by

\[
\begin{pmatrix}
H + 2J^T D^{-1}_y J & -2J^T D^{-1}_y & J^T & 0 \\
-2D^{-1}_y J & 2(D^{-1}_y + D^{-1}_w) & -I & I \\
J & -I & D_y & 0 \\
0 & I & 0 & D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta y \\
\Delta w
\end{pmatrix}
= -\begin{pmatrix}
g - J^T (\pi^y + (\pi^y - y)) \\
(2\pi^y - y) - (2\pi^w - w) \\
-D_y (\pi^y - y) \\
-D_w (\pi^w - w)
\end{pmatrix}
\]

(3.12)

Consider the nonsingular block upper-triangular matrix

\[
T = \begin{pmatrix}
I & 0 & -2J^T D^{-1}_y & 0 \\
I & 2D^{-1}_y & -2D^{-1}_w & 0 \\
0 & 0 & S^\mu & W
\end{pmatrix}
\]

Applying \( T \) to both sides of (3.13) gives

\[
\begin{pmatrix}
H & 0 & -J^T & 0 \\
0 & 0 & I & -I \\
J & -I & D_y & 0 \\
0 & W & 0 & S^\mu
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta y \\
\Delta w
\end{pmatrix}
= -\begin{pmatrix}
g - J^T y \\
y - w \\
c - s + \mu^c (y - y^e) \\
s \cdot w + \mu^e (w - w^e)
\end{pmatrix}
\]

(3.13)

Comparing these equations with the path-following Newton equations (3.5) implies that a solution of the path-following equations is also a solution of (3.14).

3.5. Computation of the shifted primal-dual penalty-barrier direction

The path-following Newton equations (3.5) may be written in symmetric form

\[
\begin{pmatrix}
H & 0 & J^T & 0 \\
0 & 0 & -I & I \\
J & -I & -D_y & 0 \\
0 & I & 0 & -D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
-\Delta y \\
-\Delta w
\end{pmatrix}
= -\begin{pmatrix}
g - J^T y \\
y - w \\
c - s + \mu^c (y - y^e) \\
W^{-1} (s \cdot w + \mu^e (w - w^e))
\end{pmatrix}
\]
Consider the following reordered subset of equations and variables involving $\Delta w$, $\Delta s$, $\Delta x$ and $\Delta y$:

\[
\begin{pmatrix}
  D_w & I & 0 & 0 \\
  -I & 0 & 0 & I \\
  0 & -I & J & D_y
\end{pmatrix}
\begin{pmatrix}
  \Delta w \\
  \Delta s \\
  \Delta x \\
  \Delta y
\end{pmatrix}
= -\begin{pmatrix}
  W^{-1} \left( w \cdot s + \mu^w (w - w^e) \right) \\
  y - w \\
  c - s + \mu^c (y - y^e)
\end{pmatrix}
= -\begin{pmatrix}
  -D_w (\pi^w - w) \\
  y - w \\
  -D_y (\pi^y - y)
\end{pmatrix}.
\]

This gives the equations

\[
\begin{pmatrix}
  I & D_w^{-1} & 0 & 0 \\
  -I & 0 & 0 & I \\
  0 & -I & J & D_y
\end{pmatrix}
\begin{pmatrix}
  \Delta w \\
  \Delta s \\
  \Delta x \\
  \Delta y
\end{pmatrix}
= -\begin{pmatrix}
  w - \pi^w \\
  y - w \\
  -D_y (\pi^y - y)
\end{pmatrix}.
\]

Applying the nonsingular matrix

\[
\begin{pmatrix}
  I \\
  I \\
  D_w \\
  D_w
\end{pmatrix}
\]

on the left-hand side of (3.15) gives the block upper-trapezoidal system

\[
\begin{pmatrix}
  I & D_w^{-1} & 0 & 0 \\
  -I & 0 & 0 & I \\
  J & D_y + D_w
\end{pmatrix}
\begin{pmatrix}
  \Delta w \\
  \Delta s \\
  \Delta x \\
  \Delta y
\end{pmatrix}
= -\begin{pmatrix}
  w - \pi^w \\
  y - \pi^w \\
  D_y (y - \pi^y) + D_w (y - \pi^w)
\end{pmatrix}.
\]

The solution of this system of equations is given by

\[
\Delta s = -D_w (y + \Delta y - \pi^w)
\]
\[
\Delta w = -(S^\mu)^{-1} \left( w \cdot (s + \Delta s) + \mu^a (w - w^e) \right),
\]

where $\Delta x$ and $\Delta y$ satisfy the KKT system

\[
\begin{pmatrix}
  H & -J^T \\
  J & D_y + D_w
\end{pmatrix}
\begin{pmatrix}
  \Delta x \\
  \Delta y
\end{pmatrix}
= -\begin{pmatrix}
  g - J^T y \\
  D_y (y - \pi^y) + D_w (y - \pi^w)
\end{pmatrix}.
\]
3.6. Summary

The results of Sections 3.1–3.5 imply that the solution of the path-following equations (3.5) may be computed as

\[ \hat{y} = y + \Delta y, \quad \Delta s = -\hat{D}_w (\hat{y} - \pi^w), \]
\[ \hat{s} = s + \Delta s, \quad \Delta w = -(S^\mu)^{-1} (w \cdot \hat{s} + \mu^\rho (w - w^e)), \]

where \( \Delta x \) and \( \Delta y \) satisfy the equations

\[
\begin{pmatrix}
H(x, y) & -J(x)^T \\
J(x) & D_v + D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix}
= -\begin{pmatrix}
g(x) - J(x)^T y \\
D_v (y - \pi^v) + D_w (y - \pi^w)
\end{pmatrix},
\]

and \( D_v, D_w, \pi^v \) and \( \pi^w \) are given by

\[ D_v = \mu^v I, \quad \pi^v = y^e - \frac{1}{\mu^v} (c(x) - s), \]
\[ D_w = S^\mu W^{-1}, \quad \pi^w = \mu^\rho (S^\mu)^{-1} w^e. \]

The associated line-search merit function \( M(x, s, y, w; \mu^v, \mu^\rho, y^e, w^e) \) is given by

\[ f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^v} ||c(x) - s||^2 + \frac{1}{2\mu^\rho} ||c(x) - s + \mu^\rho (y - y^e)||^2 - \sum_{i=1}^m \{ \mu^\rho w_i^e \ln (w_i (s_i + \mu^\rho)^2) - w_i (s_i + \mu^\rho) \}. \]
4. Fixed and Nonnegative Slacks

Next we consider nonlinear equality constraints with slacks that are either fixed or nonnegative. The variables are not subject to bounds.

4.1. Problem statement and optimality conditions

The problem has the form

\[
\begin{align*}
\text{minimize} & \quad f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad L_X s, \quad L_F s \geq 0,
\end{align*}
\]

(4.1)

where \( c : \mathbb{R}^n \to \mathbb{R}^m \) and \( f : \mathbb{R}^n \to \mathbb{R} \) are twice-continuously differentiable and \( L_X \) and \( L_F \) are fixed matrices of dimension \( m_F \times m \) and \( m_X \times m \), respectively, with \( m = m_F + m_X \). The matrices \( L_X \) and \( L_F \) are formed from rows of the identity matrix \( I_m \) in such a way that \( L_X s \) and \( L_F s \) give the fixed and “free” components of \( s \). It follows that there is an \( m \times m \) permutation matrix \( P \) such that

\[
P = \begin{pmatrix} L_F \\ L_X \end{pmatrix}
\]

with the matrices \( L_F \) and \( L_X \) satisfying the identities \( L_F L_F^T = I_F \), \( L_X L_X^T = I_X \), and \( L_F L_X^T = 0 \). The first-order KKT conditions for this problem are

\[
\begin{align*}
g(x^*) - J(x^*)^T y^* &= 0, \quad (4.2a) \\
c(x^*) - s^* &= 0, \quad L_X s^* = 0, \quad (4.2b) \\
y^* - L_X^T w_X^* - L_F^T w_F^* &= 0, \quad (4.2c) \\
L_P s^* \geq 0, \quad w_F^* \geq 0, \quad (4.2d) \\
w_F^* \cdot L_P s^* &= 0, \quad (4.2e)
\end{align*}
\]

where \( y^* \) and \( w_X^* \) are the Lagrange multipliers for the equality constraints \( c(x) - s = 0 \) and \( L_X s = 0 \), and \( w_F^* \) may be interpreted as the Lagrange multipliers for the nonnegativity constraints \( L_F s \geq 0 \).

4.2. The path-following equations

Let \( y^e \) be an estimate of the Lagrange multipliers for the nonlinear equality constraints \( c(x) - s = 0 \). Similarly, let \( w^e \) denote a nonnegative estimate of the multipliers for the inequality constraints \( L_F s \geq 0 \). Given small positive scalars \( \mu^p \) and \( \mu^a \), consider
Consider the following primal-dual path following equations given by

\[ F(x, s, y, w_x, w_p; \mu^e, \mu^a, y^e, w^e) = 0, \]

with the approximation becoming increasingly accurate as the terms \( \mu^e(y^e - y) \) and \( \mu^a(w_p - w^e) \) approach zero. For any sequence of \( y^e \) and \( w^e \) such that \( y^e \rightarrow y^* \) and \( w^e \rightarrow w^*_p \), it must hold that solutions \( (x, s, y, w_x, w_p) \) of (4.3) must satisfy \( y \cdot (c(x) - s) \rightarrow 0 \), \( w_p \cdot (L_p s) \rightarrow 0 \), and \( w_p \cdot L_p s \rightarrow 0 \). This implies that any solution \( (x, s, y, w_x, w_p) \) of (4.3) will approximate a solution of (4.2) independently of the values of \( \mu^e \) and \( \mu^a \) (i.e., it is not necessary that \( \mu^e, \mu^a \rightarrow 0 \)).

Given an approximate zero \( (x, s, y, w_x, w_p) \) of \( F \) such that \( L_p s > 0 \) and \( w_p > 0 \), the Newton equations for the change in variables \( (\Delta x, \Delta s, \Delta y, \Delta w_x, \Delta w_p) \) are given by

\[
\begin{pmatrix}
H & 0 & -J^T & 0 & 0 \\
0 & 0 & I_m & -L_{x}^T & -L_{p}^T \\
J & -I_m & D_y & 0 & 0 \\
0 & L_{x} & 0 & 0 & 0 \\
0 & W L_{p} & 0 & 0 & S^\mu
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y \\
\Delta s \\
\Delta w_x \\
\Delta w_p
\end{pmatrix}
= - \begin{pmatrix}
g - J^T y \\
y - L_{x}^T w_x - L_{p}^T w_p \\
(c(x) - s + \mu^e(y - y^e)) \\
L_{x} s \\
w_p \cdot L_p s + \mu^a(w_p - w^e)
\end{pmatrix},
\]

where \( D_y = \mu^e I, W = \text{diag}(w_p) \) and \( S^\mu = \text{diag}(s_i + \mu^a) \).

Any \( s \) may be written as \( s = L_{p}^T s_p + L_{x}^T s_x \), where \( s_p \) and \( s_x \) denote the components of \( s \) corresponding to the “free” and “fixed” components of \( s \), respectively. Throughout, we assume that \( s_x \) satisfies \( L_{x} s = 0 \), in which case the expansion of \( \Delta s \) satisfies

\[
\Delta s = L_{p}^T \Delta s_p + L_{x}^T \Delta s_x = L_{p}^T \Delta s_p.
\]
This identity allows us to write the equations (4.5) in the form
\[
\begin{pmatrix}
H & 0 & -J^T & 0 \\
0 & L_F & 0 & -I_F \\
J & -L_F^T & D_X & 0 \\
0 & W & 0 & S^\mu
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta y \\
\Delta w
\end{pmatrix}
= \begin{pmatrix}
g - J^Ty \\
y_F - w_F \\
c - s + \mu^\epsilon(y - y^\epsilon) \\
w_F \cdot L_F s + \mu^\beta(w_F - w^\epsilon)
\end{pmatrix}
\] (4.6)

The vectors $\Delta s$ and $\Delta w_X$ are recovered as $\Delta s = L_F^T \Delta s_F$ and $\Delta w_X = [y + \Delta y - w]_X$.

### 4.3. A shifted primal-dual penalty-barrier function

Problem (4.1) is equivalent to
\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad c(x) - s = 0, \quad L_X s = 0, \quad L_F s - s_F = 0, \quad s_F \geq 0.
\end{align*}
\]

Consider the shifted primal-dual penalty-barrier problem
\[
\begin{align*}
\text{minimize} & \quad M(x, s, s_F, y, w_F; \mu^\epsilon, \mu^\beta, y^\epsilon, w^\epsilon) \\
\text{subject to} & \quad L_X s = 0, \quad L_F s - s_F = 0, \quad s_F + \mu^\beta e > 0, \quad w_F > 0,
\end{align*}
\] (4.7)

where $M(x, s, s_F, y, w_F; \mu^\epsilon, \mu^\beta, y^\epsilon, w^\epsilon)$ is the shifted primal-dual penalty-barrier function
\[
\begin{align*}
f(x) - (c(x) - s)^T y^\epsilon + \frac{1}{2\mu^\epsilon} \|c(x) - s\|^2 & + \frac{1}{2\mu^\beta} \|c(x) - s + \mu^\epsilon(y - y^\epsilon)\|^2 \\
& - \sum_{i=1}^{n_F} \{\mu^\epsilon w^\epsilon_i \ln \left(\{s_F + \mu^\beta e\}_i\right) + \mu^\beta w^\epsilon_i \ln \left(\{w_F \cdot (s_F + \mu^\beta e)\}_i\right) - \left[w_F \cdot (s_F + \mu^\beta e)\right]_i\}. \tag{4.8}
\end{align*}
\]

Let $c$, $g$ and $J$ denote the quantities $c(x)$, $g(x)$ and $J(x)$. Differentiating $M(x, s, s_F, y, w_F, w_F)$ with respect to $x$, $s$, $s_F$, $y$ and $w_F$ gives
\[
\nabla M(x, s, s_F, y, w_F) = \begin{pmatrix}
g - J^T(2(y^\epsilon - \frac{1}{\mu^\epsilon}(c - s)) - y) \\
2(y^\epsilon - \frac{1}{\mu^\epsilon}(c - s)) - y \\
w_F - 2\mu^\beta(S^\mu)^{-1}w^\epsilon \\
c - s + \mu^\epsilon(y - y^\epsilon) \\
(s_F + \mu^\beta e - \mu^\beta W^{-1}w^\epsilon)
\end{pmatrix}.
\]
The gradient may be written in several equivalent forms

\[
\nabla M(x, s, s_F, y, w_F) = \begin{pmatrix}
    g - J^T(2(y^s - \frac{1}{\mu^p}(c - s)) - y) \\
    2(y^s - \frac{1}{\mu^p}(c - s)) - y \\
    w_F - 2\mu^s(S^\mu)^{-1}w^e \\
    c - s + \mu^p(y - y^s) \\
    s_F + \mu^e c - \mu W^{-1} w^e
\end{pmatrix} = \begin{pmatrix}
    g - J^T(2(y^s - \frac{1}{\mu^p}(c - s)) - y) \\
    2(y^s - \frac{1}{\mu^p}(c - s)) - y \\
    (S^\mu)^{-1}(w_F \cdot s_F + \mu^e w^e + \mu^e w_F - w^e) \\
    c - s + \mu^p(y - y^s) \\
    W^{-1}(w_F \cdot s_F + \mu^e w_F - w^e)
\end{pmatrix}
\]

\[
= \begin{pmatrix}
    g - J^T(\pi^\nu + (\pi^\nu - y)) \\
    \pi^\nu + (\pi^\nu - y) \\
    -D_F(\pi^\nu - y) \\
    -D_W(\pi^w - w_F)
\end{pmatrix},
\]

where

\[
D_F = \mu^F I_m, \quad \pi^\nu = y^s - \frac{1}{\mu^p}(c - s), \quad (4.9a)
\]

\[
D_W = S^\mu W^{-1}, \quad \pi^w = \mu^e(S^\mu)^{-1}w^e. \quad (4.9b)
\]

Similarly, the Hessian of \(M(x, s, s_1, s_2, y, w_1, w_2)\) is given by

\[
\begin{pmatrix}
    H_1 - \frac{2}{\mu^e} J^T & 0 & J^T & 0 \\
    -\frac{2}{\mu^e} J & \frac{2}{\mu^e} I_m & 0 & -I_m & 0 \\
    J & -I_m & 2\mu^e(S^\mu)^{-2}W^e & 0 & I_F \\
    0 & 0 & I_F & \mu^e I_m & 0 \\
    0 & 0 & I_F & 0 & \mu W^{-2} W^e
\end{pmatrix},
\]

where \(H_1 = H(x, 2\pi^\nu - y) + \frac{2}{\mu^e} J^T J\). Substituting \(\mu^e W^e = S^\mu \Pi^w\) from (4.9) gives the Hessian

\[
\begin{pmatrix}
    H_1 - \frac{2}{\mu^e} J^T & 0 & J^T & 0 \\
    -\frac{2}{\mu^e} J & \frac{2}{\mu^e} I_m & 0 & -I_m & 0 \\
    J & -I_m & 2(S^\mu)^{-1}\Pi^w & 0 & I_F \\
    0 & 0 & I_F & \mu^e I_m & 0 \\
    0 & 0 & I_F & 0 & S^\mu \Pi^{-2} \Pi^w
\end{pmatrix}
\]
4.4. Derivation of the shifted primal-dual penalty-barrier direction

The primal-dual penalty-barrier problem may be written in the form

$$
\begin{align*}
\minimize_{p \in \mathcal{I}} & \quad M(p) \quad \text{subject to} \quad Cp = 0, \\
\end{align*}
$$

(4.10)

where

$$\mathcal{I} = \{ p : p = (x, s, s_F, y, w_F), \quad \text{with} \quad s_F + \mu^e e > 0, \ w_F > 0 \},$$

with

$$C = \begin{pmatrix}
0 & L_X & 0 & 0 & 0 \\
0 & L_F & -I_F & 0 & 0
\end{pmatrix}.$$

Let $p \in \mathcal{I}$ be given. The Newton direction $\Delta p$ is given by the solution of the subproblem

$$
\begin{align*}
\minimize_{\Delta p} & \quad \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T \nabla^2 M(p) \Delta p \\
\text{subject to} & \quad C \Delta p = -Cp.
\end{align*}
$$

(4.11)

However, instead of solving (4.11), we define a linearly constrained modified Newton method by approximating the Hessian $\nabla^2 M(x, s, s_F, y, w_F)$ by a matrix $B(x, s, s_F, y, w_F)$. Consider the matrix defined by replacing $\pi^y$ by $y$ and $\pi^w$ by $w_F$, everywhere in the matrix $\nabla^2 M(x, s, s_F, y, w_F)$. This gives an approximate Hessian $B(x, s, s_F, y, w_F)$ of the form

$$\begin{pmatrix}
\hat{H}_1 & -2 \frac{\mu^e}{\mu} J^T \\
-2 \frac{\mu}{\mu^e} J \mu^e & 0 & J^T \\
0 & 0 & -I_m & 0 \\
J & -I_m & 0 & \mu^e I_m & 0 \\
0 & 0 & -I_m & 0 & S \mu^e W^{-1}
\end{pmatrix}
$$

where $\hat{H}_1 = H(x, y) + 2 J^T D^{-1}_y J$. The definitions of $D_y$ and $D_w$ may be used to write $B(x, s, s_F, y, w_F)$ as

$$\begin{pmatrix}
H + 2 J^T D^{-1}_y J & -2 J^T D^{-1}_y \mu^e & J^T & 0 \\
-2 D^{-1}_y \mu^e & 0 & -I_m & 0 \\
0 & 0 & 2 D^{-1}_w & 0 & I_F \\
J & -I_m & 0 & D_v & 0 \\
0 & 0 & -I_m & 0 & D_w
\end{pmatrix}
$$

where $H = H(x, y)$. Given $B(p) = B(x, s, s_F, y, w_F)$, a modified Newton direction is given by the solution of the QP subproblem

$$
\begin{align*}
\minimize_{\Delta p} & \quad \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T B(p) \Delta p \\
\text{subject to} & \quad C \Delta p = -Cp.
\end{align*}
$$
If \( p = (x, s, s_F, y, w) \) is feasible for the constraints then \( Cp = 0 \) and \( L_Xs = 0 \) and \( L_Fs - s_F = 0 \). In this case every feasible \( \Delta p \) may be written in the form \( \Delta p = Nd \), where \( N \) denote a matrix whose columns form a basis for \( \text{null}(C) \), i.e., \( CN = 0 \) and \( (C^T \quad N) \) is nonsingular. This implies that \( d \) must satisfy the reduced equations

\[
N^T B(p)Nd = -N^T \nabla M(p).
\]

Consider the particular null-space basis given by

\[
N = \begin{pmatrix}
  I_n & 0 & 0 & 0 \\
  0 & L_F^T & 0 & 0 \\
  0 & L_F & 0 & 0 \\
  0 & 0 & I_m & 0 \\
  0 & 0 & 0 & I_F
\end{pmatrix}.
\]

This definition of \( N \) gives the reduced Hessian

\[
N^T B(p)N = \begin{pmatrix}
  H + 2J^T D^{-1}_y J & -2J^T D^{-1}_y L_F^T & J^T & 0 \\
  -2L_F D^{-1}_y J & 2(L_F D^{-1}_y L_F^T + D^{-1}_w) & -L_F & I_F \\
  0 & -L_F^T & D_F & 0 \\
  0 & 0 & I_F & 0 \\
  0 & 0 & 0 & D_w
\end{pmatrix}.
\]

Similarly, the reduced gradient \( N^T \nabla M(p) \) is given by

\[
N^T \nabla M(p) = \begin{pmatrix}
  I_n & 0 & 0 & 0 & 0 \\
  0 & L_F & I_F & 0 & 0 \\
  0 & 0 & 0 & I_m & 0 \\
  0 & 0 & 0 & 0 & I_F
\end{pmatrix} \begin{pmatrix}
  g - J^T(\pi^y + (\pi^y - y)) \\
  \pi^y + (\pi^y - y) \\
  -\pi^w + (\pi^w - w_F) \\
  -D_F(\pi^y - y) \\
  -D_w(\pi^w - w_F)
\end{pmatrix} = \begin{pmatrix}
  g - J^T(\pi^y + (\pi^y - y)) \\
  \pi^y + (\pi^y - y) \\
  -\pi^w + (\pi^w - w_F) \\
  -D_F(\pi^y - y) \\
  -D_w(\pi^w - w_F)
\end{pmatrix}.
\]

The reduced modified equations \( N^T B(p)Nd = -N^T \nabla M(p) \) are then

\[
\begin{pmatrix}
  H + 2J^T D^{-1}_y J & -2J^T D^{-1}_y L_F^T & J^T & 0 \\
  -2L_F D^{-1}_y J & 2(L_F D^{-1}_y L_F^T + D^{-1}_w) & -L_F & I_F \\
  0 & -L_F^T & D_F & 0 \\
  0 & 0 & I_F & 0 \\
  0 & 0 & 0 & D_w
\end{pmatrix} \begin{pmatrix}
  d_1 \\
  d_2 \\
  d_3 \\
  d_4
\end{pmatrix} = \begin{pmatrix}
  g - J^T(\pi^y + (\pi^y - y)) \\
  \pi^y + (\pi^y - y) \\
  -\pi^w + (\pi^w - w_F) \\
  -D_F(\pi^y - y) \\
  -D_w(\pi^w - w_F)
\end{pmatrix}.
\]
Given any nonsingular matrix $R$, the direction $d$ satisfies

$$RN^T B(p) d = -RN^T \nabla M(p).$$

In particular, consider

$$R = \begin{pmatrix}
I_n & 0 & -2J^T D_Y^{-1} & 0 \\
I_F & 2L_Y D_Y^{-1} & -2D_w^{-1} & 0 \\
I_m & 0 & -2L_Y D_Y^{-1} & 2W^{-1} D_w^{-1} \\
0 & 0 & W^{-1} & 0
\end{pmatrix},$$

which is nonsingular if $W$ is positive definite, with

$$R^{-1} = \begin{pmatrix}
I_n & 0 & 2J^T D_Y^{-1} & 0 \\
I_F & -2L_Y D_Y^{-1} & 2W^{-1} D_w^{-1} & 0 \\
I_m & 0 & -2L_Y D_Y^{-1} & -2D_w^{-1} I_F \\
0 & 0 & W & 2(S^\mu)^{-1}
\end{pmatrix}.$$

For this $R$, the product $RN^T B(p)N$ is given by

$$\begin{pmatrix}
H + 2J^T D_Y^{-1} J & -2J^T D_Y^{-1} L_F^T & J^T & 0 \\
2L_Y D_Y^{-1} -2J^T D_Y^{-1} L_F^T + D_w^{-1} & -L_F & I_F & 0 \\
J & -J^T & 0 & 0 \\
0 & 0 & L_F & -I_F
\end{pmatrix} \begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4
\end{pmatrix} = \begin{pmatrix}
H & 0 & -J^T & 0 \\
0 & 0 & L_F & -I_F \\
J & -L_F^T & D_Y & 0 \\
0 & W & 0 & W D_w
\end{pmatrix} \begin{pmatrix}
g - J^T y \\
y_F - w_F \\
-D_Y (\pi^v - y) \\
-W D_w (\pi^w - w_F)
\end{pmatrix}.$$

Similarly, the transformed right-hand side $RN^T \nabla M(p)$ is given by

$$\begin{pmatrix}
g - J^T y \\
y_F - w_F \\
-D_Y (\pi^v - y) \\
-W D_w (\pi^w - w_F)
\end{pmatrix}.$$

Putting all this together gives the following transformed unsymmetric reduced modified Newton equations for the vector $d$:

$$\begin{pmatrix}
H & 0 & -J^T & 0 \\
0 & 0 & L_F & -I_F \\
J & -L_F^T & D_Y & 0 \\
0 & W & 0 & S^\mu
\end{pmatrix} \begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4
\end{pmatrix} = \begin{pmatrix}
g - J^T y \\
y_F - w_F \\
-D_Y (\pi^v - y) \\
-W D_w (\pi^w - w_F)
\end{pmatrix}.$$
4. Fixed and Nonnegative Slacks

or, equivalently,

\[
\begin{pmatrix}
H & 0 & -J^T & 0 \\
0 & 0 & L_F & -I_F \\
J & -L_F^T & D_Y & 0 \\
0 & W & 0 & S^\mu
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4
\end{pmatrix}
= -\begin{pmatrix}
g - J^Ty \\
y_F - w_F \\
c - s + \mu^p(y - y^\mu) \\
w_F \cdot L_F s + \mu^p(w_F - w^\mu)
\end{pmatrix}.
\]

(4.13)

Then, the definition of \(N\) from (4.12) implies that

\[
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta s_F \\
\Delta y \\
\Delta w_F
\end{pmatrix}
= \Delta p = Nd = \begin{pmatrix}
d_1 \\
L_F^T d_2 \\
d_2 \\
d_3 \\
d_4
\end{pmatrix}.
\]

These identities allow us to write equations (4.13) as

\[
\begin{pmatrix}
H & 0 & -J^T & 0 \\
0 & 0 & L_F & -I_F \\
J & -L_F^T & D_Y & 0 \\
0 & W & 0 & S^\mu
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta s_F \\
\Delta y \\
\Delta w_F
\end{pmatrix}
= -\begin{pmatrix}
g - J^Ty \\
y_F - w_F \\
c - s + \mu^p(y - y^\mu) \\
w_F \cdot L_F s + \mu^p(w_F - w^\mu)
\end{pmatrix}.
\]

(4.14)

If \(S\) is written in terms of \(s\), i.e., \(S = \text{diag}(L_F s)\), then the equations (4.14) are the Newton equations for the solution of the perturbed optimality conditions (4.3). The variables \(s_F\) may be computed implicitly for the line search, in which case the appropriate merit function is

\[
f - (c - s)^Ty^\varepsilon + \frac{1}{2\mu^p}\|c - s\|^2 + \frac{1}{2\mu^p}\|c - s + \mu^p(y - y^\mu)\|^2
- \sum_{i=1}^{n_F} \{\mu^p w^\mu_i \ln ([s_F + \mu^p e]_i) + \mu^p w^\mu_i \ln ([w \cdot (s_F + \mu^p e)]_i) - [w \cdot (s_F + \mu^p e)]_i\}.
\]

4.5. Computation of the shifted primal-dual penalty-barrier direction

Next we consider the solution of the modified Newton equations (4.14), which are written in the form

\[
\begin{pmatrix}
H & 0 & -J^T & 0 \\
0 & 0 & L_F & -I_F \\
J & -L_F^T & D_Y & 0 \\
0 & D_w^{-1} & 0 & I_F
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta s_F \\
\Delta y \\
\Delta w_F
\end{pmatrix}
= -\begin{pmatrix}
g - J^Ty \\
y_F - w_F \\
-c + \mu^p(y - y^\mu) \\
-w_F \cdot (\pi^\varepsilon - y)
\end{pmatrix}.
\]

(4.15)
using the identities $D_w = S^\mu W^{-1}$ and $w_F \cdot s_F + \mu^s (w_F - w^e) = -S^\mu (\pi^w - w_F)$. Consider the following reordered set of equations and variables involving (in order) $\Delta w_F$, $\Delta s_F$, $\Delta x$ and $\Delta y$:

$$
\begin{pmatrix}
I_F & D_{w_F}^{-1} & 0 & 0 \\
-I_F & 0 & 0 & L_p \\
0 & -L_p^T & J & D_Y \\
0 & 0 & H & -J^T
\end{pmatrix}
\begin{pmatrix}
\Delta w_F \\
\Delta s_F \\
\Delta x \\
\Delta y
\end{pmatrix}
= -
\begin{pmatrix}
w_F - \pi^w \\
y_F - w_F \\
D_Y (y - \pi^Y) \\
g - J^T y
\end{pmatrix}.
$$

(4.16)

Applying the nonsingular matrix

$$
\begin{pmatrix}
I_F & I_F & I_F \\
L_p^T D_w & L_p^T D_w & I_m \\
& & I_n
\end{pmatrix}
$$
on the left- and right-hand side of (4.16) yields the block upper-triangular system of equations

$$
\begin{pmatrix}
I_F & D_{w_F}^{-1} & 0 & 0 \\
D_{w_F}^{-1} & 0 & 0 & L_p \\
J & D_Y + L_p^T D_w L_p & H & -J^T
\end{pmatrix}
\begin{pmatrix}
\Delta w_F \\
\Delta s_F \\
\Delta x \\
\Delta y
\end{pmatrix}
= -
\begin{pmatrix}
w_F - \pi^w \\
y_F - \pi^w \\
D_Y (y - \pi^Y) + L_p^T D_w (y_F - \pi^w) \\
g - J^T y
\end{pmatrix}.
$$

(4.17)

Solving (4.17) while using the last block equation of (4.14) as an alternative definition of $\Delta w_F$ gives the solution of the path-following equations as

$$
\begin{align*}
\Delta s_F &= -D_w (y_F + \Delta y_F - \pi^w), \\
\Delta s &= L_p^T \Delta s_F, \\
\Delta w_F &= -(S^\mu)^{-1} (w_F \cdot (L_F (s + \Delta s) - s_F + \mu^s e) - \mu^s w^e),
\end{align*}
$$

where $\Delta x$ and $\Delta y$ satisfy the equations

$$
\begin{pmatrix}
H & -J^T \\
J & D_Y + L_p^T D_w L_p
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix}
= -
\begin{pmatrix}
g - J^T y \\
D_Y (y - \pi^Y) + L_p^T D_w (y_F - \pi^w)
\end{pmatrix}.
$$
4.6. Summary: bounded slacks

Consider the quantities

\[ D_v = \mu^v I, \quad \pi^v = y^e - \frac{1}{\mu^v} (c(x) - s), \]
\[ D_w = S^w W^{-1}, \quad \pi^w = \mu^w (S^w)^{-1} w^e, \]

then \( \Delta s, \Delta s_F \) and \( \Delta w_F \) are given by

\[ \hat{y} = y + \Delta y, \quad \Delta s_F = -D_w (\hat{y}_F - \pi^w), \]
\[ \Delta s = L_F^T \Delta s_F, \quad \Delta w_x = [\hat{y} - w]_x, \]
\[ \hat{s} = s + \Delta s, \quad \Delta w_F = -(S^w)^{-1} (w_F \cdot (L_F \hat{s} + \mu^w e) - \mu^w w^e), \]

where \( \Delta x \) and \( \Delta y \) satisfy the equations

\[
\begin{pmatrix}
H & -J^T \\
J & D_v + L_F^T D_w L_F
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix}
= - \begin{pmatrix}
g - J^T y \\
D_v (y - \pi^v) + L_F^T D_w (y_F - \pi^w)
\end{pmatrix}.
\]

The associated line-search merit function is given by

\[
f - (c - s)^T y^e + \frac{1}{2\mu^v} ||c - s||^2 + \frac{1}{2\mu^w} ||c - s + \mu^v (y - y^e)||^2
- \sum_{i=1}^{n_F} \left\{ \mu^w w_i^e \log \left( |s_F + \mu^w e| \right)_i + \mu^w w_i^e \log \left( |w \cdot (s_F + \mu^w e)| \right)_i - |w \cdot (s_F + \mu^w e)|_i \right\}.
\]
5. Nonnegativity Constraints on the Variables and Slacks

Next we consider methods for an optimization problem with nonlinear equality constraints and non-negativity constraints on the variables and slacks.

5.1. Problem statement and optimality conditions

The problem has the form

\[
\begin{align*}
\text{minimize} \quad & f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad Ax = b, \quad x \geq 0, \quad s \geq 0, \\
\end{align*}
\]

where \( c : \mathbb{R}^n \rightarrow \mathbb{R}^m \) and \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) are twice-continuously differentiable. The first-order KKT conditions for this problem are

\[
\begin{align*}
& g(x^*) - A^T v^* - J(x^*)^T y^* - z^* = 0, \quad z^* \geq 0, \quad (5.2a) \\
& y^* - w^* = 0, \quad w^* \geq 0, \quad (5.2b) \\
& c(x^*) - s^* = 0, \quad Ax^* = b, \quad x^* \geq 0, \quad s^* \geq 0, \quad (5.2c) \\
& z^* \cdot x^* = 0, \quad w^* \cdot s^* = 0. \quad (5.2d)
\end{align*}
\]

The vectors \( y^*, z^* \) and \( w^* \) constitute the Lagrange multipliers for the equality constraints \( c(x) - s = 0 \) and the nonnegativity constraints \( x \geq 0 \) and \( s \geq 0 \), respectively. A vector \( (x, y, z, w) \) is said to constitute a primal-dual estimate of the quantities \( (x^*, y^*, z^*, w^*) \) satisfying the optimality conditions for (5.1).

5.2. The path-following equations

Let \( v^e \) and \( y^e \) denote estimates of the Lagrange multipliers for the equality constraints \( Ax = b \) and \( c(x) - s = 0 \). Similarly, let \( z^e \) and \( w^e \) denote nonnegative estimates of the multipliers for the inequality constraints \( x \geq 0 \) and \( s \geq 0 \). Given small positive scalars \( \mu^A, \mu^P \) and \( \mu^B \), consider the perturbed optimality conditions

\[
\begin{align*}
& g(x) - A^T v - J(x)^T y - z = 0, \quad z \geq 0, \quad (5.3a) \\
& y - w = 0, \quad w \geq 0, \quad (5.3b) \\
& c(x) - s = \mu^P (y^e - y), \quad Ax - b = \mu^A (v^e - v), \quad x \geq 0, \quad s \geq 0, \quad (5.3c) \\
& z \cdot x = \mu^B (z^e - z), \quad w \cdot s = \mu^B (w^e - w). \quad (5.3d)
\end{align*}
\]
Consider the primal-dual path-following equations \( F(x, s, v, y, z, w; \mu_A, \mu^e, \mu^p, y^e, z^e, w^e) = 0 \), with

\[
F(x, s, v, y, z, w; \mu_A, \mu^e, \mu^p, y^e, z^e, w^e) = \begin{pmatrix}
g(x) - J(x)^T y - A^T v - z \\
y - w \\
Ax - b - \mu^e(v - v^e) \\
c(x) - s + \mu^p(y - y^e) \\
z \cdot x + \mu^p(z - z^e) \\
w \cdot s + \mu^p(w - w^e)
\end{pmatrix}.
\] (5.4)

Any zero \((x, s, v, y, z, w)\) of \( F \) that satisfies \( x > 0, s > 0, z > 0, \) and \( w > 0 \) approximates a point satisfying the optimality conditions (5.2), with the approximation becoming increasingly accurate as \( \mu^p(y - y^e) \to 0, \mu^e(v - v^e) \to 0, \mu^p(z - z^e) \to 0, \) and \( \mu^p(w - w^e) \to 0. \) For any sequence of \( v^e, y^e, z^e \) and \( w^e \) such that \( v^e \to v^*, y^e \to y^*, z^e \to z^* \) and \( w^e \to w^* \), it must hold that solutions \((s, w)\) of (5.4) must satisfy \( z \cdot x \to 0 \) and \( w \cdot s \to 0. \) This implies that a solution \((x, s, v, y, z, w)\) of (5.4) will approximate a solution of (5.2) independently of the values of \( \mu_A, \mu^e \) and \( \mu^p \) (i.e., it is not necessary that the parameters \( \mu_A, \mu^e \) and \( \mu^p \) go to zero).

If \( v = (x, s, y, w) \) is a given approximate zero of \( F \) such that \( x + \mu^e e > 0, s + \mu^p e > 0, z > 0, \) and \( w > 0, \) the Newton equations for the change in variables \((\Delta x, \Delta s, \Delta v, \Delta y, \Delta z, \Delta w)\) are given by

\[
\begin{pmatrix}
H & 0 & -A^T & -J^T & -I & 0 \\
0 & 0 & 0 & I & 0 & -I \\
A & 0 & D_A & 0 & 0 & 0 \\
J & -I & 0 & D_v & 0 & 0 \\
Z & 0 & 0 & 0 & X^\mu & 0 \\
0 & W & 0 & 0 & 0 & S^\mu
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta v \\
\Delta y \\
\Delta z \\
\Delta w
\end{pmatrix}
= -\begin{pmatrix}
g - J^T y - A^T v - z \\
y - w \\
Ax - b + \mu^e(v - v^e) \\
c - s + \mu^p(y - y^e) \\
z \cdot x + \mu^p(z - z^e) \\
w \cdot s + \mu^p(w - w^e)
\end{pmatrix},
\] (5.5)

where \( D_A = \mu^e I, D_v = \mu^p I, X^\mu = \text{diag}(x_j + \mu^p), S^\mu = \text{diag}(s_i + \mu^p), Z = \text{diag}(z_j), \) and \( W = \text{diag}(w_i). \)
5.3. A shifted primal-dual penalty-barrier function

Consider the shifted primal-dual penalty-barrier function

\[
\begin{align*}
    f(x) - (Ax - b)^T v^e + \frac{1}{2\mu_a} \|Ax - b\|^2 + \frac{1}{2\mu_a} \|Ax - b + \mu^a(v - v^e)\|^2 \\
    &- (c(x) - s)^T y^e + \frac{1}{2\mu^p} \|c(x) - s\|^2 + \frac{1}{2\mu^p} \|c(x) - s + \mu^p(y - y^e)\|^2 \\
    &- \sum_{j=1}^n \{\mu^a z^e_j \ln(x_j + \mu^a) + \mu^a z^e_j \ln(z_j(x_j + \mu^a)) + \mu^a(z^e_j - z_j) - z_j x_j \} \\
    &- \sum_{i=1}^m \{\mu^a w^e_i \ln(s_i + \mu^a) + \mu^a w^e_i \ln(w_i(s_i + \mu^a)) + \mu^a(w^e_i - w_i) - w_i s_i \},
\end{align*}
\]

which is well defined for all \(x\) and \(s\) such that \(x + \mu^e e > 0\) and \(s + \mu^e e > 0\). This function has the same gradient as the function \(M(x, s, v, y, z, w; \mu^a, \mu^p, \mu^b, v^e, y^e, z^e, w^e)\) given by

\[
\begin{align*}
    f(x) - (Ax - b)^T v^e + \frac{1}{2\mu_a} \|Ax - b\|^2 + \frac{1}{2\mu_a} \|Ax - b + \mu^a(v - v^e)\|^2 \\
    &- (c(x) - s)^T y^e + \frac{1}{2\mu^p} \|c(x) - s\|^2 + \frac{1}{2\mu^p} \|c(x) - s + \mu^p(y - y^e)\|^2 \\
    &- \sum_{j=1}^n \{\mu^a z^e_j \ln(x_j + \mu^a) + \mu^a z^e_j \ln(z_j(x_j + \mu^a)) - z_j(x_j + \mu^a) \} \\
    &- \sum_{i=1}^m \{\mu^a w^e_i \ln(s_i + \mu^a) + \mu^a w^e_i \ln(w_i(s_i + \mu^a)) - w_i(s_i + \mu^a) \}. \tag{5.6}
\end{align*}
\]

Let \(c, g\) and \(J\) denote the quantities \(c(x), g(x)\) and \(J(x)\). For clarity, the dependence of \(M\) on the parameters \(\mu^p, \mu^e, y^e, z^e\), and \(w^e\), will be suppressed, with \(M(x, s, v, y, z, w)\) being used to denote \(M(x, s, v, y, z, w; \mu^a, \mu^p, \mu^b, v^e, y^e, z^e, w^e)\). This function
5. Nonnegativity Constraints on the Variables and Slacks

may be written in the form:

\[
M(x, s, v, y, z, w) = f - (Ax - b)^T v^e + \frac{1}{2\mu^e} \|Ax - b\|^2 + \frac{1}{2\mu^a} \|Ax - b + \mu^a (v - v^e)\|^2 \\
- (c - s)^T y^e + \frac{1}{2\mu^e} \|c - s\|^2 + \frac{1}{2\mu^c} \|c - s + \mu^c (y - y^e)\|^2 \\
- \sum_{j=1}^{n} \{\mu^e z_j \ln(z_j(x_j + \mu^e)) - z_j(x_j + \mu^e)\} - \sum_{i=1}^{m} \{\mu^w w_i \ln(w_i(s_i + \mu^e)) - w_i(s_i + \mu^e)\}.
\]

(5.7)

Differentiating \(M(x, s, v, y, z, w)\) with respect to \(x, s, v, y, z\) and \(w\) gives

\[
\nabla M(x, s, v, y, z, w) = \begin{pmatrix}
    g - AT(2v^e + \frac{1}{\mu^e}(Ax - b) - v) - J^T(2(y^e - \frac{1}{\mu^e}(c - s)) - y) - 2\mu^e(X^\mu)^{-1}z^e + z \\
    2(y^e - \frac{1}{\mu^e}(c - s)) - y - 2\mu^w(S^\mu)^{-1}w^e + w \\
    Ax - b + \mu^a(v - v^e) \\
    c - s + \mu^c(y - y^e) \\
    x + \mu^e e - \mu^e Z^{-1}z^e \\
    s + \mu^e e - \mu^e W^{-1}w^e
\end{pmatrix},
\]
with $X = \text{diag}(x_1, x_2, \ldots, x_n)$, $S = \text{diag}(s_1, s_2, \ldots, s_m)$, $Z = \text{diag}(z_1, z_2, \ldots, z_n)$ and $W = \text{diag}(w_1, w_2, \ldots, w_m)$. The gradient may be written in several equivalent forms

$$
\nabla M(x, s, z, w) = \begin{pmatrix}
(g - A^T(2(v^e + \frac{1}{\mu^e}(Ax - b)) - v) - J^T(2(y^e - \frac{1}{\mu^e}(c - s)) - y) - 2\mu^e(X^\mu)^{-1}z^e + z) \\
2(y^e - \frac{1}{\mu^e}(c - s)) - y - 2\mu^e(S^\mu)^{-1}w^e + w \\
Ax - b + \mu^e(v - v^e) \\
c - s + \mu^e(y - y^e) \\
x + \mu^e - \mu^eZ^{-1}z^e \\
s + \mu^e - \mu^eW^{-1}w^e \\
g - A^T(2(v^e + \frac{1}{\mu^e}(Ax - b)) - v) - J^T(2(y^e - \frac{1}{\mu^e}(c - s)) - y) - 2\mu^e(X^\mu)^{-1}z^e + z) \\
2(y^e - \frac{1}{\mu^e}(c - s)) - y - 2\mu^e(S^\mu)^{-1}w^e + w \\
Ax - b + \mu^e(v - v^e) \\
c - s + \mu^e(y - y^e) \\
Z^{-1}(z \cdot x + \mu^e(z - z^e)) \\
W^{-1}(w \cdot s + \mu^e(w - w^e)) \\
g - A^T(\pi^A + (\pi^A - v)) - J^T(\pi^V + (\pi^V - y)) - (\pi^Z + (\pi^Z - z)) \\
(\pi^V + (\pi^V - y)) - (\pi^W + (\pi^W - w)) \\
-D_A(\pi^A - v) \\
-D_V(\pi^V - y) \\
-D_Z(\pi^Z - z) \\
-D_W(\pi^W - w)
\end{pmatrix},
$$

where

$$
D_A = \mu^e I, \quad \pi^A = v^e - \frac{1}{\mu^e}(Ax - b), \quad (5.8a)
$$

$$
D_V = \mu^e I, \quad \pi^V = y^e - \frac{1}{\mu^e}(c - s), \quad (5.8b)
$$

$$
D_Z = X^\mu Z^{-1}, \quad \pi^Z = \mu^e(X^\mu)^{-1}z^e, \quad (5.8c)
$$

$$
D_W = S^\mu W^{-1}, \quad \pi^W = \mu^e(S^\mu)^{-1}w^e. \quad (5.8d)
$$
5. Nonnegativity Constraints on the Variables and Slacks

Similarly, the Hessian of $M(x, s, v, y, z, w)$ is given by

$$
\begin{pmatrix}
H_1 & -\frac{2}{\mu} J^T & A^T & J^T & I & 0 \\
-\frac{2}{\mu^2} J & 2\left(\frac{1}{\mu^2} I + \mu\theta(S\theta)^{-2} W\right) & 0 & -I & 0 & I \\
A & 0 & \mu I & 0 & 0 & 0 \\
J & -I & 0 & \mu I & 0 & 0 \\
I & 0 & 0 & 0 & \mu Z^{-2} Z\theta & 0 \\
0 & I & 0 & 0 & 0 & \mu W^{-2} W\theta
\end{pmatrix},
$$

where $H_1 = H(x, 2\pi\gamma - y) + \frac{2}{\mu^2} A^T A + \frac{2}{\mu} J^T J + 2\mu\theta(X\theta)^{-2} Z\theta$. Substituting $\mu\theta Z\theta = X\theta \Pi z$ and $\mu\theta W\theta = S\theta \Pi w$ from (5.8) gives the Hessian

$$
\begin{pmatrix}
H_1 & -\frac{2}{\mu} J^T & A^T & J^T & I & 0 \\
-\frac{2}{\mu^2} J & 2\left(\frac{1}{\mu^2} I + (S\theta)^{-1} \Pi w\right) & 0 & -I & 0 & I \\
A & 0 & \mu I & 0 & 0 & 0 \\
J & -I & 0 & \mu I & 0 & 0 \\
I & 0 & 0 & 0 & Z^{-2} \Pi z X\theta & 0 \\
0 & I & 0 & 0 & 0 & W^{-2} \Pi w S\theta
\end{pmatrix},
$$

where $H_1 = H(x, 2\pi\gamma - y) + \frac{2}{\mu^2} A^T A + \frac{2}{\mu} J^T J + 2(X\theta)^{-1} \Pi z$.

5.4. Derivation of the shifted primal-dual penalty-barrier direction

Now consider the matrix defined by replacing $\pi\gamma$ by $y$, $\pi\theta$ by $z$, and $\pi w$ by $w$, everywhere in $\nabla^2M(x, s, v, y, z, w)$. This gives an approximate Hessian $B(x, s, v, y, z, w)$ of the form

$$
\begin{pmatrix}
\tilde{H}_1 & -\frac{2}{\mu} J^T & A^T & J^T & I & 0 \\
-\frac{2}{\mu^2} J & 2\left(\frac{1}{\mu^2} I + (S\theta)^{-1} W\right) & 0 & -I & 0 & I \\
A & 0 & \mu I & 0 & 0 & 0 \\
J & -I & 0 & \mu I & 0 & 0 \\
I & 0 & 0 & 0 & Z^{-1} X\theta & 0 \\
0 & I & 0 & 0 & 0 & W^{-1} S\theta
\end{pmatrix},
$$
where \( \hat{H}_1 = H(x, y) + \frac{2}{\mu^r} A^T A + \frac{2}{\mu^r} J^T J + 2 (X^\mu)^{-1} Z \). The definitions of \( D_Y, D_Z, \) and \( D_W \) may be used to write \( B(x, s, v, y, z, w) \) in the form

\[
\begin{pmatrix}
H + 2A^T D_A^{-1} A + 2J^T D_Y^{-1} J + 2D_Z^{-1} & -2J^T D_Y^{-1} & 2D_Y^{-1} + D_W^{-1} & A^T & J^T & I & 0 \\
-H^{-1} J & 0 & 0 & D_A & 0 & 0 & 0 \\
A & 0 & 0 & 0 & 0 & 0 & 0 \\
J & 0 & 0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & 0 & D_W
\end{pmatrix},
\]

where \( H = H(x, y) \). A modified Newton direction satisfies

\[
B(x, s, v, y, z, w) d = -\nabla M(x, s, v, y, z, w).
\]

Given any nonsingular matrix \( R \), the modified Newton direction also satisfies

\[
R B(x, s, v, y, z, w) d = -R \nabla M(x, s, v, y, z, w).
\]

In particular, consider the block upper-triangular matrix

\[
R = \begin{pmatrix}
I & 0 & -2A^T D_A^{-1} & -2J^T D_Y^{-1} & -2D_Z^{-1} & 0 & 0 \\
I & 0 & 2D_Y^{-1} & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 \\
Z & 0 & 0 & 0 & 0 & 0 & 0 \\
W & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
\]

which is nonsingular if \( Z \) and \( W \) are positive definite. For this \( R \), the product \( R B(x, s, v, y, z, w) \) is given by

\[
\begin{pmatrix}
I & 0 & -2A^T D_A^{-1} & -2J^T D_Y^{-1} & -2D_Z^{-1} & 0 & 0 \\
I & 0 & 2D_Y^{-1} & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 \\
Z & 0 & 0 & 0 & 0 & 0 & 0 \\
W & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
H + 2A^T D_A^{-1} A + 2J^T D_Y^{-1} J + 2D_Z^{-1} & -2J^T D_Y^{-1} & 2D_Y^{-1} + D_W^{-1} & A^T & J^T & I & 0 \\
H & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & A & 0 & 0 & 0 & 0 & 0 \\
J & 0 & 0 & 0 & 0 & 0 & 0 \\
Z & 0 & 0 & 0 & X^\mu & 0 & 0 \\
0 & W & 0 & 0 & 0 & 0 & S^\mu
\end{pmatrix} = \begin{pmatrix}
H & 0 & -A^T & -J^T & -I & 0 & 0 \\
0 & 0 & I & 0 & -I & 0 & 0 \\
A & 0 & D_A & 0 & 0 & 0 & 0 \\
J & 0 & 0 & D_Y & 0 & 0 & 0 \\
Z & 0 & 0 & 0 & X^\mu & 0 & 0 \\
0 & W & 0 & 0 & 0 & S^\mu & 0
\end{pmatrix}.
\]
Similarly, for the right-hand side vector $R\nabla M(x, s, y, z, w)$ we obtain

$$
\begin{pmatrix}
I & 0 & -2ATD_A^{-1} & -2JT_D^{-1} & -2D_z^{-1} & 0 \\
I & 0 & 0 & -2D_v^{-1} & 0 & -2D_w^{-1} \\
I & 0 & 0 & 0 & -2 & 0 \\
I & 0 & 0 & 0 & 0 & Z \\
0 & 0 & 0 & 0 & 0 & W
\end{pmatrix}
\begin{pmatrix}
g - A^T(\pi^y + (\pi - v)) - J^T(\pi^y + (\pi - y)) - (\pi^x + (\pi^z - z)) \\
\pi^y + (\pi - y) - (\pi^w + (\pi^w - w)) \\
-D_A(\pi^A - v) \\
-D_v(\pi^y - y) \\
-D_z(\pi^z - z) \\
-D_w(\pi^w - w)
\end{pmatrix}
= \begin{pmatrix}
g - A^Tv - J^Ty - z \\
y - w \\
-D_A(\pi^A - v) \\
-D_v(\pi^y - y) \\
-ZD_z(\pi^z - z) \\
-WD_w(\pi^w - w)
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta v \\
\Delta y \\
\Delta z \\
\Delta w
\end{pmatrix}
= \begin{pmatrix}
g - A^Tv - J^Ty - z \\
y - w \\
Ax - b + \mu^A(c - v^k) \\
-c + \mu^v(y - y^k) \\
z \cdot x + \mu^w(z - z^k) \\
w \cdot s + \mu^w(w - w^k)
\end{pmatrix}
,$$

which are equivalent to the path-following equations (5.5) associated with the perturbed optimality conditions (5.3).

### 5.5. Computation of the shifted primal-dual penalty-barrier direction

The path-following equations (5.5) may be written in symmetric form

$$
\begin{pmatrix}
H & 0 & -A^T & -J^T & I & 0 \\
0 & 0 & 0 & -I & 0 & I \\
A & 0 & -D_A & 0 & 0 & 0 \\
J & -I & 0 & -D_v & 0 & 0 \\
I & 0 & 0 & 0 & -D_z & 0 \\
0 & I & 0 & 0 & 0 & -D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
-\Delta v \\
-\Delta y \\
-\Delta z \\
-\Delta w
\end{pmatrix}
= \begin{pmatrix}
g - A^Tv - J^Ty - z \\
y - w \\
Ax - b + \mu^A(c - v^k) \\
-c + \mu^v(y - y^k) \\
z \cdot x + \mu^w(z - z^k) \\
w \cdot s + \mu^w(w - w^k)
\end{pmatrix}
,$$

where $\Delta x, \Delta s, \Delta v, \Delta y, \Delta z, \Delta w$ are the differences between successive iterates.
where \( D_a = \mu^a I, \ D_v = \mu^v I, \ D_z = X^\mu Z^{-1} \) and \( D_w = S^\mu W^{-1} \) from (5.8).

The solution of this system of equations is given by

\[
\begin{align*}
\Delta w &= y - w + \Delta y \\
\Delta s &= -W^{-1}(s \cdot (y + \Delta y) + \mu^a(y + \Delta y - w^s)) \\
\Delta v &= -D_a^{-1}(A(x + \Delta x) + \mu^a (v - v^s)) \\
\Delta z &= -(X^\mu)^{-1}(z \cdot (x + \Delta x) + \mu^a (z - z^s)),
\end{align*}
\]

where \( \Delta x \) and \( \Delta y \) satisfy the KKT system

\[
\begin{pmatrix}
H + A^T D_a^{-1} A + D_z^{-1} J & -J^T \\
J & D_v + D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix} = -\begin{pmatrix}
g - J^T (y - z - A^T \pi^a) + (X^\mu)^{-1}(z \cdot x + \mu^a (z - z^s)) \\
c - s + \mu^v (y - y^s) + W^{-1}(s \cdot y + \mu^v (y - w^v))
\end{pmatrix}.
\]

The right-hand side may be simplified using the identity

\[
(X^\mu)^{-1}(z \cdot x + \mu^a (z - z^s)) = (X + \mu^a I)^{-1}(X + \mu^a I)z - \mu^a z^s
= z - \mu^a(X + \mu^a I)^{-1}z^s
= z - \pi^s.
\]

Similarly,

\[
W^{-1}(s \cdot y + \mu^v (y - w^v)) = W^{-1}((S + \mu^v I)y - \mu^v w^v)
= (S + \mu^v I)W^{-1}(y - \mu^v (S + \mu^v I)^{-1}w^v)
= D_w(y - \mu^v (S + \mu^v I)^{-1}w^v)
= D_w(w - \pi^w).
\]

It follows that the right-hand side is given by

\[
\begin{pmatrix}
g - J^T (y - z - A^T \pi^a) + (X^\mu)^{-1}(z \cdot x + \mu^a (z - z^s)) \\
c - s + \mu^v (y - y^s) + W^{-1}(s \cdot y + \mu^v (y - w^v))
\end{pmatrix} = \begin{pmatrix}
g - J^T (y - \pi^z - A^T \pi^a) \\
c - s + \mu^v (y - y^s) + W^{-1}(s \cdot y + \mu^v (y - w^v))
\end{pmatrix}
\begin{pmatrix}
D_v(y - y^s) + D_w(w - \pi^w)
\end{pmatrix}.
\]
5. Nonnegativity Constraints on the Variables and Slacks

5.6. Summary

The results of Sections 5.1–5.5 imply that the solution of the path-following equations (5.5) may be computed as

\[
\begin{align*}
\hat{y} &= y + \Delta y, \\
\hat{s} &= s + \Delta s, \\
\hat{x} &= x + \Delta x, \\
\end{align*}
\]

\[
\begin{align*}
\Delta s &= -D_w (\hat{y} - \pi^w), \\
\Delta w &= -(S^\mu)^{-1} (w \cdot \hat{s} + \mu^a (w - w^e)), \\
\Delta z &= -(X^\mu)^{-1} (z \cdot \hat{x} + \mu^a (z - z^e)), \\
\Delta v &= \hat{\pi}^a - v,
\end{align*}
\]

where \(\Delta x\) and \(\Delta y\) satisfy the equations

\[
\begin{pmatrix} H(x, y) + A^T D_{\lambda}^{-1} A + D_z^{-1} & -J(x)^T \\ J(x) & D_v + D_w \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = - \begin{pmatrix} g(x) - J(x)^T y - A^T \pi^i - \pi^z \\ D_v (y - \pi^y) + D_w (y - \pi^w) \end{pmatrix},
\]

and \(D_A, D_v, D_z, D_w, \pi^v, \pi^z, \pi^w, \pi^a\) and \(\hat{\pi}^a\) are given by

\[
\begin{align*}
D_A &= \mu^a I, \\
\pi^a &= v^e - \frac{1}{\mu^a} (Ax - b), \\
D_v &= \mu^v I, \\
\pi^v &= y^e - \frac{1}{\mu^v} (c(x) - s), \\
D_z &= X^\mu Z^{-1}, \\
\pi^z &= \mu^a (X^\mu)^{-1} z^e, \\
D_w &= S^\mu W^{-1}, \\
\pi^w &= \mu^a (S^\mu)^{-1} w^e, \\
\hat{\pi}^a &= v^e - \frac{1}{\mu^a} (A\hat{x} - b).
\end{align*}
\]

The associated line-search merit function \(M(x, s, v, y, z, w; \mu^a, \mu^v, \mu^w, v^e, y^e, z^e, w^e)\) is given by

\[
\begin{align*}
f(x) - (Ax - b)^T v^e + \frac{1}{2\mu_a} \|Ax - b\|^2 + \frac{1}{2\mu_v} \|Ax - b + \mu^a (v - v^e)\|^2 \\
- (c(x) - s)^T y^e + \frac{1}{2\mu_v} \|c(x) - s\|^2 + \frac{1}{2\mu_v} \|c(x) - s + \mu^v (y - y^e)\|^2 \\
- \sum_{i=1}^m \{\mu^w_i w^e_i \ln(w_i (s_i + \mu^w_i)^2) - w_i (s_i + \mu^w_i)\} - \sum_{j=1}^n \{\mu^z_j z^e_j \ln(z_j (x_j + \mu^z_j)^2) - z_j (x_j + \mu^z_j)\}.
\end{align*}
\]
6. Fixed and Bounded Slacks with Linear Constraints

Next we consider nonlinear equality constraints and upper and lower bounds on the slacks. The variables are not subject to bounds.

6.1. Problem statement and optimality conditions

The problem has the form

\[
\min_{x \in \mathbb{R}^n, s \in \mathbb{R}^m} f(x) \quad \text{subject to} \quad c(x) - s = 0, \quad L_X s = h_X, \quad \ell \leq L_F s \leq u, \quad (6.1)
\]

where \( c : \mathbb{R}^n \to \mathbb{R}^m \) and \( f : \mathbb{R}^n \to \mathbb{R} \) are twice-continuously differentiable and \( L_X \) and \( L_F \) are fixed matrices of dimension \( m_X \times m \) and \( m_X \times m_F \), respectively, with \( m = m_F + m_X \). The matrices \( L_X \) and \( L_F \) are formed from rows of the identity matrix \( I_m \) in such a way that \( L_X s \) and \( L_F s \) give the fixed and “free” components of \( s \). It follows that there is an \( m \times m \) permutation matrix \( P \) such that

\[
P = \begin{pmatrix} L_F \\ L_X \end{pmatrix},
\]

with the matrices \( L_F \) and \( L_X \) satisfying the identities \( L_F L_F^T = I_F \), \( L_X L_X^T = I_X \), and \( L_F L_X^T = 0 \). The first-order KKT conditions for this problem are

\[
\begin{align*}
g(x^*) - J(x^*)^T y^* &= 0, \\
c(x^*) - s^* &= 0, \\
y^* - L_X^T w_X^* - L_F^T w_1^* + L_F^T w_2^* &= 0, \\
L_F s^* - \ell &\geq 0, \\
w_1^* &\geq 0, \\
L_F s^* - u &\geq 0, \\
w_2^* &\geq 0, \\
L_F s^* - \ell &\geq 0, \\
w_1^* &\cdot (L_F s^* - \ell) = 0, \\
w_2^* &\cdot (u - L_F s^*) = 0, \quad (6.2f)
\end{align*}
\]

where \( y^* \) and \( w_X^* \) are the Lagrange multipliers for the equality constraints \( c(x) - s = 0 \) and \( L_X s = h_X \), and \( w_1^* \) and \( w_2^* \) may be interpreted as the Lagrange multipliers for the inequality constraints \( L_F s - \ell \geq 0 \) and \( u - L_F s \geq 0 \), respectively. Given any \( s \geq 0 \), we define the index set \( \mathcal{F} \) of indices from 1, 2, \ldots, \( m \) that define the rows of \( L_F \).

6.2. The path-following equations

Let \( y^E \) be an estimate of the Lagrange multipliers for the nonlinear equality constraints \( c(x) - s = 0 \). Similarly, let \( w_1^E \) and \( w_2^E \) denote nonnegative estimates of the multipliers for the inequality constraints \( L_F s - \ell \geq 0 \) and \( u - L_F s \geq 0 \), respectively. Given
small positive scalars $\mu^e$ and $\mu^o$, consider the perturbed optimality conditions

\[
\begin{align*}
g(x) &= J(x)^Ty = 0, \\
c(x) - s &= \mu^e(y^e - y), \\
y - L_x^T x - L_p^T w_1 + L_p^T w_2 &= 0, \\
L_p s - \ell &\geq 0, \\
w_1 &\geq 0, \\
w_1 \cdot (L_p s - \ell) &= \mu^b(w_1^b - w_1), \\
w_2 &\geq 0, \\
w_2 \cdot (u - L_p s) &= \mu^o(w_2^o - w_2).
\end{align*}
\]

Consider the following primal-dual path following equations given by

\[
F(x, s, y, w_1, w_2; \mu^e, \mu^o, y^e, w_1^e, w_2^e) = 0,
\]

with

\[
F(x, s, y, w_1, w_2; \mu^e, \mu^o, y^e, w_1^e, w_2^e)
= \begin{pmatrix}
g(x) - J(x)^Ty \\
y - L_x^T x - L_p^T w_1 + L_p^T w_2 \\
c(x) - s + \mu^e(y - y^e) \\
L_x s - h_x \\
w_1 \cdot (L_p s - \ell) + \mu^b(w_1 - w_1^b) \\
w_2 \cdot (u - L_p s) + \mu^o(w_2 - w_2^o)
\end{pmatrix}.
\]

Any zero $(x, s, y, w_1, w_2)$ of $F$ satisfying $\ell < L_p s < u$, $w_1 > 0$, and $w_2 > 0$ approaches a point satisfying the optimality conditions (6.2), with the approximation becoming increasingly accurate as the terms $\mu^e(y - y^e)$, $\mu^b(w_1 - w_1^b)$ and $\mu^o(w_2 - w_2^o)$ approach zero. For any sequence of $y^e$ and $w_2^e$ such that $y^e \to y^*$, $w_1^e \to w_1^*$ and $w_2^e \to w_2^*$, it must hold that solutions $(x, s, y, w_1, w_2)$ of (6.3) must satisfy $y \cdot (c(x) - s) \to 0$, $w_1 \cdot (L_p s - \ell) \to 0$, and $w_2 \cdot (u - L_p s) \to 0$. This implies that any solution $(x, s, y, w_1, w_2)$ of (6.3) will approximate a solution of (6.2) independently of the values of $\mu^e$ and $\mu^o$ (i.e., it is not necessary that $\mu^e, \mu^o \to 0$).

Given an approximate zero $(x, s, y, w_1, w_2)$ of $F$ such that $\ell < L_p s < u$, $w_1 > 0$, and $w_2 > 0$, the Newton equations for the change in variables $(\Delta x, \Delta s, \Delta y, \Delta w_1, \Delta w_2)$ are given by

\[
\begin{pmatrix}
H & 0 & -J^T \\
0 & I_m & -L_x^T & -L_p^T & L_p^T \\
J & -I_m & D_y & 0 & 0 \\
0 & L_x & 0 & 0 & 0 \\
0 & W_1 L_p & 0 & 0 & S_1^m \\
0 & -W_2 L_p & 0 & 0 & S_2^m
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta y \\
\Delta w_1 \\
\Delta w_2
\end{pmatrix}
= -\begin{pmatrix}
g - J^Ty \\
y - L_x^T x - L_p^T w_1 + L_p^T w_2 \\
c - s + \mu^e(y - y^e) \\
L_x s - h_x \\
w_1 \cdot (L_p s - \ell) + \mu^b(w_1 - w_1^b) \\
w_2 \cdot (u - L_p s) + \mu^o(w_2 - w_2^o)
\end{pmatrix},
\]

where $D_y = \mu^o I$, $W_1 = \text{diag}([w_1]_i)$, $W_2 = \text{diag}([w_2]_i)$, $S_1^m = \text{diag}(c_i^T s - \ell_i + \mu^o)$, and $S_2^m = \text{diag}(u_i - c_i^T s + \mu^o)$.
Any $s$ may be written as $s = L_F^T s_F + L_X^T s_X$, where $s_F$ and $s_X$ denote the components of $s$ corresponding to the “free” and “fixed” components of $s$, respectively. Throughout, we assume that $s_X$ satisfies $L_X s = h_X$, in which case the expansion of $\Delta s$ satisfies

$$
\Delta s = L_F^T \Delta s_F + L_X^T \Delta s_X = L_F^T \Delta s_F.
$$

This identity allows us to write the equations (6.5) in the form

$$
\begin{pmatrix}
H & 0 & -J^T & 0 & 0 \\
0 & 0 & L_F & -I_F & I_F \\
J & -L_F^T & D_Y & 0 & 0 \\
0 & W_1 & 0 & S_1^o & 0 \\
0 & -W_2 & 0 & 0 & S_2^o
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s_F \\
\Delta y \\
\Delta w_1 \\
\Delta w_2
\end{pmatrix}
= -
\begin{pmatrix}
g - J^T y \\
y_F - w_1 + w_2 \\
c - s + \mu^p (y - y^o) \\
w_1 \cdot (L_F s - \ell) + \mu^o (w_1 - w_1^e) \\
w_2 \cdot (u - L_F s) + \mu^o (w_2 - w_2^e)
\end{pmatrix}.
$$

(6.6)

The vectors $\Delta s$ and $\Delta w_X$ are recovered as $\Delta s = L_F^T \Delta s_F$ and $\Delta w_X = [y + \Delta y - w]_X$.

### 6.3. A shifted primal-dual penalty-barrier function

Problem (6.1) may be written in the equivalent form

$$
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad c(x) - s = 0, \quad L_F s - s_1 = \ell, \quad s_1 \geq 0, \\
& \quad L_X s - h_X = 0, \quad L_F s + s_2 = u, \quad s_2 \geq 0.
\end{align*}
$$

The nonlinear equality constraints and bounds may be treated using shifted primal-dual penalty-barrier and augmented Lagrangian terms, which gives the approximate problem

$$
\begin{align*}
\text{minimize} & \quad M(x, s, s_1, s_2, y, w_1, w_2; \mu^p, \mu^o, y^o, w_1^e, w_2^e) \\
\text{subject to} & \quad L_X s - h_X = 0, \\
& \quad L_F s - s_1 = \ell, \quad s_1 + \mu^p e > 0, \quad w_1 > 0, \\
& \quad L_F s + s_2 = u, \quad s_2 + \mu^o e > 0, \quad w_2 > 0,
\end{align*}
$$

(6.7)
where \( M(x, s, s_1, s_2, y, w_1, w_2; \mu^p, \mu^b, y^e, w_1^e, w_2^e) \) is the shifted primal-dual penalty-barrier function

\[
M(x, s, s_1, s_2, y, w_1, w_2; \mu^p, \mu^b, y^e, w_1^e, w_2^e) = f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^p} \|c(x) - s\|^2 + \frac{1}{2\mu^p} \|c(x) - s + \mu^p(y - y^e)\|^2
- \sum_{i=1}^{n_L} \left\{ \mu^p[w_1^e_i]\ln\left(\left[ s_1 + \mu^b e \right]_i \right) + \mu^b[w_1^e_i]\ln\left(\left[ w_1 \cdot (s_1 + \mu^b e) \right]_i \right) - \left[ w_1 \cdot (s_1 + \mu^b e) \right]_i \right\}
- \sum_{i=1}^{n_U} \left\{ \mu^b[w_2^e_i]\ln\left(\left[ s_2 + \mu^b e \right]_i \right) + \mu^b[w_2^e_i]\ln\left(\left[ w_2 \cdot (s_2 + \mu^b e) \right]_i \right) - \left[ w_2 \cdot (s_2 + \mu^b e) \right]_i \right\}
\]  

(6.8)

Let \( c, g \) and \( J \) denote the quantities \( c(x), g(x) \) and \( J(x) \). Differentiating \( M(x, s, s_1, s_2, y, w_1, w_2) \) with respect to \( x, s, s_1, s_2, y, w_1 \) and \( w_2 \) gives

\[
\nabla M(x, s, s_1, s_2, y, w_1, w_2) = \begin{pmatrix}
\left( g - J^T(2(y^e - \frac{1}{\mu^p}(c - s)) - y) \right) \\
2(y^e - \frac{1}{\mu^p}(c - s)) - y \\
w_1 - 2\mu^b(S_1^w)^{-1}w_1^e \\
w_2 - 2\mu^b(S_2^w)^{-1}w_2^e \\
c - s + \mu^p(y - y^e) \\
s_1 + \mu^b e - \mu^b W_1^{-1}w_1^e \\
s_2 + \mu^b e - \mu^b W_2^{-1}w_2^e
\end{pmatrix}
\]
where

\[ D_v = \mu^e I_m, \quad \pi^v = y^e - \frac{1}{\mu^e}(c - s), \]  

(6.9a)

\[ D_1^w = S_1^e W_1^{-1}, \quad \pi_1^w = \mu^e (S_1^e)^{-1} w_1^e, \]  

(6.9b)

\[ D_2^w = S_2^e W_2^{-1}, \quad \pi_2^w = \mu^e (S_2^e)^{-1} w_2^e. \]  

(6.9c)

Similarly, the Hessian of \( M(x, s, s_1, s_2, y, w_1, w_2) \) is given by

\[

\begin{pmatrix}
H_1 & -\frac{2}{\mu^e} J^T & 0 & 0 & J^T & 0 & 0 \\
\frac{2}{\mu^e} J & -\frac{2}{\mu^e} J^T & 0 & 0 & J^T & 0 & 0 \\
0 & 0 & 2\mu^e (S_1^e)^{-2} W_1^e & 0 & 0 & I_F & 0 \\
0 & 0 & 0 & 2\mu^e (S_2^e)^{-2} W_2^e & 0 & 0 & I_F \\
J & -I_m & 0 & 0 & \mu^e I_m & 0 & 0 \\
0 & I_F & 0 & 0 & \mu^e W_1^{-2} W_1^e & 0 \\
0 & 0 & I_F & 0 & 0 & \mu^e W_2^{-2} W_2^e \\
0 & 0 & 0 & I_F & 0 & 0 & \mu^e W_2^{-2} W_2^e
\end{pmatrix},
\]
where $H_1 = H(x, 2\pi^y - y) + \frac{2}{\mu^2}J^TJ$. Substituting $\mu^aW_1^x = S_1^a\Pi_1^w$ and $\mu^aW_2^x = S_2^a\Pi_2^w$ from (6.9) gives the Hessian

$$H_1 = \begin{bmatrix}
H_1 & -\frac{2}{\mu^2}J^T & 0 & 0 & J^T & 0 & 0 \\
-\frac{2}{\mu^2}J & \frac{2}{\mu^2}I_m & 0 & 0 & -I_m & 0 & 0 \\
0 & 0 & 2(S_1^a)^{-1}\Pi_1^w & 0 & 0 & I_F & 0 \\
0 & 0 & 0 & 2(S_2^a)^{-1}\Pi_2^w & 0 & 0 & I_F \\
J & -I_m & 0 & 0 & \mu^aI_m & 0 & 0 \\
0 & 0 & I_F & 0 & 0 & S_1^aW_1^{-2}\Pi_1^w & 0 \\
0 & 0 & 0 & I_F & 0 & 0 & S_2^aW_2^{-2}\Pi_2^w \\
\end{bmatrix}.$$ 

### 6.4. Derivation of the shifted primal-dual penalty-barrier direction

The primal-dual penalty-barrier problem may be written in the form

$$\min_{p \in I} M(p) \quad \text{subject to} \quad Cp = b_C,$$

where

$$I = \{p : p = (x, s, s_1, s_2, y, w_1, w_2), \quad \text{with} \quad s_1 + \mu^a e > 0, s_2 + \mu^a e > 0, w_1 > 0, w_2 > 0\},$$

with

$$C = \begin{bmatrix}
0 & L_X & 0 & 0 & 0 & 0 & 0 \\
0 & L_F & -I_F & 0 & 0 & 0 & 0 \\
0 & L_F & 0 & I_F & 0 & 0 & 0 \\
\end{bmatrix}, \quad \text{and} \quad b_C = \begin{bmatrix}
\ell \\
u \\
\end{bmatrix}.$$ 

Let $p \in I$ be given. For the moment, assume that $p$ is not necessarily feasible for the linear constraints, i.e., it may not hold that $L_F s - s_1 = \ell$ and $L_F s + s_2 = u$, in which case $b_C - Cp$ may not be zero. The Newton direction $\Delta p$ is given by the solution of the subproblem

$$\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T \nabla^2 M(p) \Delta p \quad \text{subject to} \quad C \Delta p = b_C - Cp.$$ 

(6.11)

However, instead of solving (6.11), we define a linearly constrained modified Newton method by approximating the Hessian $\nabla^2 M(x, s, s_1, s_2, y, w_1, w_2)$ by a matrix $B(x, s, s_1, s_2, y, w_1, w_2)$. Consider the matrix defined by replacing $\pi^y$ by $y$, $\pi_1^w$ by $w_1$, and $\pi_2^w$ by $w_2$, everywhere in the matrix $\nabla^2 M(x, s, s_1, s_2, y, w_1, w_2)$. This gives an approximate Hessian $B(x, s, s_1, s_2, y, w_1, w_2)$.
of the form
\[
\begin{pmatrix}
\hat{H}_1 & -\frac{2}{\mu^r}J^T & 0 & 0 & J^T & 0 & 0 \\
-\frac{2}{\mu^r}J & \frac{\mu}{\mu^r}I_m & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 2(S_1^u)^{-1}W_1 & 0 & 0 & I_F & 0 \\
0 & 0 & 2(S_2^u)^{-1}W_2 & 0 & 0 & 0 & I_F \\
J & -I_m & 0 & 0 & \mu^rI_m & 0 & 0 \\
0 & 0 & I_F & 0 & 0 & S_1^uW_1^{-1} & 0 \\
0 & 0 & 0 & I_F & 0 & 0 & S_2^uW_2^{-1}
\end{pmatrix},
\]
where \(\hat{H}_1 = H(x,y) + 2J^TD_y^{-1}J\). The definitions of \(D_y, D_1^w,\) and \(D_2^w\) may be used to write \(B(x,s,s_1,s_2,y,w_1,w_2)\) as
\[
\begin{pmatrix}
H + 2J^TD_y^{-1}J & -2J^TD_y^{-1} & 0 & 0 & J^T & 0 & 0 \\
-2D_y^{-1}J & 2D_y^{-1} & 0 & 0 & 0 & -I_m & 0 \\
0 & 0 & 2(D_1^w)^{-1} & 0 & 0 & 0 & I_F \\
0 & 0 & 0 & 2(D_2^w)^{-1} & 0 & 0 & I_F \\
J & -I_m & 0 & 0 & 0 & D_y & 0 \\
0 & 0 & I_F & 0 & 0 & 0 & D_1^w \\
0 & 0 & 0 & I_F & 0 & 0 & D_2^w
\end{pmatrix},
\]
where \(H = H(x,y)\). Given \(B(p) = B(x,s,s_1,s_2,y,w_1,w_2)\), a modified Newton direction is given by the solution of the QP subproblem
\[
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T B(p) \Delta p \quad \text{subject to} \quad C\Delta p = b_c - Cp.
\]
Let \(N\) denote a matrix whose columns form a basis for \(\text{null}(C)\), i.e., \(CN = 0\) and \((C^T \ N)\) is nonsingular. The vector
\[
\Delta p_0 = \begin{pmatrix}
0 \\
0 \\
-(\ell - s + s_1) \\
(u - s - s_2) \\
0 \\
0 \\
0 \\
0
\end{pmatrix} \triangleq \begin{pmatrix}
0 \\
0 \\
r_u \\
r_u \\
0 \\
0 \\
0 \\
0
\end{pmatrix} \quad (6.12)
\]
satisfies \(C\Delta p_0 = b_c - Cp\), and it follows that every feasible \(\Delta p\) may be written in the form
\[
\Delta p = \Delta p_0 + Nd.
\]
This implies that $d$ must satisfy the reduced equations
\[ N^T B(p) N d = -N^T (\nabla M(p) + B(p) \Delta p_0). \]

Consider the particular null-space basis given by
\[ N = \begin{pmatrix}
  I_n & 0 & 0 & 0 & 0 \\
  0 & L_F^T & 0 & 0 & 0 \\
  0 & I_F & 0 & 0 & 0 \\
  0 & -I_F & 0 & 0 & 0 \\
  0 & 0 & I_m & 0 & 0 \\
  0 & 0 & 0 & I_F & 0 \\
  0 & 0 & 0 & 0 & I_F
\end{pmatrix}. \quad (6.13) \]

The definition of $N$ of (6.13) gives the reduced Hessian
\[
N^T B(p) N = 
\begin{pmatrix}
  H + 2J^T D^{-1}_Y J & -2J^T D^{-1}_Y L_F^T & J^T & 0 & 0 \\
  -2L_F D^{-1}_Y J & 2(L_F D^{-1}_Y L_F^T + D^{-1}_w) & -L_F & I_F & -I_F \\
  J & -L_F^T & D_Y & 0 & 0 \\
  0 & I_F & 0 & D^{w}_1 & 0 \\
  0 & -I_F & 0 & 0 & D^{w}_2
\end{pmatrix},
\]
where $D^{-1}_w = (D^{w}_1)^{-1} + (D^{w}_2)^{-1}$. Similarly, the reduced gradient is
\[
N^T \nabla M(p) = 
\begin{pmatrix}
  I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & L_F & I_F & -I_F & 0 & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & I_F & 0 & 0 \\
  0 & 0 & 0 & 0 & 0 & 0 & I_F & 0 \\
  0 & 0 & 0 & 0 & 0 & 0 & 0 & I_F
\end{pmatrix}
\begin{pmatrix}
  g - J^T(\pi^Y + (\pi^Y - y)) \\
  \pi^Y + (\pi^Y - y) \\
  -\pi^w_1 + \pi^w - w_1 \\
  -\pi^w_2 + \pi^w - w_2 \\
  -D_Y(\pi^Y - y) \\
  -D^{w}_1(\pi^w_1 - w_1) \\
  -D^{w}_2(\pi^w_2 - w_2)
\end{pmatrix}
= 
\begin{pmatrix}
  g - J^T(\pi^Y + (\pi^Y - y)) \\
  \pi^w - (\pi^Y - y) - (\pi^w_1 + (\pi^w_1 - w_1)) + (\pi^w_2 + (\pi^w_2 - w_2)) \\
  -D_Y(\pi^Y - y) \\
  -D^{w}_1(\pi^w_1 - w_1) \\
  -D^{w}_2(\pi^w_2 - w_2)
\end{pmatrix}.
Moreover

\[ B(p) \Delta p_0 = \begin{pmatrix}
0 \\
neg2(D^w_1)^{-1} (\ell - L_F s + s_1) \\
neg2(D^w_2)^{-1} (u - L_F s - s_2) \\
0 \\
-(\ell - L_F s + s_1) \\
(u - L_F s - s_2)
\end{pmatrix}
= \begin{pmatrix}
0 \\
neg2(D^w_1)^{-1} r_L \\
neg2(D^w_2)^{-1} r_U \\
0 \\
-neg r_L \\
-r_U
\end{pmatrix},
\]

where \( r_L = \ell - L_F s + s_1 \) and \( r_U = u - L_F s - s_2 \). This implies that \( N^T B(p) \Delta p_0 \) is given by

\[
\begin{pmatrix}
I_n & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & L_F & I_F & -I_F & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_m & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I_F & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_F
\end{pmatrix}
\begin{pmatrix}
0 \\
neg2(D^w_1)^{-1} r_L \\
neg2(D^w_2)^{-1} r_U \\
0 \\
-neg r_L \\
-r_U
\end{pmatrix}
= \begin{pmatrix}
0 \\
neg2((D^w_1)^{-1} r_L + (D^w_2)^{-1} r_U)
\end{pmatrix}.
\]

This gives \( N^T (\nabla M(p) + B(p) \Delta p_0) \) such that

\[
\begin{pmatrix}
g - J^T(\pi^y + (\pi^y - y)) \\
\pi^y + (\pi^y - y_F) - (\pi^w_1 + (\pi^w_1 - w_1)) + (\pi^w_2 + (\pi^w_2 - w_2)) - 2((D^w_1)^{-1} r_L + (D^w_2)^{-1} r_U)
\end{pmatrix}
\]

\[
\begin{pmatrix}
-D_y(\pi^y - y) \\
-D_1^w(\pi^w_1 - w_1) - r_L \\
-D_2^w(\pi^w_2 - w_2) + r_U
\end{pmatrix}.
\]

\[.
\]
The reduced modified equations $N^T B(p) N d = -N^T (\nabla M(p) + B(p) \Delta p_0)$ are then

$$
\begin{pmatrix}
H + 2J^T D_Y^{-1} J & -2J^T D_Y^{-1} L_F^T \\
-2L_F D_Y^{-1} J & 2(L_F D_Y^{-1} L_F^T + D_Y^{-1}) \\
J & -I_F \\
0 & D_Y \\
0 & -I_F
\end{pmatrix}
\begin{pmatrix}
J^T \\
-I_F \\
0 \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5
\end{pmatrix}
= \begin{pmatrix}
\pi^v_F + (\pi^v_F - y_F) - (\pi^w_1 + (\pi^w_1 - w_1)) + (\pi^w_2 + (\pi^w_2 - w_2)) - 2(D^w_1)^{-1} r_L + (D^w_2)^{-1} r_U \\
g - J^T (\pi^v + (\pi^v - y)) \\
-2(D^w_1)^{-1} - 2(D^w_2)^{-1} \\
W_1 \\
W_2
\end{pmatrix}.
$$

Given any nonsingular matrix $R$, the direction $d$ satisfies

$$
RN^T B(p) N d = -RN^T (\nabla M(p) + B(p) \Delta p_0).
$$

In particular, consider

$$
R = \begin{pmatrix}
I_n & 0 & -2J^T D_Y^{-1} \\
I_F & 2L_F D_Y^{-1} & 0 \\
0 & 0 & 0 \\
I_m & 0 & 0 \\
0 & W_1 & 0 \\
W_2 & 0 & 0
\end{pmatrix},
$$

which is nonsingular if $W_1$ and $W_2$ are positive definite, with

$$
R^{-1} = \begin{pmatrix}
I_n & 0 & 2J^T D_Y^{-1} \\
I_F & -2L_F D_Y^{-1} & 0 \\
0 & 0 & W_1^{-1} \\
I_m & 0 & 0 \\
0 & 0 & W_2^{-1}
\end{pmatrix}.
$$
For this $R$, the product $RN^{TB}(p)N$ is given by

$$
\begin{pmatrix}
I_n & 0 & -2I^T D^{-1}_V & 0 \\
I_F & 2L_p D^{-1}_V & -2(D^w_1)^{-1} & 2(D^w_2)^{-1} \\
I_m & 0 & 0 & 0 \\
W_1 & 0 & W_2 & 0
\end{pmatrix}
\begin{pmatrix}
H + 2J^T D^{-1}_V & -2J^T D^{-1}_V L^T_F & J^T & 0 & 0 \\
-2L_p D^{-1}_V & 2(L_p D^{-1}_V L^T_F + D^{-1}_W) & -L_p F & I_F & -I_F \\
J & -L^T_F & D_V & 0 & 0 \\
0 & I_F & 0 & D^w_2 & 0 \\
0 & 0 & -I_F & 0 & 0 \\
0 & 0 & 0 & D^w_2 & 0
\end{pmatrix}
\begin{pmatrix}
H & 0 & -J^T & 0 & 0 \\
0 & 0 & L_F & -I_F & I_F \\
J & -L^T_F & D_V & 0 & 0 \\
0 & W_1 & 0 & W_1 D^w_1 & 0 \\
0 & -W_2 & 0 & 0 & W_2 D^w_2
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
0 \\
0 \\
-W_1 r_L \\
W_2 r_U
\end{pmatrix}
$$

Similarly $RN^{TB}(p)\Delta p_0$ is

$$
\begin{pmatrix}
I_n & 0 & -2I^T D^{-1}_V & 0 \\
I_F & 2L_p D^{-1}_V & -2(D^w_1)^{-1} & 2(D^w_2)^{-1} \\
I_m & 0 & 0 & 0 \\
W_1 & 0 & W_2 & 0
\end{pmatrix}
\begin{pmatrix}
-2((D^w_1)^{-1} r_L + (D^w_2)^{-1} r_U) \\
0 \\
0 \\
0
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
0 \\
-W_1 r_L \\
W_2 r_U
\end{pmatrix}
$$

Putting all this together gives the following transformed unsymmetric reduced modified Newton equations for $d$

$$
\begin{pmatrix}
H & 0 & -J^T & 0 & 0 \\
0 & 0 & L_F & -I_F & I_F \\
J & -L^T_F & D_V & 0 & 0 \\
0 & W_1 & 0 & S^u_1 & 0 \\
0 & -W_2 & 0 & 0 & S^u_2
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5
\end{pmatrix}
= -\begin{pmatrix}
g - J^T y \\
y_F - w_1 + w_2 \\
-D_V (\pi^y - y) \\
-W_1 (D^w_1 (\pi^y_1 - w_1) + r_L) \\
-W_2 (D^w_2 (\pi^y_2 - w_2) - r_U)
\end{pmatrix},
$$
with \(\Delta s = \Delta s_f, \Delta s_1 = \Delta s - (\ell - L_y s + s_1)\) and \(\Delta s_2 = -\Delta s + (u - L_y s - s_2)\). The Newton equations have been derived for arbitrary interior \(s_1\) and \(s_2\), i.e., it is not assumed that \(s_1\) and \(s_2\) satisfy the linear constraints \(L_y s - s_1 = \ell\) and \(L_y s + s_2 = u\). However, unless an extra term is added to the objective function of (6.10) that forces the linear constraints to become feasible, it is necessary to choose feasible \(s_1\) and \(s_2\). In this case, \(L_y s - s_1 = \ell\) and \(L_y s + s_2 = u\), and it follows that \(\Delta s_1 = \Delta s_f\) and \(\Delta s_2 = -\Delta s_f\). This assumption is made for the remainder of this section.

Under the feasibility assumption, if \(S_1\) and \(S_2\) are written in terms of \(s\), i.e., \(S_1 = \text{diag}(e^T_i s - \ell_i)\) and \(S_2 = \text{diag}(u_i - e^T_i s)\), then the equations (6.15) are the Newton equations for the solution of the perturbed optimality conditions (6.3). The variables
s_1 and s_2 may be computed implicitly for the line search, in which case the appropriate merit function is

\[
f - (c - s)^T y^e + \frac{1}{2\mu^e} ||c - s||^2 + \frac{1}{2\mu^e} ||c + s + \mu^p(y - y^e)||^2
\]

\[- \sum_{i \in \mathcal{F}} \{ \mu^p[w_{1i}^e] \ln (s_i - \ell_i + \mu^a) + \mu^a[w_{1i}^e] \ln ([w_{1i}^e](s_i - \ell_i + \mu^a)) - [w_{1i}^e](s_i - \ell_i + \mu^a) \}
\]

\[- \sum_{i \in \mathcal{F}} \{ \mu^a[w_{2i}^e] \ln (u_i - s_i + \mu^a) + \mu^a[w_{2i}^e] \ln ([w_{2i}^e](u_i - s_i + \mu^a)) - [w_{2i}^e](u_i - s_i + \mu^a) \},
\]

where \( \mathcal{F} \) denotes the index set of slacks with upper and lower bounds.

### 6.5. Computation of the shifted primal-dual penalty-barrier direction

Next we consider the solution of the modified Newton equations (6.15), which are written in the form

\[
\begin{pmatrix}
H & 0 & -J^T & 0 & 0 \\
0 & 0 & L_p & -I_p & I_p \\
J & -L_p^T & D_v & 0 & 0 \\
0 & W_2 & 0 & W_1 D_1^w & 0 \\
0 & -W_2 & 0 & 0 & W_2 D_2^w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s_p \\
\Delta y \\
\Delta w_1 \\
\Delta w_2
\end{pmatrix}
=
\begin{pmatrix}
g - J^T y \\
y_p - w_1 + w_2 \\
-D_v(\pi_v^y - y) \\
-W_1(D_1^w(\pi_1^w - w_1)) \\
-W_2(D_2^w(\pi_2^w - w_2))
\end{pmatrix}.
\]

(6.16)

Consider the following reordered set of equations and variables involving (in order) \( \Delta w_1, \Delta w_2, \Delta s_p, \Delta x \) and \( \Delta y \):

\[
\begin{pmatrix}
I_p & 0 & (D_1^w)^{-1} & 0 & 0 \\
0 & I_p & -(D_2^w)^{-1} & 0 & 0 \\
-I_p & I_p & 0 & 0 & L_p \\
0 & 0 & -L_p^T & J & D_v \\
0 & 0 & 0 & -J^T & H
\end{pmatrix}
\begin{pmatrix}
\Delta w_1 \\
\Delta w_2 \\
\Delta s_p \\
\Delta x \\
\Delta y
\end{pmatrix}
=
\begin{pmatrix}
w_1 - \pi_1^w \\
w_2 - \pi_2^w \\
y_p - w_1 + w_2 \\
D_v(y - \pi_v^y) \\
g - J^T y
\end{pmatrix}.
\]

(6.17)

If, as above, \( \tilde{D}_w \) denotes the matrix \( \tilde{D}_w = ((D_1^w)^{-1} + (D_2^w)^{-1})^{-1} \), then applying the nonsingular matrix

\[
\begin{pmatrix}
I_p & 0 & I_p \\
I_p & -I_p & I_p \\
L_p^T \tilde{D}_w & -L_p^T D_w & L_p^T D_w & I_m
\end{pmatrix}
\]

\[I_n\]
on the left and right-hand side of (6.17) gives the block upper-trapezoidal system

\[
\begin{pmatrix}
I & 0 & (D_Y^0)^{-1} & 0 & 0 \\
I & -(D_Y^0)^{-1} & 0 & 0 \\
\bar{D}_W^{-1} & 0 & L_F & D_Y + L_F^T \bar{D}_W L_F \\
J & D_Y & -J^T \\
H & -J^T \\
\end{pmatrix}
\begin{pmatrix}
\Delta w_1 \\
\Delta w_2 \\
\Delta s_F \\
\Delta x \\
\Delta y \\
\end{pmatrix}
= - \begin{pmatrix}
w_1 - \pi_1^w \\
w_2 - \pi_2^w \\
y_F - \pi^w \\
g - J^T y \\
\end{pmatrix},
\]  

(6.18)

where \( \pi^w = \pi_1^w - \pi_2^w \). Solving (6.18) for \( \Delta y \) and \( \Delta s_F \), and using the last two block equations of (6.15) for \( \Delta w_1 \) and \( \Delta w_2 \) gives the solution of the path-following equations as

\[
\Delta s_F = -D_W (y_F + \Delta y_F - \pi^w), \\
\Delta s = L_F^T \Delta s_F, \\
\Delta w_1 = -(S_1^u)^{-1} (w_1 \cdot (L_F(s + \Delta s) - \ell + \mu^v e) - \mu^v w_1^e), \\
\Delta w_2 = -(S_2^u)^{-1} (w_2 \cdot (u - L_F(s + \Delta s) + \mu^v e) - \mu^v w_2^e),
\]

where \( \Delta x \) and \( \Delta y \) satisfy the equations

\[
\begin{pmatrix}
H & -J^T \\
J & D_Y + L_F^T \bar{D}_W L_F \\
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y \\
\end{pmatrix}
= - \begin{pmatrix}
g - J^T y \\
D_Y (y - \pi^v) + L_F^T \bar{D}_W (y_F - \pi^w) \\
\end{pmatrix}.
\]

6.6. Summary: bounded slacks

Consider the quantities

\[
D_Y = \mu^v I, \\
\pi^v = y^k - \frac{1}{\mu^v} (c(x) - s), \\
D_Y^0 = S_Y^u W_Y^{-1}, \\
\pi_1^w = \mu^w (S_1^u)^{-1} w_1^e, \\
D_Y^1 = S_Y^u W_Y^{-1}, \\
\pi_2^w = \mu^w (S_2^u)^{-1} w_2^e, \\
\bar{D}_W = (D_Y^0)^{-1} + (D_Y^1)^{-1}, \\
\pi^w = \pi_1^w - \pi_2^w,
\]
then \( \Delta s, \Delta s_1, \Delta s_2, \Delta w_1 \) and \( \Delta w_1 \) are given by

\[
\hat{y} = y + \Delta y, \quad \Delta s_F = -\bar{D}_w (\hat{y}_F - \pi^w), \\
\Delta s = L_{\bar{p} F} \Delta s_F, \\
\Delta w_x = [\hat{y} - w]_x, \\
\hat{s} = s + \Delta s, \\
\Delta w_1 = -(S_1^u)^{-1} (w_1 \cdot (L_{\bar{p}} \hat{s} - \ell + \mu^s e) - \mu^s w_1^f), \\
\Delta w_2 = -(S_2^u)^{-1} (w_2 \cdot (u - L_{\bar{p}} \hat{s} + \mu^s e) - \mu^s w_2^f),
\]

where \( \Delta x \) and \( \Delta y \) satisfy the equations

\[
\begin{pmatrix}
H \\
- J^T
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix} = 
\begin{pmatrix}
g - J^T y \\
D_{\nu} + L_{\bar{p} F} \bar{D}_w (y_F - \pi^w)
\end{pmatrix}.
\]

The associated line-search merit function is given by

\[
f(x) - (c(x) - s)^T y^u + \frac{1}{2 \mu^u} \|c(x) - s\|^2 + \frac{1}{2 \mu^v} \|c(x) - s + \mu^v (y - y^u)\|^2
\]

\[
- \sum_{i \in F} \mu^s [w_1^e]_i \ln ([w_1^e]_i(c_i^T s - \ell_i + \mu^s)^2) - [w_1^e]_i(c_i^T s - \ell_i + \mu^s))
\]

\[
- \sum_{i \in \mathcal{F}} \mu^u [w_2^f]_i \ln ([w_2^f]_i(u_i - c_i^T s + \mu^u)^2) - [w_2^f]_i(u_i - c_i^T s + \mu^u)) \right) + \frac{1}{2 \mu^v} \|c(x) - s + \mu^v (y - y^u)\|^2
\]

(6.19)
7. Fixed and Bounded Variables

Next we consider nonlinear equality constraints and upper and lower bounds on the variables but only nonnegativity constraints for the slacks.

7.1. Problem statement and optimality conditions

The problem has the form

\[
\begin{align*}
\text{minimize} \quad & f(x) & \quad \text{subject to} \quad & c(x) - s = 0, & s \geq 0, & E_X x = b_X, & \ell \leq E_F x \leq u,
\end{align*}
\]

(7.1)

where \(c: \mathbb{R}^n \rightarrow \mathbb{R}^m\) and \(f: \mathbb{R}^n \rightarrow \mathbb{R}\) are twice-continuously differentiable and \(E_X\) and \(E_F\) are fixed matrices of dimension \(n_F \times n\) and \(n_X \times n\), respectively, with \(n = n_F + n_X\). The matrices \(E_X\) and \(E_F\) are formed from rows of the identity matrix \(I_n\) in such a way that \(E_X x\) and \(E_F x\) give the fixed and “free” components of \(x\). It follows that there is an \(n \times n\) permutation matrix \(P\) such that

\[
P = \begin{pmatrix} E_F \\ E_X \end{pmatrix},
\]

with the matrices \(E_F\) and \(E_X\) satisfying the identities \(E_F E_F^T = I_F\), \(E_X E_X^T = I_X\), and \(E_F E_X^T = 0\). The first-order KKT conditions for this problem are

\[
\begin{align*}
g(x^*) - J(x^*)^T y^* - E_X^T z_X^* - E_F^T z_1^* + E_F^T z_2^* = 0, & \quad z_1^* \geq 0, & z_2^* \geq 0, & \text{(7.2a)} \\
y^* - w^* = 0, & \quad w^* \geq 0, & \text{(7.2b)} \\
c(x^*) - s^* = 0, & \quad s^* \geq 0, & \text{(7.2c)} \\
E_F x^* - \ell \geq 0, & \quad u - E_F x^* \geq 0, & \text{(7.2d)} \\
z_1^* \cdot (E_F x^* - \ell) = 0, & \quad z_2^* \cdot (u - E_F x^*) = 0, & \text{(7.2e)} \\
w^* \cdot s^* = 0, & \text{(7.2f)} \\
E_X x^* - b_X = 0, & \text{(7.2g)}
\end{align*}
\]

where \(y^*\) and \(z_X^*\) are the multipliers for the equality constraints \(c(x) - s = 0\) and \(E_X x = b_X\), and \(z_1^*, z_2^*\) and \(w^*\) may be interpreted as the Lagrange multipliers for the constraints \(E_F x - \ell \geq 0\), \(u - E_F x \geq 0\), and \(s \geq 0\) respectively.
7.2. The path-following equations

Let \( y^e, z_1^e, z_2^e, \) and \( w^e \) denote nonnegative estimates of the Lagrange multipliers for the inequality constraints \( E_\ell x - \ell \geq 0, \) \( u - E_\ell x \geq 0, \) and \( s \geq 0, \) respectively. Given small positive scalars \( \mu^e \) and \( \mu^0, \) consider the perturbed optimality conditions

\[
g(x) - J(x)^T y - E_\ell^T z_x - E_\ell^T z_1 + E_\ell^T z_2 = 0, \quad z_1 \geq 0, \quad z_2 \geq 0, \quad (7.3a)\]

\[
y - w = 0, \quad w \geq 0, \quad (7.3b)\]

\[
c(x) - s = \mu^e (y^e - y), \quad s \geq 0, \quad (7.3c)\]

\[
E_\ell x - \ell \geq 0, \quad u - E_\ell x \geq 0, \quad (7.3d)\]

\[
z_1 \cdot (E_\ell x - \ell) = \mu^0 (z_1^e - z_1), \quad z_2 \cdot (u - E_\ell x) = \mu^0 (z_2^e - z_2), \quad (7.3e)\]

\[
w \cdot s = \mu^e (w^e - w), \quad (7.3f)\]

\[
E_\ell x - b_x = 0. \quad (7.3g)\]

Consider the following primal-dual path following equations given by \( F(x, s, y, z_x, z_1, z_2, w; \mu^e, \mu^0, y^e, z_1^e, z_2^e, w^e) = 0, \) with

\[
F(x, s, y, z_x, z_1, z_2, w; \mu^e, \mu^0, y^e, z_1^e, z_2^e, w^e) = \begin{pmatrix} g(x) - J(x)^T y - E_\ell^T z_x - E_\ell^T z_1 + E_\ell^T z_2 \\ y - w \\ c(x) - s + \mu^e (y^e - y) \\ z_1 \cdot (E_\ell x - \ell) + \mu^0 (z_1 - z_1^e) \\ z_2 \cdot (u - E_\ell x) + \mu^0 (z_2 - z_2^e) \\ w \cdot s + \mu^e (w - w^e) \\ E_\ell x - b_x \end{pmatrix}. \quad (7.4)\]

Any zero \((x, s, y, z_x, z_1, z_2, w)\) of \( F \) that satisfies \( \ell < E_\ell x < u, \) \( z_1 > 0, \) \( z_2 > 0, \) and \( w > 0 \) approximates a point satisfying the optimality conditions (7.2), with the approximation becoming increasingly accurate as the terms \( \mu^e (y - y^e), \mu^0 (z_1 - z_1^e), \mu^0 (z_2 - z_2^e), \) and \( \mu^0 (w - w^e) \) approach zero. For any sequence of \( z_1^e, z_2^e, w^e \) and \( y^e \) such that \( z_1^e \to z_1^*, \) \( z_2^e \to z_2^*, \) \( w^e \to w^*, \) and \( y^e \to y^*; \) it must hold that solutions \((x, s, y, z_x, z_1, z_2, w)\) of (7.3) must satisfy \( z_1 \cdot (E_\ell x - \ell) \to 0, \) \( z_2 \cdot (u - E_\ell x) \to 0, \) and \( w \cdot s \to 0, \) This implies that any solution \((x, s, y, z_x, z_1, z_2, w)\) of (7.3) will approximate a solution of (7.2) independently of the values of \( \mu^e \) and \( \mu^0 \) (i.e., it is not necessary that \( \mu^e \to 0 \) and \( \mu^0 \to 0 \)).

If \((x, s, y, z_x, z_1, z_2, w)\) is a given approximate zero of \( F \) such that \( \ell - \mu^e e < E_\ell x < u + \mu^e e, \) \( s + \mu^0 e > 0, \) \( z_1 > 0, \) \( z_2 > 0, \) and
Once these equations are solved, $\Delta x$ and $\Delta z_X$ are recovered as $\Delta x = E_F^T \Delta x_F$ and $\Delta z_X = [g + H \Delta x - J^T(y + \Delta y)]_X - z_X$.

### 7.3. A shifted primal-dual penalty-barrier function

Problem (7.1) is equivalent to

$$
\min_{x,s_1,s_2,s} \quad f(x)
$$

subject to

$$
c(x) - s = 0, \quad s \geq 0,
E_x x - b_x = 0,
E_F x - x_1 = \ell, \quad x_1 \geq 0,
E_F x + x_2 = u, \quad x_2 \geq 0.
$$
Consider the shifted primal-dual penalty-barrier problem

\[
\begin{align*}
\text{minimize} & \quad M(x, x_1, x_2, s, y, z_1, z_2, w; \mu^e, \mu^s, y^e, z_1^e, z_2^e, w^e) \\
\text{subject to} & \quad E_x x = b_x, \quad E_{x^T} x - x_1 = \ell, \quad x_1 + \mu^s e > 0, \quad z_1 > 0, \\
& \quad E_{x^T} x + x_2 = u, \quad x_2 + \mu^s e > 0, \quad z_2 > 0,
\end{align*}
\]

(7.7)

where \( M(x, x_1, x_2, s, y, z_1, z_2, w; \mu^e, \mu^s, y^e, z_1^e, z_2^e, w^e) \) is the barrier function

\[
f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^s} ||c(x) - s||^2 + \frac{1}{2\mu^s} ||c(x) - s + \mu^s(y - y^e)||^2 \\
- \sum_{j \in F} \left\{ \mu^s z_1^e, [z_1^e]_j \ln \left( (|x_1|_j + \mu^s) + \mu^s [z_1^e]_j \ln \left( ([x_1]_j + \mu^s) - [z_1]_j ([x_1]_j + \mu^s) \right) \right) \right\} \\
- \sum_{j \in F} \left\{ \mu^s z_2^e, [z_2^e]_j \ln \left( ([x_2]_j + \mu^s) + \mu^s [z_2^e]_j \ln \left( ([x_2]_j + \mu^s) - [z_2]_j ([x_2]_j + \mu^s) \right) \right) \right\} \\
- \sum_{i=1}^m \left\{ \mu^s w_i^e, \ln (s_i + \mu^s) + \mu^s w_i^e, \ln \left( w_i (s_i + \mu^s) \right) - w_i (s_i + \mu^s) \right\}.
\]

(7.8)

Let \( c, g \) and \( J \) denote the quantities \( c(x), g(x) \) and \( J(x) \). Differentiating \( M(x, x_1, x_2, s, y, z_1, z_2, w) \) with respect to \( x, x_1, x_2, s, y, z_1, z_2 \), and \( w \) gives

\[
\nabla M(x, x_1, x_2, s, y, z_1, z_2, w) = \left( \begin{array}{c} \\
\end{array} \begin{array}{c} \\
g - J^T (2(y^e - \frac{1}{\mu^s} (c - s)) - y) \\
\end{array} \right) \\
\begin{array}{c} \\
z_1 - 2\mu^s (X_1^T)^{-1} z_1^e \\
z_2 - 2\mu^s (X_2^T)^{-1} z_2^e \\
2(y^e - \frac{1}{\mu^s} (c - s)) - y - 2\mu^s (S^T)^{-1} w^e + w \\
\end{array} \\
\begin{array}{c} \\
c - s + \mu^s (y - y^e) \\
x_1 + \mu^s e - \mu^s Z_1^{-1} z_1^e \\
x_2 + \mu^s e - \mu^s Z_2^{-1} z_2^e \\
s + \mu^s e - \mu^s W^{-1} w^e \\
\end{array} \right).
\]
The gradient may be written in several equivalent forms

\[
\nabla M(x, x_1, x_2, s, y, z_1, z_2, w) = \begin{pmatrix}
    g - J^T (2(y^e - \frac{1}{\mu^e}(c - s)) - y) \\
    z_1 - 2\mu^e (X_1^\mu)^{-1}z_1^e \\
    z_2 - 2\mu^e (X_2^\mu)^{-1}z_2^e \\
    2(y^e - \frac{1}{\mu^e}(c - s)) - y - 2\mu^e (S^\mu)^{-1}w^e + w \\
    c - s + \mu^e (y - y^e) \\
    x_1 + \mu^e e - \mu^e Z_1^{-1}z_1^e \\
    x_2 + \mu^e e - \mu^e Z_2^{-1}z_2^e \\
    s + \mu^e e - \mu^e W^{-1}w^e
\end{pmatrix}
\]

where now, \(X_1^\mu = \text{diag}([x_1])\) and \(X_2^\mu = \text{diag}([x_2])\), with

\[
\begin{align*}
    D_v &= \mu^e I, & \pi^v &= y^e - \frac{1}{\mu^e}(c - s), \\
    D_1^\gamma &= X_1^\mu Z_1^{-1}, & \pi_1^\gamma &= \mu^e (X_1^\mu)^{-1}z_1^e, \\
    D_2^\gamma &= X_2^\mu Z_2^{-1}, & \pi_2^\gamma &= \mu^e (X_2^\mu)^{-1}z_2^e, \\
    D_w &= S^\mu W^{-1}, & \pi^w &= \mu^e (S^\mu)^{-1}w^e.
\end{align*}
\]
Similarly, the Hessian of $M(x, x_1, x_2, s, y, z_1, z_2, w)$ is given by

$$
\begin{pmatrix}
H_1 & 0 & 0 & -\frac{2}{\mu^e} J^T & J^T & 0 & 0 & 0 \\
0 & 2\mu^e(X_1^e)^{-2} Z_1^e & 0 & 0 & 0 & I_F & 0 & 0 \\
0 & 0 & 2\mu^e(X_2^e)^{-2} Z_2^e & 0 & 0 & 0 & I_F & 0 \\
-\frac{2}{\mu^e} J & 0 & 0 & 2(\frac{1}{\mu^e} I + \mu^e(S^e)^{-2} W^e) & -I_m & 0 & 0 & I_m \\
J & 0 & 0 & -I_m & \mu^e I_m & 0 & 0 & 0 \\
0 & I_F & 0 & 0 & 0 & \mu^e Z_1^e Z_1^e & 0 & 0 \\
0 & 0 & I_F & 0 & 0 & 0 & \mu^e Z_2^e Z_2^e & 0 \\
0 & 0 & 0 & I_m & 0 & 0 & 0 & \mu^e W^{-2} W^e \\
\end{pmatrix},
$$

where $H_1 = H(x, 2\pi^v - y) + \frac{2}{\mu^e} J^T J$. Substituting $\mu^e Z_1^e = X_1^e I_1^e$, $\mu^e Z_2^e = (X_2 + \mu^e I) I_2^e$, and $\mu^e W^e = S^e I^e$ from (7.9) gives the Hessian

$$
\begin{pmatrix}
H_1 & 0 & 0 & -\frac{2}{\mu^e} J^T & J^T & 0 & 0 & 0 \\
0 & 2(X_1^e)^{-1} I_1^e & 0 & 0 & 0 & I_F & 0 & 0 \\
0 & 0 & 2(X_2^e)^{-1} I_2^e & 0 & 0 & 0 & I_F & 0 \\
-\frac{2}{\mu^e} J & 0 & 0 & 2(\frac{1}{\mu^e} I + (S^e)^{-1} I^e) & -I_m & 0 & 0 & I_m \\
J & 0 & 0 & -I_m & \mu^e I_m & 0 & 0 & 0 \\
0 & I_F & 0 & 0 & 0 & X_1^e Z_1^e Z_1^e & 0 & 0 \\
0 & 0 & I_F & 0 & 0 & 0 & X_2^e Z_2^e Z_2^e & 0 \\
0 & 0 & 0 & I_m & 0 & 0 & 0 & S^e W^{-2} W^e \\
\end{pmatrix}.
$$

### 7.4. Derivation of the shifted primal-dual penalty-barrier direction

The primal-dual penalty-barrier problem may be written in the form

$$
\min_{p \in I} M(p) \quad \text{subject to} \quad Cp = b_C,
$$

(7.10)

where

$$
I = \{ p : p = (x, x_1, x_2, s, y, z_1, z_2, w), \text{ with } x_1 + \mu^e e > 0, x_2 + \mu^e e > 0, s + \mu^e e > 0, z_1 > 0, z_2 > 0, w > 0 \},
$$

and

$$
C = \begin{pmatrix}
E_x & 0 & 0 & 0 & 0 & 0 & 0 \\
E_F & -I_F & 0 & 0 & 0 & 0 & 0 \\
E_F & 0 & I_F & 0 & 0 & 0 & 0 \\
\end{pmatrix}, \quad \text{and} \quad b_C = \begin{pmatrix} b_x \\ \ell \\ u \end{pmatrix}.$$
Let \( p \in I \) be given. Assume that \( x \) is feasible for the equality constraints \( E_x x = b_x \), but not necessarily for the linear inequality constraints, i.e., it may not hold that \( E_x x - x_1 = \ell \) and \( E_x x + x_2 = u \). The Newton direction \( \Delta p \) is given by the solution of the subproblem

\[
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T \nabla^2 M(p) \Delta p \quad \text{subject to} \quad C \Delta p = b_c - C p.
\] (7.11)

However, instead of solving (7.11), we define a linearly constrained modified Newton method by approximating the Hessian \( \nabla^2 M(x,x_1,x_2,s,y,z_1,z_2,w) \) by a matrix \( B(x,x_1,x_2,s,y,z_1,z_2,w) \). Consider the matrix defined by replacing \( \pi_Y \) by \( y \), \( \pi_Z^1 \) by \( z_1 \), \( \pi_Z^2 \) by \( z_2 \), and \( \pi_W \) by \( w \) everywhere in the matrix \( \nabla^2 M(x,x_1,x_2,s,y,z_1,z_2,w) \). This gives an approximate Hessian \( B(x,x_1,x_2,s,y,z_1,z_2,w) \) of the form

\[
\begin{pmatrix}
\hat{H}_1 & 0 & 0 & -\frac{2}{\mu} J^T & J^T & 0 & 0 & 0 \\
0 & 2(X_1^\mu)^{-1} Z_1 & 0 & 0 & 0 & I & 0 & 0 \\
0 & 0 & 2(X_2^\mu)^{-1} Z_2 & 0 & 0 & 0 & I & 0 \\
-\frac{2}{\mu^2} J & 0 & 0 & 2(\frac{1}{\mu^2} I + (S^\mu)^{-1} W) & -I & 0 & 0 & I \\
J & 0 & 0 & -I & \mu^2 I & 0 & 0 & 0 \\
0 & I & 0 & 0 & 0 & (X_1^\mu) Z_1^{-1} & 0 & 0 \\
0 & 0 & I & 0 & 0 & 0 & X_2^\mu Z_2^{-1} & 0 \\
0 & 0 & 0 & I & 0 & 0 & 0 & S^\mu W^{-1}
\end{pmatrix},
\]

where \( \hat{H}_1 = H(x,y) + \frac{2}{\mu} J^T J \). The definitions of \( D_Y \), \( D_1^Z \), and \( D_2^Z \) may be used to write \( B(x,x_1,x_2,s,y,z_1,z_2,w) \) in the form

\[
\begin{pmatrix}
H + 2J^T D_Y^{-1} J & 0 & 0 & -2J^T D_1^Z & J^T & 0 & 0 & 0 \\
0 & 2(D_1^Z)^{-1} & 0 & 0 & 0 & I_p & 0 & 0 \\
0 & 0 & 2(D_2^Z)^{-1} & 0 & 0 & 0 & I_p & 0 \\
-2D_1^{-1} J & 0 & 0 & 2(D_1^{-1} + D_2^{-1}) & -I_m & 0 & 0 & I_m \\
J & 0 & 0 & -I & D_Y & 0 & 0 & 0 \\
0 & I_p & 0 & 0 & 0 & D_1^Z & 0 & 0 \\
0 & 0 & I_p & 0 & 0 & 0 & D_2^Z & 0 \\
0 & 0 & 0 & I_m & 0 & 0 & 0 & D_w
\end{pmatrix},
\]

where \( H = H(x,y) \). Given \( B(p) = B(x,x_1,x_2,s,y,z_1,z_2,w) \), a modified Newton direction is given by the solution of the QP subproblem

\[
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T B(p) \Delta p \quad \text{subject to} \quad C \Delta p = b_c - C p.
\] (7.12)
Let $N$ denote a matrix whose columns form a basis for null($C$), i.e., the columns of $N$ are linearly independent and $CN = 0$. The vector

$$
\Delta p_0 = \begin{pmatrix}
0 \\
-(\ell - E_F x + x_1) \\
(u - E_F x - x_2) \\
0 \\
0 \\
0 \\
0 \\
0
\end{pmatrix} \oplus \begin{pmatrix}
0 \\
-r_c \\
r_v \\
0 \\
0 \\
0 \\
0 \\
0
\end{pmatrix}
$$

(7.13)

satisfies $C\Delta p_0 = b_c - Cp$, and every feasible $\Delta p$ may be written in the form

$$\Delta p = \Delta p_0 + Nd.$$ 

This implies that $d$ satisfies the reduced equations

$$N^TB(p)Nd = -N^T(\nabla M(p) + B(p)\Delta p_0).$$

Consider the null-space basis defined from the columns of

$$N = \begin{pmatrix}
E_F^T & 0 & 0 & 0 & 0 & 0 \\
I_F & 0 & 0 & 0 & 0 & 0 \\
-I_F & 0 & 0 & 0 & 0 & 0 \\
0 & I_m & 0 & 0 & 0 & 0 \\
0 & 0 & I_m & 0 & 0 & 0 \\
0 & 0 & 0 & I_F & 0 & 0 \\
0 & 0 & 0 & 0 & I_F & 0 \\
0 & 0 & 0 & 0 & 0 & I_m
\end{pmatrix}. 
$$

(7.14)

The definition of $N$ of (7.14) gives the reduced Hessian $N^TB(p)N$ such that

$$
\begin{pmatrix}
H_F + 2J_F^TD^{-1}_y J_F + 2((D^2_1)^{-1} + (D^2_2)^{-1}) & -2J_F^TD^{-1}_y & J_F^T & I_F & -I_F & 0 \\
-J_F^T & -I_m & 0 & 0 & 0 & I_m \\
0 & 0 & D^2_1 & 0 & 0 \\
I_F & 0 & 0 & D^2_2 & 0 \\
-I_F & 0 & 0 & 0 & D_w \\
0 & I_m & 0 & 0 & 0 & D_w
\end{pmatrix}.
$$
Similarly, the reduced gradient $N^T\nabla M(p)$ is given by

$$
N^T \begin{pmatrix}
    g_p - J_p^T \left( \pi^y + (\pi^y - y) \right) \\
    -2\pi_1^z - z_1 \\
    -2\pi_2^z - z_2 \\
    (2\pi^v - y) - (2\pi^w - w) \\
    -D_v(\pi^v - y) \\
    -D_1^z(\pi_1^z - z_1) \\
    -D_2^z(\pi_2^z - z_2) \\
    -D_w(\pi^w - w)
\end{pmatrix} =
\begin{pmatrix}
    g_p - J_p^T \left( \pi^y + (\pi^y - y) \right) - (2\pi_1^z - z_1) + (2\pi_2^z - z_2) \\
    (2\pi^v - y) - (2\pi^w - w) \\
    -D_v(\pi^v - y) \\
    -D_1^z(\pi_1^z - z_1) \\
    -D_2^z(\pi_2^z - z_2) \\
    -D_w(\pi^w - w)
\end{pmatrix}.
$$

Moreover

$$
B\Delta p_0 = \begin{pmatrix}
    0 \\
    -2(D_1^z)^{-1}r_L \\
    2(D_2^z)^{-1}r_U \\
    0 \\
    0 \\
    -r_L \\
    r_U \\
    0
\end{pmatrix},
$$

where $r_L = \ell - E_p x + x_1$ and $r_U = u - E_p x - x_2$. This implies that $N^TB(p)\Delta p_0$ is given by

$$
\begin{pmatrix}
    -2((D_1^z)^{-1}r_L + (D_2^z)^{-1}r_U) \\
    0 \\
    0 \\
    -r_L \\
    r_U \\
    0
\end{pmatrix}.
$$

This gives the reduced gradient $N^T(\nabla M(p) + B(p)\Delta p_0)$ such that

$$
N^T(\nabla M(p) + B(p)\Delta p_0) =
\begin{pmatrix}
    g_p - J_p^T \left( \pi^y + (\pi^y - y) \right) - (2\pi_1^z - z_1) + (2\pi_2^z - z_2) - 2((D_1^z)^{-1}r_L + (D_2^z)^{-1}r_U) \\
    (2\pi^v - y) - (2\pi^w - w) \\
    -D_v(\pi^v - y) \\
    -D_1^z(\pi_1^z - z_1) - r_L \\
    -D_2^z(\pi_2^z - z_2) + r_U \\
    -D_w(\pi^w - w)
\end{pmatrix}.
$$
The reduced modified Newton equations \( N^T B(p) N d = -N^T (\nabla M(p) + B(p) \Delta p_0) \) are then

\[
\begin{pmatrix}
H_p + 2 J_p^T D_v^{-1} J_p + 2((D_1^z)^{-1} + (D_2^z)^{-1}) & -2 J_p^T D_v^{-1} & J_p & I_p & -I_p & 0 \\
2 D_v^{-1} & 2(D_v^{-1} + D_w^{-1}) & -I_m & 0 & 0 & I_m \\
J_p & -I_m & D_v & 0 & 0 & 0 \\
I_p & 0 & 0 & D_1^z & 0 & 0 \\
-I_p & 0 & 0 & D_2^z & 0 & 0 \\
0 & I_m & 0 & 0 & 0 & D_w
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5 \\
d_6
\end{pmatrix}
= -
\begin{pmatrix}
g_f - J_p^T (\pi^y + (\pi^y - y)) - 2\pi_1^z - z_1)^2 + (2\pi_2^z - z_2)^2 - 2(\pi_1^z)^{-1} r_1 + (\pi_2^z)^{-1} r_2) \\
(2\pi^y - y) - (2\pi^w - w) \\
-2\pi^y - y \\
-D_1^z(z_1 - r_l) \\
-D_2^z(z_2^2 - z_2 + r_v) \\
-D_1^z(z_1^w - y)
\end{pmatrix}.
\]

Given any nonsingular matrix \( R \), the direction \( d \) satisfies

\[
RN^T B(p) N d = -RN^T (\nabla M(p) + B(p) \Delta p_0).
\]

In particular, if \( R \) is the block upper-triangular matrix \( R \) such that

\[
R =
\begin{pmatrix}
I_p & 0 & -2 J_p^T D_v^{-1} & 2(D_1^z)^{-1} & 2(D_2^z)^{-1} & 0 \\
I_m & 0 & 0 & 0 & 0 & -2D_w^{-1} \\
J_p & -I_m & D_v & 0 & 0 & 0 \\
Z_1 & 0 & 0 & 0 & 0 & 0 \\
Z_2 & 0 & 0 & 0 & 0 & 0 \\
W & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
\]

then \( R \) is nonsingular because \( Z_1, Z_2 \) and \( W \) are positive definite, and

\[
RN^T B(p) N = \begin{pmatrix}
H_p & 0 & -J_p^T & -I_p & I_p & 0 \\
0 & 0 & I_m & 0 & 0 & -I_m \\
J_p & -I_m & D_v & 0 & 0 & 0 \\
Z_1 & 0 & 0 & Z_1 D_1^z & 0 & 0 \\
-Z_2 & 0 & 0 & Z_2 D_2^z & 0 & 0 \\
0 & W & 0 & 0 & 0 & WD_w
\end{pmatrix} = \begin{pmatrix}
H_p & 0 & -J_p^T & -I_p & I_p & 0 \\
0 & 0 & I_m & 0 & 0 & -I_m \\
J_p & -I_m & D_v & 0 & 0 & 0 \\
Z_1 & 0 & 0 & X_1^\mu & 0 & 0 \\
-Z_2 & 0 & 0 & X_2^\mu & 0 & 0 \\
0 & W & 0 & 0 & 0 & S^\mu
\end{pmatrix}.
\]
Also, $RNT(T\nabla M(p) + B(p)\Delta p_0)$ is given by

\[
\begin{pmatrix}
I_f & 0 & -2J_f^T D_f^{-1} & -2(D_f^2)^{-1} & 2(D_f^2)^{-1} & 0 \\
I_m & 2D_f^{-1} & 0 & 0 & 0 & -2D_f^{-1} \\
I_m & 0 & 0 & 0 & 0 & Z_1 \\
I_m & 0 & 0 & 0 & 0 & Z_2 \\
I_m & 0 & 0 & 0 & 0 & W
\end{pmatrix}
\begin{pmatrix}
g_f - J_f^T (\pi^V + (\pi^V - y)) - (2\pi_1^V - z_1) + (2\pi_2^V - z_2) - 2((D_f^2)^{-1} r_l + (D_f^2)^{-1} r_u) \\
-2D_f (\pi^V - y) \\
-D_f (\pi^V - y) \\
-D_f (\pi^V - y) \\
-D_w (\pi^V - y)
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta x_1 \\
\Delta s \\
\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w
\end{pmatrix}
\begin{pmatrix}
\Delta p \equiv \Delta p_0 + N d
\end{pmatrix}.

This gives the following (unsymmetric) reduced modified Newton equations for $d$

\[
\begin{pmatrix}
H_f & 0 & -J_f^T & -I_f^T & -I_f & 0 \\
0 & 0 & I_m & 0 & 0 & -I_m \\
J_f & -I_m & D_f & 0 & 0 & 0 \\
Z_1 & 0 & 0 & X_1^\mu & 0 & 0 \\
-Z_2 & 0 & 0 & 0 & X_2^\mu & 0 \\
0 & W & 0 & 0 & 0 & S^\mu
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5 \\
d_6
\end{pmatrix}
\begin{pmatrix}
g_f - J_f^T y - z_1 + z_2 \\
-2D_f (\pi^V - y) \\
-D_f (\pi^V - y) \\
-D_f (\pi^V - y) \\
-D_w (\pi^V - w)
\end{pmatrix}.

(7.15)

Then, (7.13) implies that

\[
\begin{pmatrix}
\Delta x \\
\Delta x_1 \\
\Delta s \\
\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w
\end{pmatrix} = \Delta p = \Delta p_0 + N d =
\begin{pmatrix}
E_f^T d_1 \\
(d_1 - r_l) \\
-(d_1 - r_u) \\
d_2 \\
d_3 \\
d_4 \\
d_5 \\
d_6
\end{pmatrix}.
\]
These identities allow us to write equations (7.15) in the form
\[
\begin{pmatrix}
H_p & 0 & -J_p^T & -I_p & I_p & 0 \\
0 & 0 & I_m & 0 & 0 & -I_m \\
J_p & -I_m & D_p & 0 & 0 & 0 \\
-I_p & 0 & 0 & D_p^2 & 0 & 0 \\
-I_p & 0 & 0 & 0 & D_w & 0 \\
0 & I_m & 0 & 0 & 0 & D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x_p \\
\Delta s \\
\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w
\end{pmatrix}
= -
\begin{pmatrix}
g_p - J_p^T y - z_1 + z_2 \\
g_p - y - w \\
c - s + \mu^e (y - y^e) \\
z_1 \cdot (E_p x - \ell) + \mu^e (z_1 - z_1^e) \\
z_2 \cdot (u - E_p x) + \mu^e (z_2 - z_2^e) \\
w \cdot s + \mu^e (w - w^e)
\end{pmatrix},
\] (7.16)

with \(\Delta x = E_p^T \Delta x_p\), \(\Delta x_1 = \Delta x - (\ell - E_p x + x_1)\) and \(\Delta x_2 = -\Delta x + (u - E_p x - x_2)\).

As in the bounded slack case, it is necessary to choose feasible \(x_1\) and \(x_2\), which gives \(E_p x - x_1 = \ell\) and \(E_p x + x_2 = u\), and it follows that \(\Delta x_1 = \Delta x\) and \(\Delta x_2 = -\Delta x\). (This assumption is made for the remainder of this section.) Under this feasibility assumption, if \(X_1\) and \(X_2\) are written in terms of \(x\), i.e., \(X_1 = \text{diag}(c^T x - \ell)\) and \(X_2 = \text{diag}(u - c^T x)\), respectively, then equations (7.16) are the Newton equations for a solution of the perturbed optimality conditions (7.3). The variables \(x_1\) and \(x_2\) may be computed implicitly for the line search, in which case the appropriate merit function is

\[
f - (c - s)^T y^e + \frac{1}{2 \mu^e} ||c - s||^2 + \frac{1}{2 \mu^e} ||c - s + \mu^e (y - y^e)||^2
- \sum_{j \in \mathcal{F}} \{ \mu^e[z_1^e]_j \ln (x_j - \ell_j + \mu^e) + \mu^e[z_2^e]_j \ln (z_1_j (x_j - \ell_j + \mu^e) - [z_1]_j (x_j - \ell_j + \mu^e))
- \sum_{j \in \mathcal{F}} \{ \mu^e[z_2^e]_j \ln (u_j - x_j + \mu^e) + \mu^e[z_2^e]_j \ln ([z_2]_j (u_j - x_j + \mu^e) - [z_2]_j (u_j - x_j + \mu^e))
- \sum_{i=1}^m \{ \mu^e w_i^e \ln (s_i + \mu^e) + \mu^e w_i^e \ln (w_i (s_i + \mu^e)) - w_i (s_i + \mu^e)\}.
\]

### 7.5. Computation of the shifted primal-dual penalty-barrier direction

Next we consider the solution of the path-following modified Newton equations (7.16), which may be written in the form

\[
\begin{pmatrix}
H_p & 0 & -J_p^T & -I_p & I_p & 0 \\
0 & 0 & I_m & 0 & 0 & -I_m \\
J_p & -I_m & D_p & 0 & 0 & 0 \\
-I_p & 0 & 0 & D_p^2 & 0 & 0 \\
-I_p & 0 & 0 & 0 & D_w & 0 \\
0 & I_m & 0 & 0 & 0 & D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x_p \\
\Delta s \\
\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w
\end{pmatrix}
= -
\begin{pmatrix}
g_p - J_p^T y - z_1 + z_2 \\
g_p - y - w \\
c - s + \mu^e (y - y^e) \\
z_1 \cdot (E_p x - \ell) + \mu^e (z_1 - z_1^e) \\
z_2 \cdot (u - E_p x) + \mu^e (z_2 - z_2^e) \\
w \cdot s + \mu^e (w - w^e)
\end{pmatrix}.
\]
Consider the following reordered equations and variables:

\[
\begin{pmatrix}
  D^z_1 & 0 & 0 & 0 & I_p & 0 \\
  0 & D^z_2 & 0 & 0 & -I_p & 0 \\
  0 & 0 & D_w & I_m & 0 & 0 \\
  0 & 0 & -I_m & 0 & 0 & I_m \\
  -I_p & I_p & 0 & 0 & H_F & -J_F^T
\end{pmatrix}
\begin{pmatrix}
  \Delta z_1 \\
  \Delta z_2 \\
  \Delta w \\
  \Delta s \\
  \Delta x
\end{pmatrix}
= 
\begin{pmatrix}
  D^z_1(z_1 - \pi^z_1) \\
  D^z_2(z_2 - \pi^z_2) \\
  D_w(w - \pi^w) \\
  y - w \\
  \varrho_F - J_F^T y - z_1 + z_2
\end{pmatrix}.
\]

(7.17)

Applying the nonsingular matrix

\[
\begin{pmatrix}
  I_p & 0 & I_p & 0 & I_m & D_{w}^{-1} & I_m & D_{w} & I_m \\
  0 & I_p & 0 & I_m & D_{w}^{-1} & I_m & D_{w} & I_m \\
  (D^z_1)^{-1} & -(D^z_2)^{-1}
\end{pmatrix}
\]

to both sides of (7.17) gives the block upper-trapezoidal system

\[
\begin{pmatrix}
  D^z_1 & 0 & 0 & 0 & I_p & 0 \\
  0 & D^z_2 & 0 & 0 & -I_p & 0 \\
  0 & 0 & D_w & I_m & 0 & 0 \\
  0 & 0 & D_{w}^{-1} & 0 & I_m & 0 \\
  0 & 0 & 0 & J & D_{v} + D_w & 0 \\
  0 & 0 & 0 & 0 & H_F + D_{z}^{-1} & -J_F^T
\end{pmatrix}
\begin{pmatrix}
  \Delta z_1 \\
  \Delta z_2 \\
  \Delta w \\
  \Delta s \\
  \Delta x_F \\
  \Delta y
\end{pmatrix}
= 
\begin{pmatrix}
  D^z_1(z_1 - \pi^z_1) \\
  D^z_2(z_2 - \pi^z_2) \\
  D_w(w - \pi^w) \\
  y - \pi^w \\
  \varrho_F - J_F^T y - \pi^z
\end{pmatrix}.
\]

with \( \pi^z = \pi^z_1 - \pi^z_2 \), and \( D_{z}^{-1} = (D^z_1)^{-1} + (D^z_2)^{-1} \). It follows that the solution of the Newton path-following equations (7.5) is given by

\[
\begin{align*}
\Delta x &= E_F^T \Delta x_F, \\
\Delta z_X &= [g + H \Delta x - J^T(y + \Delta y)]_X - z_X \\
\Delta w &= y + \Delta y - w, \\
\Delta s &= -W^{-1}((w + \Delta w) \cdot s + \mu^\theta(w + \Delta w - w^e)), \\
\Delta z_1 &= -(X^*_1)^{-1}(z_1 \cdot (E_F(x + \Delta x) - \ell + \mu^\theta e) - \mu^\theta z_1^e), \\
\Delta z_2 &= -(X^*_2)^{-1}(z_2 \cdot (u - E_F(x + \Delta x) + \mu^\theta e) - \mu^\theta z_2^e),
\end{align*}
\]
where \( \Delta x_F \) and \( \Delta y \) satisfy the equations

\[
\left( \begin{array}{cc}
H_F + D_z^{-1} & -J_F^T \\
J_F & D_V + D_W
\end{array} \right) \left( \begin{array}{c}
\Delta x_F \\
\Delta y
\end{array} \right) = - \left( \begin{array}{c}
g_F - J_F^T y - \pi^z \\
D_V (y - \pi^y) + D_w (y - \pi^w)
\end{array} \right).
\]

### 7.6. Summary: bounded variables

Define the quantities

\[
D_V = \mu^v I, \\
\pi^v = y^v - \frac{1}{\mu^v} (c(x) - s),
\]

\[
D_1^z = X_1^v Z_1^{-1}, \\
\pi_1^z = \mu^z (X_1^v)^{-1} z_1^v,
\]

\[
D_2^z = X_2^v Z_2^{-1}, \\
\pi_2^z = \mu^z (X_2^v)^{-1} z_2^v,
\]

\[
D_Z = ((D_1^z)^{-1} + (D_2^z)^{-1})^{-1}, \\
\pi^z = \pi_1^z - \pi_2^z,
\]

\[
D_W = S^\mu W^{-1}, \\
\pi^w = \mu^w (S^\mu)^{-1} w^e,
\]

then \( \Delta s, \Delta w, \Delta x_1, \Delta x_2, \Delta z_1 \) and \( \Delta z_2 \) are given by

\[
\Delta x = E_F^T \Delta x_F, \\
\tilde{y} = y + \Delta y, \\
\Delta w = \tilde{y} - w, \\
\tilde{w} = w + \Delta w, \\
\Delta s = -W^{-1} (\tilde{w} \cdot s + \mu^w (\tilde{w} - w^e)), \\
\tilde{x} = x + \Delta x, \\
\Delta z_1 = -(X_1^v)^{-1} (z_1^v \cdot (\tilde{x} - e + \mu^e) - \mu^e z_1^e), \\
\Delta z_2 = -(X_2^v)^{-1} (z_2^v \cdot (u - \tilde{x} + \mu^e) - \mu^e z_2^e),
\]

where \( \Delta x \) and \( \Delta y \) satisfy the equations

\[
\left( \begin{array}{cc}
H_F + D_z^{-1} & -J_F^T \\
J_F & D_V + D_W
\end{array} \right) \left( \begin{array}{c}
\Delta x_F \\
\Delta y
\end{array} \right) = - \left( \begin{array}{c}
g_F - J_F^T y - \pi^z \\
D_V (y - \pi^y) + D_w (y - \pi^w)
\end{array} \right).
\]
The line-search merit function is

\[
    f - (c - s)^T y^\varepsilon + \frac{1}{2\mu^p}||c - s||^2 + \frac{1}{2\mu^p}||c - s + \mu^p(y - y^\varepsilon)||^2
\]

\[
    - \sum_{j \in \mathcal{F}} \{ \mu^u[z^E_1]_j \ln (x_j - \ell_j + \mu^u) + \mu^u[z^E_1]_j \ln (\{z_1\}_j(x_j - \ell_j + \mu^u)) - \{z_1\}_j(x_j - \ell_j + \mu^u) \}
\]

\[
    - \sum_{j \in \mathcal{F}} \{ \mu^u[z^E_2]_j \ln (u_j - x_j + \mu^u) + \mu^u[z^E_2]_j \ln (\{z_2\}_j(u_j - x_j + \mu^u)) - \{z_2\}_j(u_j - x_j + \mu^u) \}
\]

\[
    - \sum_{i=1}^{m} \{ \mu^u w^E_i \ln (s_i + \mu^u) + \mu^u w^E_i \ln (w_i(s_i + \mu^u)) - w_i(s_i + \mu^u) \}. \quad (7.18)
\]
8. Upper and Lower Bounds on all Variables and Slacks

Next we consider the case with upper and lower bounds on all the variables and slacks.

8.1. Problem statement and optimality conditions

The definition of the problem is

$$\begin{align*}
\text{minimize} \quad & f(x) \\
\text{subject to} \quad & c(x) - s = 0, \quad \ell^x \leq x \leq u^x, \quad \ell^s \leq s \leq u^s,
\end{align*}$$

where $c : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^n \to \mathbb{R}$ are twice-continuously differentiable. The first-order KKT conditions for this problem are

$$\begin{align*}
g(x^*) - J(x^*)^T y^* - z_1^* + z_2^* &= 0, & z_1^* &\geq 0, & z_2^* &\geq 0, \\
y^* - w_1^* + w_2^* &= 0, & w_1^* &\geq 0, & w_2^* &\geq 0, \\
c(x^*) - s^* &= 0, & x^* - \ell^x &\geq 0, & u^x - x^* &\geq 0, \\
\ell^s - s^* &\geq 0, & u^s - s^* &\geq 0, \\
z_1^* \cdot (x^* - \ell^x) &= 0, & z_2^* \cdot (u^x - x^*) &= 0, \\
w_1^* \cdot (s^* - \ell^s) &= 0, & w_2^* \cdot (u^s - s^*) &= 0,
\end{align*}$$

where $y^*$ are the multipliers for the equality constraints $c(x) - s = 0$, and $z_1^*$, $z_2^*$, $w_1^*$, and $w_2^*$ may be interpreted as the Lagrange multipliers for the inequality constraints $x - \ell^x \geq 0$, $u^x - x \geq 0$, $s - \ell^s \geq 0$ and $u^s - s \geq 0$, respectively.

8.2. The path-following equations

Let $z_1^\ell$, $z_2^\ell$, $w_1^\ell$, and $w_2^\ell$ denote nonnegative estimates of the Lagrange multipliers for the inequality constraints $x_1 \geq 0$, $x_2 \geq 0$, $s_1 \geq 0$ and $s_2 \geq 0$, respectively. Given small positive scalars $\mu^\ell$ and $\mu^s$, consider the perturbed optimality conditions

$$\begin{align*}
g(x) - J(x)^T y - z_1 + z_2 &= 0, & z_1 &\geq 0, & z_2 &\geq 0, \\
y - w_1 + w_2 &= 0, & w_1 &\geq 0, & w_2 &\geq 0, \\
c(x) - s = \mu^\ell (y^\ell - y), & x - \ell^x &\geq 0, & u^x - x &\geq 0, \\
\ell^s - s &\geq 0, & u^s - s &\geq 0, \\
z_1 \cdot (x - \ell^x) = \mu^\ell (z_1^\ell - z_1), & z_2 \cdot (u^x - x) = \mu^\ell (z_2^\ell - z_2), \\
w_1 \cdot (s - \ell^s) = \mu^\ell (w_1^\ell - w_1), & w_2 \cdot (u^s - s) = \mu^\ell (w_2^\ell - w_2),
\end{align*}$$

8. Upper and Lower Bounds on all Variables and Slacks 72
Consider the primal-dual path-following equations
\[ F(x, s, y, z_1, z_2, w_1, w_2 : \mu^p, \mu^a, y^e, z_1^p, z_2^p, w_1^p, w_2^p) = 0, \]
with
\[
F(x, s, y, z_1, z_2, w_1, w_2 : \mu^p, \mu^a, y^e, z_1^p, z_2^p, w_1^p, w_2^p) = \begin{pmatrix}
g(x) - J(x)^T y - z_1 + z_2 \\
y - w_1 + w_2 \\
c(x) - s + \mu^a(y - y^e) \\
z_1 \cdot (x - \ell^x) + \mu^a(z_1 - z_1^p) \\
z_2 \cdot (x - \ell^x) + \mu^a(z_2 - z_2^p) \\
w_1 \cdot (s - \ell^s) + \mu^a(w_1 - w_1^p) \\
w_2 \cdot (s - \ell^s) + \mu^a(w_2 - w_2^p)
\end{pmatrix}.
\] (8.4)

Any zero \((x, s, y, z_1, z_2, w_1, w_2)\) of \(F\) that satisfies \(\ell^x < x < u^x, \ell^s < s < u^s, z_1 > 0, z_2 > 0, w_1 > 0,\) and \(w_2 > 0\) approximates a point satisfying the optimality conditions (8.2), with the approximation becoming increasingly accurate as the terms \(\mu^p(y - y^e), \mu^a(z_1 - z_1^p), \mu^a(z_2 - z_2^p), \mu^a(w_1 - w_1^p)\) and \(\mu^a(w_2 - w_2^p)\) approach zero. For any sequence of \(z_1^p, z_2^p, w_1^p, w_2^p\) and \(y^e\) such that \(z_1^p \to z_1^*, z_2^p \to z_2^*, w_1^p \to w_1^*, w_2^p \to w_2^*,\) and \(y^e \to y^*,\) and it must hold that solutions \((x, s, y, z_1, z_2, w_1, w_2)\) of (8.3) must satisfy \(z_1 \cdot (x - \ell^x) \to 0, z_2 \cdot (u^x - x) \to 0, w_1 \cdot (s - \ell^s) \to 0,\) and \(w_2 \cdot (u^s - s) \to 0.\) This implies that any solution \((x, s, y, z_1, z_2, w_1, w_2)\) of (8.3) will approximate a solution of (8.2) independently of the values of \(\mu^p\) and \(\mu^a\) (i.e., it is not necessary that \(\mu^p \to 0\) and \(\mu^a \to 0).\)

If \(v = (x, s, y, z_1, z_2, w_1, w_2)\) is a given approximate zero of \(F\) such that \(\ell^x - \mu^a < x < u^x + \mu^a, \ell^s - \mu^a < s < u^s + \mu^a, z_1 > 0, z_2 > 0, w_1 > 0,\) and \(w_2 > 0,\) the Newton equations for the change in variables \((\Delta x, \Delta s, \Delta y, \Delta z_1, \Delta z_2, \Delta w_1, \Delta w_2)\) are given by

\[
\begin{pmatrix}
H \\ 0 \\ 0 \\ Z_1 \\ -Z_2 \\ 0 \\ 0 \\ 0
\end{pmatrix}
\begin{pmatrix}
I \\ 0 \\ 0 \\ X_1^a \\ 0 \\ 0 \\ W_1 \\ W_2
\end{pmatrix}
\begin{pmatrix}
\Delta x \\ \Delta z_1 \\ \Delta z_2 \\ \Delta w_1 \\ \Delta w_2
\end{pmatrix}
= \begin{pmatrix}
g - J^T y - z_1 + z_2 \\ y - w_1 + w_2 \\ c - s + \mu^a(y - y^e) \\ z_1 \cdot (x - \ell^x) + \mu^a(z_1 - z_1^p) \\ z_2 \cdot (x - \ell^x) + \mu^a(z_2 - z_2^p) \\ w_1 \cdot (s - \ell^s) + \mu^a(w_1 - w_1^p) \\ w_2 \cdot (s - \ell^s) + \mu^a(w_2 - w_2^p)
\end{pmatrix},
\] (8.5)

where \(D_v = \mu^p I, X_1^a = \text{diag}(x_j - \ell^x + \mu^a), X_2^a = \text{diag}(u^x - x_j + \mu^a), Z_1 = \text{diag}([z_1]_j), Z_2 = \text{diag}([z_2]_j), W_1 = \text{diag}([w_1]_j), W_2 = \text{diag}([w_2]_j), S_1^a = \text{diag}(s_i - \ell^s + \mu^a),\) and \(S_2^a = \text{diag}(u^s - s_i + \mu^a).\)
8.3. A shifted primal-dual penalty-barrier function

Problem (8.1) is equivalent to

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad c(x) - s = 0, \\
& \quad x - x_1 = \ell^x, \quad s - s_1 = \ell^s, \quad x_1 \geq 0, \quad s_1 \geq 0, \\
& \quad x + x_2 = u^x, \quad s + s_2 = u^s, \quad x_2 \geq 0, \quad s_2 \geq 0.
\end{align*}
\]

Consider the shifted primal-dual penalty-barrier problem

\[
\begin{align*}
\text{minimize} & \quad M(x, x_1, x_2, s, s_1, s_2, y, w_1, w_2; \mu^e, \mu^s, y^e, w_1^e, w_2^e) \\
\text{subject to} & \quad x - x_1 = \ell^x, \quad s - s_1 = \ell^s, \quad x_1 + \mu^e > 0, \quad z_1 > 0, \quad s_1 + \mu^e > 0, \quad w_1 > 0, \\
& \quad x + x_2 = u^x, \quad s + s_2 = u^s, \quad x_2 + \mu^e > 0, \quad z_2 > 0, \quad s_2 + \mu^e > 0, \quad w_2 > 0,
\end{align*}
\]

where \( M(x, x_1, x_2, s, s_1, s_2, y, z_1, z_2, w_1, w_2; \mu^e, \mu^s, y^e, z_1^e, z_2^e, w_1^e, w_2^e) \) is the shifted primal-dual penalty-barrier function

\[
\begin{align*}
f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^e} \|c(x) - s\|^2 + \frac{1}{2\mu^e} \|c(x) - s + \mu^e(y - y^e)\|^2 \\
- \sum_{j=1}^n \{\mu^e[z_1^e] \ln ([x_1 + \mu^e e]\_j) + \mu^e[z_1^e] \ln ([z_1 \cdot (x_1 + \mu^e e)]_j) - [z_1 \cdot (x_1 + \mu^e e)]_j\} \\
- \sum_{j=1}^n \{\mu^e[z_2^e] \ln ([x_2 + \mu^e e]\_j) + \mu^e[z_2^e] \ln ([z_2 \cdot (x_2 + \mu^e e)]_j) - [z_2 \cdot (x_2 + \mu^e e)]_j\} \\
- \sum_{i=1}^m \{\mu^e[w_1^e] \ln ([s_1 + \mu^e e]\_i) + \mu^e[w_1^e] \ln ([w_1 \cdot (s_1 + \mu^e e)]_i) - [w_1 \cdot (s_1 + \mu^e e)]_i\} \\
- \sum_{i=1}^m \{\mu^e[w_2^e] \ln ([s_2 + \mu^e e]\_i) + \mu^e[w_2^e] \ln ([w_2 \cdot (s_2 + \mu^e e)]_i) - [w_2 \cdot (s_2 + \mu^e e)]_i\}.
\end{align*}
\]
Let $c$, $g$ and $J$ denote the quantities $c(x)$, $g(x)$ and $J(x)$. The gradient of the merit function as a function of $x$, $x_1$, $x_2$, $s$, $s_1$, $s_2$, $y$, $z_1$, $z_2$, $w_1$, and $w_2$, is
\[
\nabla M = \begin{pmatrix}
    g - J^T(2(y^e - \frac{1}{\mu^p}(c - s)) - y) \\
    z_1 - 2\mu^a(X^\mu_1)^{-1}z_1^e \\
    z_2 - 2\mu^a(X^\mu_2)^{-1}z_2^e \\
    2(y^e - \frac{1}{\mu^p}(c - s)) - y \\
    w_1 - 2\mu^a(S^\mu_1)^{-1}w_1^e \\
    w_2 - 2\mu^a(S^\mu_2)^{-1}w_2^e \\
    c - s + \mu^e(y - y^e) \\
    x_1 + \mu^e c - \mu^a Z_1^{-1}z_1^e \\
    x_2 + \mu^e c - \mu^a Z_2^{-1}z_2^e \\
    s_1 + \mu^e c - \mu^a W_1^{-1}w_1^e \\
    s_2 + \mu^e c - \mu^a W_2^{-1}w_2^e
\end{pmatrix}.
\]

Similarly, the Hessian of $M(x, x_1, x_2, s, s_1, s_2, y, w_1, w_2)$ is given by
\[
\begin{pmatrix}
    H_1 & 0 & 0 & -\frac{2}{\mu^p}J^T & 0 & 0 & J^T & 0 & 0 & 0 & 0 \\
    0 & 2\mu^a(X^\mu_1)^{-2}Z_1^e & 0 & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 \\
    0 & 0 & 2\mu^a(X^\mu_2)^{-2}Z_2^e & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 \\
    -\frac{2}{\mu^p}J & 0 & 0 & 2\mu^a S^\mu_1^{-1}W_1^e & 0 & 0 & 0 & I_m & 0 & 0 & 0 \\
    0 & 0 & 0 & 0 & 2\mu^a S^\mu_2^{-1}W_2^e & 0 & 0 & 0 & I_m & 0 & 0 \\
    J & 0 & 0 & -I_m & 0 & 0 & 0 & 0 & I_m & 0 & 0 \\
    0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
    0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & \mu^a Z_1^{-2}Z_1^e & 0 \\
    0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & \mu^a Z_2^{-2}Z_2^e & 0 \\
    0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & \mu^a W_1^{-2}W_1^e \\
    0 & 0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & \mu^a W_2^{-2}W_2^e
\end{pmatrix}.
\]
where \( H_1 = H(x, 2\pi y - y) + \frac{2}{\mu} J^T J \). Substituting \( \mu^w Z^\pi_1 = X^\mu_1 \Pi^\pi_1 \), \( \mu^w Z^\pi_2 = X^\mu_2 \Pi^\pi_2 \), \( \mu^w W^\pi_1 = S^\mu \Pi^w_1 \) and \( \mu^w W^\pi_2 = S^\mu \Pi^w_2 \) gives

\[
\begin{pmatrix}
H_1 & 0 & 0 & -\frac{2}{\mu^w} J^T & 0 & 0 & J^T & 0 & 0 & 0 & 0 \\
0 & 2(X_1^\mu)^{-1} \Pi^\pi_1 & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 & 0 \\
0 & 0 & 2(X_2^\mu)^{-1} \Pi^\pi_2 & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 \\
-\frac{2}{\mu^w} J & 0 & 0 & 0 & -I_m & 0 & I_m & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 2(S_1^w)^{-1} \Pi^\pi_1 & 0 & 0 & 0 & I_m & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 \\
0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X_1^\mu Z_1^{-2} \Pi^\pi_1 & 0 & 0 \\
0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & X_2^\mu Z_2^{-2} \Pi^\pi_2 & 0 & 0 \\
0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & S_1^w W_1^{-2} \Pi^w_1 & 0 & 0 \\
0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & S_2^w W_2^{-2} \Pi^w_2 \\
\end{pmatrix}
\]

### 8.4. Derivation of the shifted primal-dual penalty-barrier direction

The primal-dual penalty-barrier problem may be written in the form

\[
\min_{p \in \mathcal{I}} M(p) \quad \text{subject to} \quad C p = 0,
\]

(8.8)

where \( \mathcal{I} \) is the set of vectors \( p = (x, x_1, x_2, s_1, s_2, y, z_1, z_2, w_1, w_2) \) such that \( x_1 + \mu^e e > 0, x_2 + \mu^w e > 0, s_1 + \mu^s e > 0, s_2 + \mu^w e > 0, z_1 > 0, z_2 > 0 > 0, w_1 > 0, \) and \( w_2 > 0, \) and

\[
C = \begin{pmatrix}
I_n & -I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
I_n & -I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_m & -I_m & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_m & -I_m & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}.
\]

Given \( p \in \mathcal{I} \), the Newton direction \( \Delta p \) is the solution of the subproblem

\[
\min_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T \nabla^2 M(p) \Delta p \quad \text{subject to} \quad C \Delta p = -C p.
\]

(8.9)

However, instead of solving (8.9), we define a linearly constrained modified Newton method by approximating the Hessian \( \nabla^2 M(x, x_1, x_2, s_1, s_2, y, z_1, z_2, w_1, w_2) \) by a matrix \( B(x, x_1, x_2, s_1, s_2, y, z_1, z_2, w_1, w_2) \). Consider the matrix defined by replacing \( \pi^y \) by \( y, \pi^1_1 \) by \( z_1, \pi^2_2 \) by \( z_2, \pi^1_2 \) by \( w_1 \) and \( \pi^2_2 \) by \( w_2 \) everywhere in the matrix \( \nabla^2 M(x, x_1, x_2, s_1, s_2, y, z_1, z_2, w_1, w_2) \). This
gives an approximate Hessian $B(x, s, s_p, y, w_p)$ of the form

$$
\begin{pmatrix}
H_1 & 0 & 0 & -\frac{2}{\mu^2}J^T & 0 & 0 & J^T & 0 & 0 & 0 & 0 \\
0 & 2(X_1^0)^{-1}Z_1 & 0 & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 \\
0 & 0 & 2(X_2^0)^{-1}Z_2 & 0 & 0 & 0 & 0 & 0 & I_n & 0 & 0 \\
-\frac{2}{\mu^2}J & 0 & 0 & 0 & 0 & 0 & 0 & -I_m & 0 & I_n & 0 \\
0 & 0 & 0 & 0 & 0 & 2(S_1^0)^{-1}W_1 & 0 & 0 & 0 & 0 & I_m \\
0 & 0 & 0 & 0 & 0 & 0 & 2(S_2^0)^{-1}W_2 & 0 & 0 & 0 & I_m \\
0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & 0 & S_1^0W_1^{-1} \\
0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & S_2^0W_2^{-1}
\end{pmatrix},
$$

where

$$
D_1^x = X_1^0Z_1^{-1}, \quad \pi_1^x = \mu^a(X_1^0)^{-1}z_1^0, \\
D_2^x = X_2^0Z_2^{-1}, \quad \pi_2^x = \mu^a(X_2^0)^{-1}z_2^0, \\
D_1^w = X_1^0W_1^{-1}, \quad \pi_1^w = \mu^a(S_1^0)^{-1}w_1^0, \\
D_2^w = X_2^0W_2^{-1}, \quad \pi_2^w = \mu^a(S_2^0)^{-1}w_2^0.
$$

These definitions of $D_1^x$, $D_2^x$, $D_1^w$ and $D_2^w$ can be used to write $B(x, x_1, x_2, s, s_1, s_2, y, z_1, z_2, w_1, w_2)$ in the form

$$
\begin{pmatrix}
H + 2J^TD_1^{-1}J & 0 & 0 & -2J^TD_1^{-1} & 0 & 0 & 0 & J^T & 0 & 0 & 0 & 0 \\
0 & 2(D_1^0)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 \\
0 & 0 & 2(D_2^0)^{-1} & 0 & 0 & 0 & 0 & 0 & 0 & I_n & 0 & 0 \\
-2D_1^{-1}J & 0 & 0 & 0 & 0 & 2(D_1^0)^{-1} & 0 & 0 & 0 & 0 & I_m & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2(D_2^0)^{-1} & 0 & 0 & 0 & I_m & 0 \\
0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & D_1^x & 0 & 0 \\
0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & D_2^x & 0 & 0 \\
0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & 0 & D_1^w & 0 \\
0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & D_2^w & 0
\end{pmatrix},
$$

for $8. Upper and Lower Bounds on all Variables and Slacks 77$.
where $H = H(x,y)$. Given $B(p) = B(x,x_1,x_2,s,s_1,s_2,y,z_1,z_2,w_1,w_2)$, a modified Newton direction is given by the solution of the QP subproblem

$$
\begin{equation}
\minimize_{\Delta p} \nabla M(p)^T \Delta p + \frac{1}{2} \Delta p^T B(p) \Delta p \quad \text{subject to} \quad C \Delta p = -Cp.
\end{equation}
$$

(8.11)

If $p = (x,x_1,x_2,s,s_1,s_2,y,z_1,z_2,w_1,w_2)$ is feasible for the constraints then $Cp = 0$. In this case every feasible $\Delta p$ may be written in the form $\Delta p = Nd$, where $N$ is a matrix whose columns form a basis for null($C$), i.e., $CN = 0$ and $(C^T \quad N)$ is nonsingular. This implies that $d$ must satisfy the reduced equations

$$
N^T B(p) N d = -N^T \nabla M(p).
$$

Consider the particular null-space basis given by

$$
N = \begin{pmatrix}
I_n & 0 & 0 & 0 & 0 & 0 & 0 \\
I_n & 0 & 0 & 0 & 0 & 0 & 0 \\
-I_n & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_m & 0 & 0 & 0 & 0 & 0 \\
0 & I_m & 0 & 0 & 0 & 0 & 0 \\
0 & -I_m & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_m & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_n & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_n & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I_m & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_m
\end{pmatrix}.
$$

(8.12)
With this definition, the reduced modified Newton equations \( N^T B(p) N d = -N^T \nabla M(p) \) for the linearly constrained problem (8.6) are

\[
\begin{pmatrix}
H + 2J^T D^{-1}_Y J + 2D^{-1}_Z & -2J^T D^{-1}_Y & J^T & I_n & -I_n & 0 & 0 \\
-2D^{-1}_Y J & 2(D^{-1}_Y + D^{-1}_W) & -I_m & 0 & 0 & I_m & -I_m \\
J & I_n & D_Y & 0 & 0 & 0 & 0 \\
I_n & 0 & 0 & D^2_Z & 0 & 0 & 0 \\
-I_n & 0 & 0 & 0 & D^2_W & 0 & 0 \\
0 & I_m & 0 & 0 & 0 & D^w_Y & 0 \\
0 & -I_m & 0 & 0 & 0 & 0 & D^w_Z
\end{pmatrix}
\begin{pmatrix}
d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7
\end{pmatrix}
= - \begin{pmatrix}
g - J^T(2\pi^Y - y) - (2\pi^Y_1 - z_1) + (2\pi^Y_2 - z_2) \\
(2\pi^Y - y) - (2\pi^W_1 - w_1) + (2\pi^W_2 - w_2) \\
D_Y(y - \pi^Y) \\
D^2_Y(z_1 - \pi^Y_1) \\
D^2_Z(z_2 - \pi^Y_2) \\
D^w_Y(w_1 - \pi^W_1) \\
D^w_Z(w_2 - \pi^W_2)
\end{pmatrix},
\]

where \( D_w = ((D^1_w)^{-1} + (D^2_w)^{-1})^{-1} \), and \( D_z = ((D^2_z)^{-1} + (D^2_Z)^{-1})^{-1} \). Given any nonsingular matrix \( R \), the direction \( d \) satisfies

\[
RN^T B(p) Nd = -RN^T \nabla M(p).
\]

In particular, if \( R \) is the block upper-triangular matrix

\[
R = \begin{pmatrix}
I & 0 & -2J^T D^{-1}_Y & -2(D^2_z)^{-1} & 2(D^2_Z)^{-1} & 0 & 0 \\
& I & 2D^{-1}_Y & 0 & 0 & -2(D^2_Y)^{-1} & 2(D^w_Z)^{-1} \\
& & I & 0 & 0 & 0 & 0 \\
& & & Z_1 & 0 & 0 & 0 \\
& & & & Z_2 & 0 & 0 \\
& & & & & W_1 & 0 \\
& & & & & & W_2
\end{pmatrix},
\]
then

\[
RN^T B(p) N = \begin{pmatrix} H & 0 & J^T & -I_n & I_n & 0 & 0 \\ 0 & 0 & -I_m & 0 & 0 & -I_m & I_m \\ J & -I_n & D_y & 0 & 0 & 0 & 0 \\ Z_1 & 0 & 0 & Z_1 D_1^w & 0 & 0 & 0 \\ -Z_2 & 0 & 0 & Z_2 D_2^w & 0 & 0 & 0 \\ 0 & W_1 & 0 & 0 & 0 & W_1 D_1^w & 0 \\ 0 & -W_2 & 0 & 0 & 0 & W_2 D_2^w & 0 \end{pmatrix} = \begin{pmatrix} H & 0 & J^T & -I_n & I_n & 0 & 0 \\ 0 & 0 & -I_m & 0 & 0 & -I_m & I_m \\ J & -I_m & D_y & 0 & 0 & 0 & 0 \\ Z_1 & 0 & 0 & X_1^w & 0 & 0 & 0 \\ -Z_2 & 0 & 0 & X_2^w & 0 & 0 & 0 \\ 0 & W_1 & 0 & 0 & 0 & S_1^w & 0 \\ 0 & -W_2 & 0 & 0 & 0 & S_2^w & 0 \end{pmatrix} \]

and

\[
RN^T \nabla M(p) = - \begin{pmatrix} g - J^T y - z_1 + z_2 \\ y - w_1 + w_2 \\ D_y (y - \pi^w) \\ X_1^w (z_1 - \pi_1^w) \\ X_2^w (z_2 - \pi_2^w) \\ S_1^w (w_1 - \pi_1^w) \\ S_2^w (w_2 - \pi_2^w) \end{pmatrix} .
\]

Giving the transformed modified Newton system \(RN^T B(p) Nd = -RN^T \nabla M(p)\) as

\[
\begin{pmatrix} H & 0 & J^T & -I_n & I_n & 0 & 0 \\ 0 & 0 & -I_m & 0 & 0 & -I_m & I_m \\ J & -I_m & D_y & 0 & 0 & 0 & 0 \\ Z_1 & 0 & 0 & X_1^w & 0 & 0 & 0 \\ -Z_2 & 0 & 0 & X_2^w & 0 & 0 & 0 \\ 0 & W_1 & 0 & 0 & 0 & S_1^w & 0 \\ 0 & -W_2 & 0 & 0 & 0 & S_2^w & 0 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ d_6 \\ d_7 \end{pmatrix} = \begin{pmatrix} g - J^T y - z_1 + z_2 \\ y - w_1 + w_2 \\ D_y (y - \pi^w) \\ X_1^w (z_1 - \pi_1^w) \\ X_2^w (z_2 - \pi_2^w) \\ S_1^w (w_1 - \pi_1^w) \\ S_2^w (w_2 - \pi_2^w) \end{pmatrix} .
\]

Identities of the form

\[
X_1^w (z_1 - \pi_1^w) = X_1^w (z_1 - \mu^w (X_1^w)^{-1} z_1^i) = Z_1 (x_1 + \mu^w e) - \mu^w z_1^i = z_1 \cdot (x - \ell) + \mu^w (z_1 - z_1^i)
\]
for each of the terms $X^u_1(z_1 - \pi^u_1)$, $X^u_2(z_2 - \pi^u_2)$, $S^u_1(w_1 - \pi^u_1)$, $S^u_2(w_2 - \pi^u_2)$ give

\[

\begin{pmatrix}
H & 0 & J^T & -I_n & I_n & 0 & 0 \\
0 & 0 & -I_m & 0 & 0 & -I_m & I_m \\
J & -I_m & D_y & 0 & 0 & 0 & 0 \\
Z_1 & 0 & 0 & X^u_1 & 0 & 0 & 0 \\
-Z_2 & 0 & 0 & 0 & X^u_2 & 0 & 0 \\
0 & W_1 & 0 & 0 & 0 & S^u_1 & 0 \\
0 & -W_2 & 0 & 0 & 0 & 0 & S^u_2 \\
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5 \\
d_6 \\
d_7 \\
\end{pmatrix}
= 
\begin{pmatrix}
g - J^T y - z_1 + z_2 \\
y - w_1 + w_2 \\
D_y (y - \pi^y) \\
z_1 \cdot (x - \ell^e) + \mu^v(z_1 - z^e_1) \\
z_2 \cdot (u^x - x) + \mu^v(z_1 - z^e_1) \\
w_1 \cdot (s - \ell^e) + \mu^v(w_1 - w^e_1) \\
w_2 \cdot (u^s - s) + \mu^v(w_2 - w^e_2) \\
\end{pmatrix}
\tag{8.13}
\]

Then, the definition of $N$ from (8.12) implies that

\[

\begin{pmatrix}
\Delta x \\
\Delta x_1 \\
\Delta x_2 \\
\Delta s \\
\Delta s_1 \\
\Delta s_2 \\
\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\end{pmatrix}
= \Delta p = Nd = 
\begin{pmatrix}
I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_m & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & I_m & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -I_m & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & I_n & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
d_1 \\
d_2 \\
d_3 \\
d_4 \\
d_5 \\
d_6 \\
d_7 \\
\end{pmatrix}
\tag{8.14}
\]

Using these identities to substitute for the components of $d$ in (8.13) yields

\[

\begin{pmatrix}
H & 0 & J^T & -I_n & I_n & 0 & 0 \\
0 & 0 & -I_m & 0 & 0 & -I_m & I_m \\
J & -I_m & D_y & 0 & 0 & 0 & 0 \\
Z_1 & 0 & 0 & X^u_1 & 0 & 0 & 0 \\
-Z_2 & 0 & 0 & 0 & X^u_2 & 0 & 0 \\
0 & W_1 & 0 & 0 & 0 & S^u_1 & 0 \\
0 & -W_2 & 0 & 0 & 0 & 0 & S^u_2 \\
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\end{pmatrix}
= 
\begin{pmatrix}
g - J^T y - z_1 + z_2 \\
y - w_1 + w_2 \\
D_y (y - \pi^y) \\
z_1 \cdot (x - \ell^e) + \mu^v(z_1 - z^e_1) \\
z_2 \cdot (u^x - x) + \mu^v(z_1 - z^e_1) \\
w_1 \cdot (s - \ell^e) + \mu^v(w_1 - w^e_1) \\
w_2 \cdot (u^s - s) + \mu^v(w_2 - w^e_2) \\
\end{pmatrix}
\tag{8.14}
\]

This system is identical to the Newton equations (8.5) for a solution of the path-following equations (8.3).
8.5. Computation of the shifted primal-dual penalty-barrier direction

The symmetric path-following equations are

\[
\begin{pmatrix}
H & 0 & J^T & I & -I & 0 & 0 \\
0 & 0 & -I & 0 & 0 & I & -I \\
J & -I & -D_y & 0 & 0 & 0 & 0 \\
I & 0 & 0 & -Z_1^{-1}X_1^\mu & 0 & 0 & 0 \\
-I & 0 & 0 & 0 & -Z_2^{-1}X_2^\mu & 0 & 0 \\
0 & I & 0 & 0 & 0 & -W_1^{-1}S_1^w & 0 \\
0 & -I & 0 & 0 & 0 & 0 & -W_2^{-1}S_2^w \\
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta s \\
-\Delta y \\
\Delta z_1 \\
\Delta z_2 \\
-\Delta w_1 \\
-\Delta w_2 \\
\end{pmatrix}
= -
\begin{pmatrix}
g - J^Ty - z_1 + z_2 \\
y - w_1 + w_2 \\
c - s + \mu^\nu(y - y^\nu) \\
Z_1^{-1}(z_1 \cdot (x - \ell^\nu) + \mu^\nu(z_1 - z_1^\nu)) \\
Z_2^{-1}(z_2 \cdot (u^\nu - x) + \mu^\nu(z_2 - z_2^\nu)) \\
W_1^{-1}(w_1 \cdot (s - \ell^\nu) + \mu^\nu(w_1 - w_1^\nu)) \\
W_2^{-1}(w_2 \cdot (u^\nu - s) + \mu^\nu(w_2 - w_2^\nu)) \\
\end{pmatrix}
\]  

(8.15)

After collecting terms and reordering the equations and unknowns, we obtain

\[
\begin{pmatrix}
D_1^z & 0 & 0 & 0 & 0 & I_n & 0 \\
0 & D_2^z & 0 & 0 & 0 & -I_n & 0 \\
0 & 0 & D_3^y & 0 & I_m & 0 & 0 \\
0 & 0 & 0 & D_4^y & -I_m & 0 & 0 \\
0 & 0 & -I_m & I_m & 0 & 0 & I_m \\
0 & 0 & 0 & 0 & -I_m & J & D_y \\
-I_n & I_n & 0 & 0 & 0 & 0 & H - J^T \\
\end{pmatrix}
\begin{pmatrix}
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s \\
\Delta x \\
\Delta y \\
\end{pmatrix}
= -
\begin{pmatrix}
D_1^z(z_1 - \pi_1^z) \\
D_2^z(z_2 - \pi_2^z) \\
D_3^y(w_1 - \pi_1^y) \\
D_4^y(w_2 - \pi_2^y) \\
D_5^z(w_1 - \pi_1^z) \\
D_6^z(w_2 - \pi_2^z) \\
y - w_1 + w_2 \\
D_y(y - \pi^\nu) \\
\end{pmatrix}
\]  

(8.16)

where

\[
D_y = \mu^\nu I_m, \quad \pi^\nu = y^\nu - \frac{1}{\mu^\nu}(c - s),
\]

\[
D_1^z = S_1^wW_1^{-1}, \quad \pi_1^z = \mu^\nu(S_1^w)^{-1}w_1^\nu, \quad D_1^z = X_1^\mu Z_1^{-1}, \quad \pi_1^z = \mu^\nu(X_1^\mu)^{-1}z_1^\nu,
\]

\[
D_2^w = S_2^wW_2^{-1}, \quad \pi_2^w = \mu^\nu(S_2^w)^{-1}w_2^\nu, \quad D_2^w = X_2^\mu Z_2^{-1}, \quad \pi_2^w = \mu^\nu(X_2^\mu)^{-1}z_2^\nu.
\]

If we define \( \bar{H} = H + (D_1^z)^{-1} + (D_2^w)^{-1} \), \( \bar{D}_y = D_y + D_w \) and \( D_w = ((D_1^w)^{-1} + (D_2^w)^{-1})^{-1} \), then premultiplying the equations (8.16) by the matrix

\[
\begin{pmatrix}
I_n & 0 & I_n & 0 & I_m & 0 & 0 & 0 \\
0 & 0 & 0 & I_m & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & (D_1^w)^{-1} & -(D_2^w)^{-1} & I_m & 0 \\
0 & 0 & 0 & D_w(D_1^w)^{-1} & -D_w(D_2^w)^{-1} & I_m & 0 & 0 \\
(D_1^z)^{-1} & -(D_2^z)^{-1} & 0 & 0 & 0 & 0 & 0 & I_m \\
\end{pmatrix}
\]
gives the block upper-triangular system

\[
\begin{pmatrix}
D_1^z & 0 & 0 & 0 & 0 & I_n & 0 \\
0 & D_2^z & 0 & 0 & 0 & -I_n & 0 \\
0 & 0 & D_1^w & 0 & I_m & 0 & 0 \\
0 & 0 & 0 & D_2^w & -I_m & 0 & 0 \\
0 & 0 & 0 & 0 & I_m & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & J & \bar{D}_y \\
0 & 0 & 0 & 0 & 0 & 0 & \bar{H}_F - J^T
\end{pmatrix}
\begin{pmatrix}
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s \\
\Delta x \\
\Delta y
\end{pmatrix}
= 
\begin{pmatrix}
D_1^z(z_1 - \pi_1^z) \\
D_2^z(z_2 - \pi_2^z) \\
D_1^w(w_1 - \pi_1^w) \\
D_2^w(w_2 - \pi_2^w) \\
D_w(y - \pi^w) \\
D_w(y - \pi^w) + D_v(y - \pi^v) \\
g - J^T y - \pi^z
\end{pmatrix},
\]

where \( \pi^w = \pi_1^w - \pi_2^w \) and \( \pi^z = \pi_1^z - \pi_2^z \). Using block back substitution, we may compute \( \Delta x \) and \( \Delta y \) by solving the equations

\[
\begin{pmatrix}
\bar{H}(x, y) + D_2^{-1} - J(x)^T \\
J(x) \\
D_v + D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix}
= 
\begin{pmatrix}
g(x) - J(x)^T y - \pi^z \\
D_w(y - \pi^w) + D_v(y - \pi^v)
\end{pmatrix},
\]

with \( D_x = ((D_1^z)^{-1} + (D_2^z)^{-1})^{-1} \) and \( D_w = ((D_1^w)^{-1} + (D_2^w)^{-1})^{-1} \). The fifth block of equations gives

\[\Delta s = -D_w(y + \Delta y - w_1 + w_2).\]

There are several ways of computing \( \Delta w_1 \) and \( \Delta w_2 \). Instead of using the block upper-triangular system above, we use the last two blocks of equations of (8.5) to give

\[\Delta w_1 = -(S_1^w)^{-1}(w_1 \cdot (s + \Delta s - \ell^s + \mu^w e) - \mu^s w_1^w) \quad \text{and} \quad \Delta w_2 = -(S_2^w)^{-1}(w_2 \cdot (u^s - (s + \Delta s) + \mu^e) - \mu^w w_2^w).\]

Similarly, using (8.5) to solve for \( \Delta z_1 \) and \( \Delta z_2 \) yields

\[\Delta z_1 = -(X_1^z)^{-1}(z_1 \cdot (x + \Delta x - \ell^x + \mu^e) - \mu^s z_1^s) \quad \text{and} \quad \Delta z_2 = -(X_2^z)^{-1}(z_2 \cdot (u^x - (x + \Delta x) + \mu^e) - \mu^w z_2^w).\]
The variables $x_1$, $x_2$, $s_1$ and $s_2$ may be computed implicitly for the line search, and the appropriate merit function is

$$f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^e} \|c(x) - s\|^2 + \frac{1}{2\mu^v} \|c(x) - s + \mu^v (y - y^e)\|^2$$

$$- \sum_{j=1}^{n} \{\mu^e [z_1^e]_j \ln (x_j - \ell_j^e + \mu^e) + \mu^a [z_1^a]_j \ln ([z_1]_j(x_j - \ell_j^a + \mu^a)) - [z_1]_j(x_j - \ell_j^e + \mu^e)\}$$

$$- \sum_{j=1}^{n} \{\mu^e [z_2^e]_j \ln (u_j^e - x_j + \mu^e) + \mu^a [z_2^a]_j \ln ([z_2]_j(u_j^e - x_j + \mu^a)) - [z_2]_j(u_j^e - x_j + \mu^e)\}$$

$$- \sum_{i=1}^{m} \{\mu^e [w_1^e]_i \ln (s_i - \ell_i^e + \mu^e) + \mu^a [w_1^a]_i \ln ([w_1]_i(s_i - \ell_i^a + \mu^a)) - [w_1]_i(s_i - \ell_i^e + \mu^e)\}$$

$$- \sum_{i=1}^{m} \{\mu^a [w_2^a]_i \ln (u_i^a - s_i + \mu^a) + \mu^a [w_2^e]_i \ln ([w_2]_i(u_i^a - s_i + \mu^a)) - [w_2]_i(u_i^a - s_i + \mu^a)\} \cdot (8.18)$$
8.6. Summary: upper and lower bounds on all variables and slacks

The results of Sections 6.5 and 7.5 imply that the solution of the path-following equations (8.5) may be computed as follows. Let \( x \) and \( s \) be given primal variables such that

\[
\ell^x - \mu^u < x < u^x + \mu^u, \quad \text{and} \quad \ell^s - \mu^u < s < u^s + \mu^u,
\]

and dual variables \( y, w_1, w_2, z_1, \) and \( z_2 \) such that

\[
w_1 > 0, \quad w_2 > 0, \quad z_1 > 0, \quad \text{and} \quad z_2 > 0.
\]

Let \( X_1, X_2, S_1, \) and \( S_2 \) denote the matrices \( \text{diag}(x_j - \ell^x_j) \), \( \text{diag}(u^x_j - x_j) \), \( \text{diag}(s_i - \ell^s_i) \) and \( \text{diag}(u^s_i - s_i) \), respectively. If \( D^x_1, D^x_2, D^y_1, D^y_2, D_D, D_S, D_w, \pi^x, \pi^x_1, \pi^x_2, \pi^y \) and \( \pi^y_2 \) denote the quantities

\[
D_D = \mu^I, \quad D_D^x = Z^{-1}S_1, \quad D_D^y = W^{-1}S_2, \\
D^x_1 = Z^{-1}X^x_1, \quad D^x_2 = Z^{-1}X^x_2, \\
D^y_1 = W^{-1}S_2, \quad D^y_2 = W^{-1}S_2, \\
D = ((D^x_1)^{-1} + (D^y_2)^{-1})^{-1}, \quad \pi^x = \mu^x(Y^x_1)^{-1}z^x_1, \quad \pi^x_2 = \mu^x(Y^x_2)^{-1}z^x_2, \\
D = ((D^y_1)^{-1} + (D^y_2)^{-1})^{-1}, \quad \pi^y = \mu^y(S^y_1)^{-1}w^y_1, \quad \pi^y_2 = \mu^y(S^y_2)^{-1}w^y_2,
\]

then \( \Delta x, \Delta y, \Delta s, \Delta w_1, \Delta w_2, \Delta z_1 \) and \( \Delta z_2 \), can be computed using the equations

\[
\begin{align*}
\pi^w &= \pi^w_1 - \pi^w_2, \\
\hat{y} &= y + \Delta y, \\
\Delta s &= -D_w(\hat{y} - \pi^w), \\
\pi^x &= \pi^x_1 - \pi^x_2, \\
\bar{x} &= x + \Delta x, \\
\Delta s_1 &= -D_w(w_1 \cdot (\bar{x} - \ell^x_1) + \mu^u(z_1 - z^x_1)), \\
\Delta s_2 &= -D_w(w_2 \cdot (u^x - \bar{x}) + \mu^u(z_2 - z^x_2)), \\
\Delta s &= s + \Delta s, \\
\Delta w_1 &= -D_w(w_1 \cdot (\bar{s} - \ell^y) + \mu^y(w_1 - w^y_1)), \\
\Delta w_2 &= -D_w(w_2 \cdot (u^y - \bar{s}) + \mu^y(w_2 - w^y_2)),
\end{align*}
\]

where \( \Delta x \) and \( \Delta y \) satisfy the equations

\[
\begin{pmatrix}
H(x, y) + D^{-1}_x \quad -J(x)T \\
J(x) \\
D_D + D_w
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta y
\end{pmatrix}
= -\begin{pmatrix}
g(x) - J(x)y - \pi^y \\
D_D(y - \pi^x) + D_w(y - \pi^w)
\end{pmatrix}. \tag{8.19}
\]

As \((x, s) \to (x^*, s^*)\) it holds that \( \|D_w\| \to \infty \), which implies that the matrix and right-hand side of this system goes to infinity. If \( D_w \) is the diagonal matrix such that \( D_w^2 = D_w^{-1} \), equations (8.19) may be written in the form

\[
\begin{pmatrix}
H(x, y) + D^{-1}_x \quad -D_wJ(x)T \\
D_wJ(x) \\
D_w^2D_D + I
\end{pmatrix}
\begin{pmatrix}
\Delta x \\
\Delta \hat{y}
\end{pmatrix}
= -\begin{pmatrix}
g(x) - J(x)y - \pi^y \\
D_w(D_D(y - \pi^x) + D_w(y - \pi^w))
\end{pmatrix}, \quad \Delta y = D_w \Delta \hat{y}. \tag{8.20}
\]
In this case, the scaled KKT matrix remains bounded if $H(x,y)$ is bounded. Similarly, the right-hand side remains bounded if $\|D_w D_w (y - \pi^w)\|$ is bounded.

The associated line-search merit function $M(x,s,y,z_1,z_2,w_1,w_2)$ of (8.18) can be written as

$$f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^p} \|c(x) - s\|^2 + \frac{1}{2\mu^p} \|c(x) - s + \mu^p (y - y^e)\|^2$$

$$- \sum_{j=1}^n \left\{ \mu^p [z_1^e]_j \ln \left( \left[ z_1 \right]_j (x_j - \ell_j^x + \mu^e) \right) - \left[ z_1 \right]_j (x_j - \ell_j^x + \mu^e) \right\}$$

$$- \sum_{j=1}^m \left\{ \mu^p [z_2^e]_j \ln \left( \left[ z_2 \right]_j (u_j^x - x_j + \mu^e) \right) - \left[ z_2 \right]_j (u_j^x - x_j + \mu^e) \right\}$$

$$- \sum_{i=1}^n \left\{ \mu^p [w_1^e]_i \ln \left( \left[ w_1 \right]_i (s_i - \ell_i^s + \mu^e) \right) - \left[ w_1 \right]_i (s_i - \ell_i^s + \mu^e) \right\}$$

$$- \sum_{i=1}^m \left\{ \mu^p [w_2^e]_i \ln \left( \left[ w_2 \right]_i (u_i^s - s_i + \mu^e) \right) - \left[ w_2 \right]_i (u_i^s - s_i + \mu^e) \right\},$$

for which the gradient $\nabla M(x,s,y,z_1,z_2,w_1,w_2)$ can be written as

$$\begin{pmatrix}
g - J^T (\pi^y + (\pi^y - y)) - (\pi^x + (\pi^x - z)) \\
\pi^y + (\pi^y - y) - (\pi^w + (\pi^w - w)) \\
-D_y (\pi^y - y) \\
-D_1^w (\pi_1^w - z_1) \\
-D_2^w (\pi_2^w - z_2) \\
-D_3^w (\pi_1^w - w_1) \\
-D_2^w (\pi_2^w - w_2)
\end{pmatrix},$$

where $z = z_1 - z_2$, $w = w_1 - w_2$. 
The residuals of the unsymmetric path-following equations (8.5) may be written as

\[
\begin{pmatrix}
g - J^T y - z \\
y - w \\
c - s + \mu^y(y - y^e) \\
z_1 \cdot (x - \ell^x) + \mu^x(z_1 - z_1^e) \\
z_2 \cdot (u^x - x) + \mu^x(z_2 - z_2^e) \\
w_1 \cdot (s - \ell^s) + \mu^y(w_1 - w_1^e) \\
w_2 \cdot (u^s - s) + \mu^y(w_2 - w_2^e)
\end{pmatrix}
= \begin{pmatrix}
g - J^T y - z \\
y - w \\
c - s + \mu^y(y - y^e) \\
z_1 \cdot (x - \ell^x) + \mu^x(z_1 - z_1^e) \\
z_2 \cdot (u^x - x) + \mu^x(z_2 - z_2^e) \\
w_1 \cdot (s - \ell^s) + \mu^y(w_1 - w_1^e) \\
w_2 \cdot (u^s - s) + \mu^y(w_2 - w_2^e)
\end{pmatrix},
\]

with \( z = z_1 - z_2 \) and \( w = w_1 - w_2 \).
9. General case: upper and lower Bounds on all variables

Next we consider the case with upper and lower bounds on both the variables and slacks, together with both implicit and explicit bounds on the variables.

9.1. Problem statement and optimality conditions

The definition of the problem is

\[
\text{minimize } f(x) \quad \text{subject to} \quad \begin{align*}
    c(x) - s &= 0, \\
    L_X s &= h_X, \\
    \ell^s &\leq L_F s \leq u^s, \\
    Ax - b &= 0, \\
    E_X x &= b_X, \\
    \ell^x &\leq E_P x \leq u^x,
\end{align*}
\] (9.1)

where \( c : \mathbb{R}^n \to \mathbb{R}^m \) and \( f : \mathbb{R}^n \to \mathbb{R} \) are twice-continuously differentiable. Throughout the discussion, the functions \( c : \mathbb{R}^n \to \mathbb{R}^m \) and \( f : \mathbb{R}^n \to \mathbb{R} \) are assumed to be twice-continuously differentiable. The matrices associated with the linear constraints \( E_X x = b_X \) and \( Ax = b \) have linearly independent rows. The matrices \( L_X \) and \( L_F \) are formed from rows of the identity matrix \( I_m \) in such a way that \( s_X = L_X s \) and \( s_F = L_F s \) are the vectors of “fixed” and “free” components of \( s \). It follows that there is an \( m \times m \) permutation matrix \( P \) such that

\[
P^* = \begin{pmatrix} L_F \\ L_X \end{pmatrix},
\]

with the matrices \( L_F \) and \( L_X \) satisfying the identities \( L_F L_F^T = I_F \), \( L_X L_X^T = I_X \), and \( L_F L_F^T = 0 \). The matrices \( E_X \) and \( E_P \) define an analogous partition of \( x \) into fixed and free components \( x_F \) and \( x_X \) of \( x \). The bound constraints involving \( E_X \) and \( L_X \) are enforced exactly. The linear constraints \( Ax - b = 0 \) are imposed using a shifted primal-dual penalty method.

The first-order KKT conditions for problem (9.1) are

\[
\begin{align*}
    g(x^*) - J(x^*) y^* - A^T u^* - E_X^T z_X^* - E_P^T z_F^* + E_P^T z_F^* &= 0, \\
    y^* - L_X^T w_X^* - L_F^T w_F^* + L_F^T w_F^* &= 0, \\
    c(x^*) - s^* &= 0, \\
    Ax^* - b &= 0, \\
    E_P x^* - \ell^x &\geq 0, \\
    u^x - E_P x^* &\geq 0, \\
    L_F s^* - \ell^s &\geq 0, \\
    u^s - L_F s^* &\geq 0, \\
    z_X^* \cdot (E_P x^* - \ell^x) &= 0, \\
    z_F^* \cdot (u^x - E_P x^*) &= 0, \\
    w_X^* \cdot (L_F s^* - \ell^s) &= 0, \\
    w_F^* \cdot (u^s - L_F s^*) &= 0,
\end{align*}
\] (9.2a-9.2h)

where \( y^* \) are the multipliers for the equality constraints \( c(x) - s = 0 \), and \( z_X^* \), \( z_F^* \), \( w_X^* \), and \( w_F^* \) may be interpreted as the Lagrange multipliers for the inequality constraints \( E_P x - \ell^x \geq 0 \), \( u^x - E_P x \geq 0 \), \( L_F s - \ell^s \geq 0 \) and \( u^s - L_F s \geq 0 \), respectively. The
components of $v^*$ are the multipliers for the linear equality constraints $Ax = b$. If $x_1 = E_{\ell^x} x - \ell^x$, $x_2 = u^x - E_{\ell^x} x$, $s_1 = L_{\ell^s} s - \ell^s$, and $s_2 = u^s - L_{\ell^s} s$, then $z_1^*, z_2^*, w_1^*$, and $w_2^*$ are the Lagrange multipliers for the inequality constraints $x_1 \geq 0$, $x_2 \geq 0$, $s_1 \geq 0$, and $s_2 \geq 0$, respectively.

9.2. The path-following equations

Let $z_1^*$ and $z_2^*$, $w_1^*$ and $w_2^*$ denote nonnegative estimates of $z_1^*$ and $z_2^*$, $w_1^*$ and $w_2^*$. Given small positive scalars $\mu^c$, $\mu^s$ and $\mu^p$, consider the perturbed optimality conditions

\[
\begin{align*}
g(x) - J(x)^T y - A^T v - E_x^T z_x - E_{\ell^x} z_1 + E_T^T z_2 &= 0, \\
y - L_x^T w_x - L_{\ell^s} w_1 + L_T^T w_2 &= 0, \\
c(x) - s &= \mu^c (y^e - y), \\
Ax - b &= \mu^s (v^e - v), \\
E_{\ell^x} x - \ell^x &\geq 0, \\
L_{\ell^s} s - \ell^s &\geq 0, \\
z_1 \cdot (E_{\ell^x} x - \ell^x) &= \mu^p (z_1^e - z_1^*), \\
z_2 \cdot (u^x - E_{\ell^x} x) &= \mu^p (z_2^e - z_2^*), \\
w_1 \cdot (L_{\ell^s} s - \ell^s) &= \mu^p (w_1^e - w_1^*), \\
w_2 \cdot (u^s - L_{\ell^s} s) &= \mu^p (w_2^e - w_2^*),
\end{align*}
\]

Consider the primal-dual path-following equations $F(x, s, y, v, w_x, z_x, z_1, z_2, w_1, w_2; \mu^c, \mu^s, \mu^p, y^e, v^e, z_1^e, z_2^e, w_1^e, w_2^e) = 0$, with

\[
F(x, s, y, v, z_1, z_2, w_1, w_2; \mu^c, \mu^s, \mu^p, y^e, v^e, z_1^e, z_2^e, w_1^e, w_2^e) = \begin{pmatrix}
g(x) - J(x)^T y - A^T v - E_x^T z_x - E_{\ell^x} z_1 + E_T^T z_2 \\
y - L_x^T w_x - L_{\ell^s} w_1 + L_T^T w_2 \\
c(x) - s + \mu^c (y - y^e) \\
Ax - b + \mu^s (v - v^e) \\
z_1 \cdot (E_{\ell^x} x - \ell^x) + \mu^p (z_1^e - z_1^*) \\
z_2 \cdot (u^x - E_{\ell^x} x) + \mu^p (z_2^e - z_2^*) \\
w_1 \cdot (L_{\ell^s} s - \ell^s) + \mu^p (w_1^e - w_1^*) \\
w_2 \cdot (u^s - L_{\ell^s} s) + \mu^p (w_2^e - w_2^*) \\
E_{\ell^x} x - b_x
\end{pmatrix}.
\]

Any zero $(x, s, y, v, w_x, z_x, z_1, z_2, w_1, w_2)$ of $F$ such that $\ell^x < E_{\ell^x} x < u^x$, $\ell^s < L_{\ell^s} s < u^s$, $z_1 > 0$, $z_2 > 0$, $w_1 > 0$, and $w_2 > 0$ approximates a point satisfying the optimality conditions (9.2), with the approximation becoming increasingly accurate as the terms $\mu^c (y - y^e)$, $\mu^s (v - v^e)$, $\mu^p (z_1 - z_1^*)$, $\mu^p (z_2 - z_2^*)$, $\mu^p (w_1 - w_1^*)$ and $\mu^p (w_2 - w_2^*)$ approach zero. For any sequence of $z_1^e$, $z_2^e$, $w_1^e$, $w_2^e$, $v^e$ and $y^e$ such that $z_1^e \to z_1^*$, $z_2^e \to z_2^*$, $w_1^e \to w_1^*$, $w_2^e \to w_2^*$, $v^e \to v^*$ and $y^e \to y^*$, and it must hold that solutions
(x, s, y, v, z_1, z_2, w_1, w_2) of (9.3) must satisfy \( z_1 \cdot (x - \ell^x) \to 0, z_2 \cdot (u^x - x) \to 0, w_1 \cdot (s - \ell^s) \to 0, \) and \( w_2 \cdot (u^s - s) \to 0. \) This implies that any solution \((x, s, y, v, w_X, z_X, z_1, z_2, w_1, w_2)\) of (9.3) will approximate a solution of (9.2) independently of the values of \( \mu^p, \mu^A \) and \( \mu^B \) (i.e., it is not necessary that \( \mu^p \to 0, \mu^A \to 0 \) and \( \mu^B \to 0 \)).

9.3. A shifted primal-dual penalty-barrier function

Problem (9.1) is equivalent to

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad c(x) - s = 0, \quad Ax - b = 0, \\
& \quad E_F x - x_1 = \ell^x, \quad L_F s - s_1 = \ell^s, \quad x_1 \geq 0, \quad s_1 \geq 0, \\
& \quad E_F x + x_2 = u^x, \quad L_F s + s_2 = u^s, \quad x_2 \geq 0, \quad s_2 \geq 0, \\
& \quad E_X x - b_X = 0, \quad L_X s - h_X = 0.
\end{align*}
\]

Consider the primal-dual shifted penalty-barrier problem

\[
\begin{align*}
\text{minimize} & \quad M(x, x_1, x_2, s, s_1, s_2, y, v, w_1, w_2; \mu^p, \mu^A, y^p, v^p, w_1^p, w_2^p) \\
\text{subject to} & \quad E_F x - x_1 = \ell^x, \quad L_F s - s_1 = \ell^s, \quad x_1 + \mu^p e > 0, \quad s_1 + \mu^A e > 0, \quad w_1 > 0, \\
& \quad E_F x + x_2 = u^x, \quad L_F s + s_2 = u^s, \quad x_2 + \mu^{A^c} e > 0, \quad s_2 + \mu^{B^c} e > 0, \quad w_2 > 0, \\
& \quad E_X x - b_X = 0, \quad L_X s - h_X = 0.
\end{align*}
\]
where $M(x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2; \mu^\epsilon, \mu^b, y^\epsilon, v^\epsilon, z_1^\epsilon, z_2^\epsilon, w_1^\epsilon, w_2^\epsilon)$ is the shifted penalty-barrier function

$$f(x) - (c(x) - s)^T y^\epsilon + \frac{1}{2\mu^\epsilon} \|c(x) - s\|^2 + \frac{1}{2\mu^b} \|c(x) - s + \mu^\epsilon (y - y^\epsilon)\|^2$$

$$- (Ax - b)^T v^\epsilon + \frac{1}{2\mu^\epsilon} \|Ax - b\|^2 + \frac{1}{2\mu^b} \|Ax - b + \mu^\epsilon (v - v^\epsilon)\|^2$$

$$- \sum_{j=1}^{m_F} \{\mu^b [z_1^\epsilon]_j \ln ([z_1]_j [x_1 + \mu^\epsilon e]) - [z_1 \cdot (x_1 + \mu^\epsilon e)]_j\}$$

$$- \sum_{j=1}^{m_F} \{\mu^\epsilon [z_2^\epsilon]_j \ln ([z_2]_j [x_2 + \mu^\epsilon e]) - [z_2 \cdot (x_2 + \mu^\epsilon e)]_j\}$$

$$- \sum_{i=1}^{m_F} \{\mu^\epsilon [w_1^\epsilon]_i \ln ([w_1]_i [s_1 + \mu^\epsilon e]) - [w_1 \cdot (s_1 + \mu^\epsilon e)]_i\}$$

$$- \sum_{i=1}^{m_F} \{\mu^\epsilon [w_2^\epsilon]_i \ln ([w_2]_i [s_2 + \mu^\epsilon e]) - [w_2 \cdot (s_2 + \mu^\epsilon e)]_i\}\} .$$

(9.6)

$$\nabla M(x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2) =$$

$$\begin{pmatrix}
  g - AT(2(v^\epsilon + \frac{1}{\mu^\epsilon}(Ax - b) - v)) - JT(2(y^\epsilon - \frac{1}{\mu^b}(c - s) - y)) \\
  z_1 - 2\mu^\epsilon (X_1^\epsilon)^{-1}z_1^\epsilon \\
  z_2 - 2\mu^\beta (X_2^\epsilon)^{-1}z_2^\epsilon \\
  2(y^\epsilon - \frac{1}{\mu^b}(c - s) - y) \\
  w_1 - 2\mu^\epsilon (S_1^\epsilon)^{-1}w_1^\epsilon \\
  w_2 - 2\mu^\beta (S_2^\epsilon)^{-1}w_2^\epsilon \\
  c - s + \mu^\epsilon (y - y^\epsilon) \\
  Ax - b + \mu^b (v - v^\epsilon) \\
  x_1 + \mu^\beta e - \mu^\epsilon Z_1^{-1}z_1^\epsilon \\
  x_2 + \mu^\beta e - \mu^\epsilon Z_2^{-1}z_2^\epsilon \\
  s_1 + \mu^\epsilon e - \mu^\beta W_1^{-1}w_1^\epsilon \\
  s_2 + \mu^\epsilon e - \mu^\beta W_2^{-1}w_2^\epsilon
\end{pmatrix}.$$
The gradient may be written in several equivalent forms

\[
\nabla M = \begin{pmatrix}
    g - A^T(2(v^e + \frac{1}{\mu^e}(Ax - b)) - v) - J^T(2(y^e - \frac{1}{\mu^e}(c - s)) - y) \\
    z_1 - 2\mu^e(X_1^I)^{-1}z_1^e \\
    z_2 - 2\mu^e(X_2^I)^{-1}z_2^e \\
    2(y^e - \frac{1}{\mu^e}(c - s)) - y \\
    w_1 - 2\mu^e(S_1^I)^{-1}w_1^e \\
    w_2 - 2\mu^e(S_2^I)^{-1}w_2^e \\
    c - s + \mu^e(y - y^e) \\
    Ax - b + \mu^e(v - v^e) \\
    x_1 + \mu^e(z_1 - z_1^e) \\
    x_2 + \mu^e(z_2 - z_2^e) \\
    s_1 + \mu^e(w_1 - w_1^e) \\
    s_2 + \mu^e(w_2 - w_2^e)
\end{pmatrix}
\]

9.4. Derivation of the shifted primal-dual penalty-barrier direction

9.5. Computation of the shifted primal-dual penalty-barrier direction

Next we consider the solution of the path-following Newton equations (9.4). If \(v = (x, s, y, v, w_x, z_x, z_1, z_2, w_1, w_2)\) is a given approximate zero of \(F(v)\) such that \(\ell^x - \mu^e < E_x < u^x + \mu^e, \ell^s - \mu^e < L_s < u^s + \mu^e, z_1 > 0, z_2 > 0, w_1 > 0,\) and
$w_2 > 0$, the Newton equations for the change in variables $\Delta v = (\Delta x, \Delta s, \Delta y, \Delta v, \Delta w_x, \Delta z_x, \Delta z_1, \Delta z_2, \Delta w_1, \Delta w_2)$ are given by $F'(v)\Delta v = -F(v)$, where

$$F(v) = \begin{pmatrix}
g(x) - J(x)^Ty - A^Tv - E_x^Tz_x - E_p^Tz_1 + E_p^Tz_2 \\
y - L_x^Tw_x - L_p^Tw_1 + L_p^Tw_2 \\
c(x) - s + \mu^e(y - y^e) \\
Ax - b + \mu^u(v - v^s) \\
L_x s - h_x \\
E_x x - b_x \\
z_1 \cdot (E_p x - \ell^x) + \mu^h(z_1 - z_1^e) \\
z_2 \cdot (u^x - E_p x) + \mu^h(z_2 - z_2^e) \\
w_1 \cdot (L_p s - \ell^x) + \mu^u(w_1 - w_1^s) \\
w_2 \cdot (u^x - L_p s) + \mu^u(w_2 - w_2^s)
\end{pmatrix},$$

and

$$F'(v) = 
\begin{pmatrix}
H & 0 & -J^T & -A^T & 0 & -E_x^T & -E_p^T & 0 & 0 \\
0 & 0 & I_m & 0 & -L_x^T & 0 & 0 & 0 & -L_p^T \\
J & -I_m & D_v & 0 & 0 & 0 & 0 & 0 & 0 \\
A & 0 & 0 & D_A & 0 & 0 & 0 & 0 & 0 \\
0 & L_p & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
E_x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
Z_1 E_p & 0 & 0 & 0 & 0 & X_1^a & 0 & 0 & 0 \\
-Z_2 E_p & 0 & 0 & 0 & 0 & 0 & X_2^a & 0 & 0 \\
0 & W_1 L_p & 0 & 0 & 0 & 0 & 0 & S_1^u & 0 \\
0 & -W_2 L_p & 0 & 0 & 0 & 0 & 0 & 0 & S_2^u
\end{pmatrix},$$

where $D_v = \mu^e I_m$, $D_A = \mu^s I_A$, $X_1^a = \text{diag}(E_p x - \ell^x + \mu^u e)$, $X_2^a = \text{diag}(u^x - E_p x + \mu^u e)$, $Z_1 = \text{diag}(z_1)$, $Z_2 = \text{diag}(z_2)$, $W_1 = \text{diag}([w_1]_i)$, $W_2 = \text{diag}([w_2]_i)$, $S_1^u = \text{diag}(L_p s - \ell^x + \mu^u e)$, and $S_2^u = \text{diag}(u^x - L_p s + \mu^u e)$.

Any $s$ may be written as $s = L_p^T s_p + L_x^T s_x$, where $s_p$ and $s_x$ denote the components of $s$ corresponding to the “free” and “fixed” components of $s$, respectively. Throughout, we assume that $s$ satisfies $L_x s - h_x = 0$, in which case $\Delta s_x = 0$ and $\Delta s$ satisfies

$$\Delta s = L_p^T \Delta s_p + L_x^T \Delta s_x = L_p^T \Delta s_p.$$ 

Similarly, any $x$ may be written as $x = E_p^T x_p + E_x^T x_x$, where $x_p$ and $x_x$ denote the components of $x$ corresponding to the “free” and “fixed variables”, respectively. Throughout, we assume that $x_x$ satisfies $E_x x - b_x = 0$, in which case $\Delta x_x = 0$ and $\Delta x$ satisfies

$$\Delta x = E_p^T \Delta x_p + E_x^T \Delta x_x = E_p^T \Delta x_p.$$
After premultiplying the first and fourth block of equations by $L_F$ and $A_F$ respectively, these identities allow us to write the equations (7.5) in the reduced form $\hat{F}^T \Delta v_F = -\hat{F}$, where $\Delta v_F = (\Delta x_F, \Delta s_F, \Delta y, \Delta w, \Delta x, \Delta z_1, \Delta z_2, \Delta w_1, \Delta w_2)$,

\[
\begin{pmatrix}
H_F & 0 & -J_F^T & -A_F^T & -I_F^T & I_F^T & 0 & 0 \\
0 & L_F & 0 & 0 & 0 & 0 & -I_F^T & I_F^T \\
J_F & -L_F & D_V & 0 & 0 & 0 & 0 & 0 \\
A_F & 0 & 0 & D_A & 0 & 0 & 0 & 0 \\
Z_1 & 0 & 0 & X_1^T & 0 & 0 & 0 & 0 \\
-Z_2 & 0 & 0 & 0 & X_2^T & 0 & 0 & 0 \\
0 & W_1 & 0 & 0 & 0 & 0 & S_1^T & 0 \\
0 & -W_2 & 0 & 0 & 0 & 0 & 0 & S_2^T \\
\end{pmatrix}
\begin{pmatrix}
\Delta x_F \\
\Delta s_F \\
\Delta v \\
\Delta w \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\end{pmatrix} = 
\begin{pmatrix}
g_F - J_F^T y - A_F^T v - z_1 + z_2 \\
y_F - w_1 + w_2 \\
c(x) - s + \mu^\ell(y - y^\ell) \\
Ax - b + \mu^\mu(v - v^\mu) \\
z_1 \cdot (E_F x - \ell^\xi) + \mu^\mu(z_1 - z_1^\ell) \\
z_2 \cdot (u^x - E_F x) + \mu^\mu(z_2 - z_2^\ell) \\
w_1 \cdot (L_F s - \ell^s) + \mu^\mu(w_1 - w_1^\ell) \\
w_2 \cdot (u^x - L_F s) + \mu^\mu(w_2 - w_2^\ell) \\
\end{pmatrix},
\]

(9.9)

where $A_F = AE_F^T$ are the columns of $A$ associated with the “free” components of $x$. The vectors $\Delta s$ and $\Delta w$ are recovered as $\Delta s = L_{F}^T \Delta s_F$, and $\Delta w = [y + \Delta y - w]_x$. Similarly, $\Delta x$ and $\Delta z$ are recovered as $\Delta x = L_{F}^T \Delta x_F$ and $\Delta z = [g + H \Delta x - J^T (y + \Delta y) - z]_x$. After scaling the last four blocks of equations by (respectively) $Z_1^{-1}$, $Z_2^{-1}$, $W_1^{-1}$ and $W_2^{-1}$, collecting terms and reordering the equations and unknowns, we obtain

\[
\begin{pmatrix}
D_A & 0 & 0 & 0 & 0 & A_F & 0 \\
0 & D_\ell^T & 0 & 0 & 0 & -I_F^T & 0 \\
0 & 0 & D_2^T & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & D_1^T & -I_F^T & 0 & 0 \\
0 & 0 & 0 & -I_F^T & I_F^T & 0 & 0 \\
0 & 0 & 0 & 0 & -L_F & J_F & D_V \\
0 & 0 & 0 & 0 & 0 & 0 & H_F - J_F^T \\
\end{pmatrix}
\begin{pmatrix}
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s_F \\
\Delta x_F \\
\Delta y \\
\end{pmatrix} = 
\begin{pmatrix}
D_A(v - \pi^\ell) \\
D_\ell^T(z_1 - \pi_1^\ell) \\
D_2^T(z_2 - \pi_2^\ell) \\
D_1^T(w_1 - \pi_1^\ell) \\
D_2^T(w_2 - \pi_2^\ell) \\
y_F - w_1 + w_2 \\
D_V(y - \pi^\mu) \\
g_F - J_F^T y - A_F^T v - z_1 + z_2 \\
\end{pmatrix},
\]

(9.10)

where $A_F = AE_F^T$ are the columns of $A$ associated with the “free” components of $x$, and

\[
D_V = \mu^\ell I_m, \quad \pi^\ell = y^\ell - \frac{1}{\mu^\ell}(c - s), \quad D_A = \mu^\mu I_A, \quad \pi^\mu = v^\mu - \frac{1}{\mu^\mu}(Ax - b), \\
D_1^W = S_1^T W_1^{-1}, \quad \pi_1^W = \mu^\mu(S_1^T)^{-1} w_1^\ell, \quad D_1^T = X_1^T Z_1^{-1}, \quad \pi_1^\ell = \mu^\mu(X_1^T)^{-1} z_1^\ell, \\
D_2^W = S_2^T W_2^{-1}, \quad \pi_2^W = \mu^\mu(S_2^T)^{-1} w_2^\ell, \quad D_2^T = X_2^T Z_2^{-1}, \quad \pi_2^\ell = \mu^\mu(X_2^T)^{-1} z_2^\ell.
\]

If we define $\tilde{H}_F = H_F + A_F^T D_A^{-1} A_F + (D_1^T)^{-1} + (D_2^T)^{-1}, \quad \tilde{D}_V = D_V + L_F^T D_W L_F$ and $D_W = ((D_1^W)^{-1} + (D_2^W)^{-1})^{-1}$, then
premultiplying the equations (9.10) by the matrix
\[
\begin{pmatrix}
I_A & I_p^x & I_p^y & I_p^z \\
0 & I_p^x & I_p^y & I_p^z \\
0 & 0 & I_p^x & I_p^z \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
A_p^T D_p^{-1} & (D_p^w)^{-1} & -(D_p^w)^{-1} & I_p^x \\
(D_p^w)^{-1} & -(D_p^w)^{-1} & I_p^x & 0 \\
0 & 0 & 0 & 0 \\
\end{pmatrix}
\]
gives the block upper-triangular system
\[
\begin{pmatrix}
D_x & 0 & 0 & 0 & 0 & 0 & A_p & 0 \\
0 & D_1^z & 0 & 0 & 0 & 0 & I_p^x & 0 \\
0 & 0 & D_2^z & 0 & 0 & 0 & -I_p^x & 0 \\
0 & 0 & 0 & D_3^z & 0 & I_p^x & 0 & 0 \\
0 & 0 & 0 & 0 & I_p^x & 0 & D_w L_p & 0 \\
0 & 0 & 0 & 0 & 0 & J_p & D_p & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & H_p & -J_p^T \\
\end{pmatrix}
\begin{pmatrix}
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s_f \\
\Delta x_F \\
\Delta y \\
\end{pmatrix}
= -
\begin{pmatrix}
D_A (v - \pi^w) \\
D_1^z (z_1 - \pi_1^w) \\
D_2^z (z_2 - \pi_2^w) \\
D_3^z (w_1 - \pi_1^w) \\
D_3^z (w_2 - \pi_2^w) \\
D_w (y - \pi^w) \\
J_p^T y - A_p^T \pi^w - \pi_p^w \\
\end{pmatrix},
\]
where \( \pi^w = L_p^T \pi_1^w - L_p^T \pi_2^w \) and \( \pi^z = E_p^T \pi_1^z - E_p^T \pi_2^z \). Using block back substitution, we may compute \( \Delta x \) and \( \Delta y \) by solving the equations
\[
\begin{pmatrix}
\bar{H}_p(x, y) + A_p^T D_x^{-1} A_p + D_x^{-1} & -J_p(x)^T \\
J_p(x) & D_x + L_p^T D_w L_p
\end{pmatrix}
\begin{pmatrix}
\Delta x_F \\
\Delta y
\end{pmatrix}
= -
\begin{pmatrix}
g_p(x) - J_p(x)^T (y - A_p \bar{\pi} - \pi_p^w) \\
L_p^T D_w (y - \pi_p^w) + D_y (y - \pi^y)
\end{pmatrix},
\]
with \( D_x = ((D_1^z)^{-1} + (D_2^z)^{-1})^{-1} \) and \( D_w = ((D_3^z)^{-1} + (D_2^z)^{-1})^{-1} \). The full vector \( \Delta x \) is then computed as \( \Delta x = E_p^T \Delta x_F \). Using the identity \( \Delta s = L_p^T \Delta s_p \) in the fifth block of equations gives
\[
\Delta s = -L_p^T D_w (y + \Delta y - \pi^w).
\]
There are several ways of computing \( \Delta w_1 \) and \( \Delta w_2 \). Instead of using the block upper-triangular system above, we use the last two blocks of equations of (9.9) to give
\[
\Delta w_1 = -(S_1^y)^{-1} (w_1 \cdot (L_p(s + \Delta s) - \ell^e + \mu^e e) - \mu^w w_1^0) \quad \text{and} \quad \Delta w_2 = -(S_2^y)^{-1} (w_2 \cdot (u^e - L_p(s + \Delta s) + \mu^e e) - \mu^w w_2^0),
\]
Similarly, using (9.9) to solve for $\Delta z_1$ and $\Delta z_2$ yields
\[
\Delta z_1 = -(X_1^u)^{-1}(z_1 \cdot (E_x(x + \Delta x) - \ell^x + \mu^a e) - \mu^a z_1^f) \quad \text{and} \quad \Delta z_2 = -(X_2^u)^{-1}(z_2 \cdot (u^x - E_x(x + \Delta x) + \mu^a e) - \mu^a z_2^f).
\]

Similarly, using the fourth and fifth block of equations of the Newton equations for a zero of (9.7) to solve for $\Delta v$ gives $\Delta v = -(v - \bar{v}^s)$, with $\bar{v}^s = v^s - \frac{1}{\mu^a}(A(x + \Delta x) - \bar{b})$. Finally, the vectors $\Delta w_x$ and $\Delta z_x$ are recovered as $\Delta w_x = [y + \Delta y - w]_x$ and $\Delta z_x = [g + H\Delta x - J^T(y + \Delta y) - z]_x$.

9.6. Summary: computations associated with the general problem

The results of the preceding section implies that the solution of the path-following equations $F'(v)\Delta v = -F(v)$ with $F$ and $F'$ given by (9.7) and (9.8) may be computed as follows. Let $x$ and $s$ be given primal variables such that $E_x x = b_x$, $L_x s = h_x$, with
\[
\ell^x - \mu^a e < E_p x < u^x + \mu^a e, \quad \ell^s - \mu^a e < L_p s < u^s + \mu^a e,
\]
and dual variables $y$, $w_1$, $w_2$, $z_1$, and $z_2$ such that $w_1 > 0$, $w_2 > 0$, $z_1 > 0$, and $z_2 > 0$. Let $X_1$, $X_2$, $S_1$, and $S_2$ denote the matrices diag([$x_p - \ell^x$], diag([$u^x - x_p$]), diag([$s_p - \ell^s$]), and diag([$u^s - s_p$]), respectively, and define the quantities
\[
D_v = \mu^a 1_m, \quad \pi_v = y^e - \frac{1}{\mu^a}(c - s),
\]
\[
D_\lambda = \mu^a 1_n, \quad \pi_\lambda = \hat{v} - \frac{1}{\mu^a}(Ax - b),
\]
\[
(D_1^x)^{-1} = (X_1^u)^{-1} Z_1, \quad (D_1^y)^{-1} = (S_1^u)^{-1} W_1,
\]
\[
(D_2^x)^{-1} = (X_2^u)^{-1} Z_2, \quad (D_2^y)^{-1} = (S_2^u)^{-1} W_2,
\]
\[
(D_2^z)^{-1} = (D_1^z)^{-1} + (D_2^z)^{-1}, \quad (D_2^w)^{-1} = (D_1^w)^{-1} + (D_2^w)^{-1},
\]
\[
\pi_1^z = \mu^a (X_1^u)^{-1} z_1^e, \quad \pi_1^w = \mu^a (S_1^u)^{-1} w_1^e,
\]
\[
\pi_2^z = \mu^a (X_2^u)^{-1} z_2^e, \quad \pi_2^w = \mu^a (S_2^u)^{-1} w_2^e,
\]
\[
\pi^z = E_f^T \pi_1^z - E_f^T \pi_2^z, \quad \pi^w = L_f^T \pi_1^w - L_f^T \pi_2^w.
\]

Solve the KKT system
\[
\begin{pmatrix}
H_f(x, y) + A_f^T D_\lambda^{-1} A_f + D_z^{-1} J_x(x)^T & -J_x(x)^T \\
J_x(x) & D_v + L_f^T D_w L_f
\end{pmatrix}
\begin{pmatrix}
\Delta x_f \\
\Delta y
\end{pmatrix}
= -\begin{pmatrix}
E_f(g(x) - J(x)^T y - A^T \pi^\lambda - \pi^z) \\
L_f^T D_w (y - \pi^w) + D_v(y - \pi^\lambda)
\end{pmatrix}.
\]
\[ \Delta x = E_p^T \Delta x_p \quad \hat{x} = x + \Delta x, \]
\[ \Delta z_1 = -(X^x_1)^{-1}(z_1 \cdot (E_p x - \ell^x + \mu^x e) - \mu^x z_1^x), \]
\[ \Delta z_2 = -(X^y_2)^{-1}(z_2 \cdot (u^x - E_p \hat{x} + \mu^y e) - \mu^y z_2^x), \]
\[ \hat{y} = y + \Delta y, \]
\[ \Delta s = -L^T_D \Delta \ell - \Delta \lambda, \]
\[ \hat{s} = s + \Delta s, \]
\[ \Delta w_1 = -(S^x_1)^{-1}(w_1 \cdot (L_p \hat{s} - \ell^x + \mu^x e) - \mu^x w_1^x), \]
\[ \Delta w_2 = -(S^y_2)^{-1}(w_2 \cdot (u^x - L_p \hat{s} + \mu^y e) - \mu^y w_2^x), \]
\[ \hat{v} = v + \Delta v \]
\[ \Delta w_x = \Delta v = \hat{v} - v, \]
\[ \Delta s_x = \Delta v = \hat{v} - v, \]
\[ \Delta w_x = \hat{v} - v, \]
\[ \Delta z_x = \hat{v} - v. \]

As \((x, s) \to (x^*, s^*)\) it holds that \(\|D^{-1}\|\) is bounded, but \(\|D_w\| \to \infty\) and \(\|A_p^T D^{-1}_z A_p\| \to \infty\). This implies that the matrix and right-hand side of this system goes to infinity. In the situation where \(A_p^T D^{-1}_z A_p\) is diagonal, then the KKT system can be rescaled so that the equations to be solved are bounded. If \(\hat{D}_z\) and \(\hat{D}_w\) denote diagonal matrices such that \(\hat{D}_z^2 = (A_p^T D^{-1}_z A_p)^{-1}\) and \(\hat{D}_w^2 = (L_p^T D_w L_p)^{-1}\), then \(\|\hat{D}_z\|\) and \(\|\hat{D}_w\|\) are bounded as \((x, s) \to (x^*, s^*)\). The equations (9.12) may be written in the form

\[
\begin{pmatrix}
\hat{D}_z H_p(x, y) \hat{D}_z + \hat{D}_z^2 D^{-1}_z + I & -(\hat{D}_w J_p(x) \hat{D}_z)^T \\
\hat{D}_w J_p(x) \hat{D}_z & \hat{D}_v + L_p^T D_w L_p
\end{pmatrix}
\begin{pmatrix}
\Delta x_p \\
\Delta y
\end{pmatrix}
= -\begin{pmatrix}
\hat{D}_z E_p (g(x) - J(x)^T y - A^T \pi^x - \pi^y) \\
\hat{D}_w (L_p^T D_w L_p (y - \pi^w) + D_v (y - \pi^v))
\end{pmatrix},
\]

with \(\Delta x_p = \hat{D}_z \Delta \hat{x}_p\) and \(\Delta y = \hat{D}_w \Delta \hat{y}\). In this case, the scaled KKT matrix remains bounded if \(H(x, y)\) is bounded. Similarly, the right-hand side remains bounded if \(\|\hat{D}_w L_p^T D_w L_p (y - \pi^w)\|\) is bounded.
The associated line-search merit function (9.6) can be written as

\[
\begin{align*}
 f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu_e}\|c(x) - s\|^2 + \frac{1}{2\mu^v}||c(x) - s + \mu^v(y - y^e)||^2 \\
&- (Ax - b)^Tv^e + \frac{1}{2\mu^x}||Ax - b||^2 + \frac{1}{2\mu^x}||Ax - b + \mu^x(v - v^e)||^2 \\
&- \sum_{j=1}^{n_F} \{\mu^e [z_1^e]_j \ln ([z_1]_j [x_F - \ell^x + \mu^e e]_j^2) - [z_1 \cdot (x_F - \ell^x + \mu^e e)]_j\} \\
&- \sum_{j=1}^{n_F} \{\mu^e [z_2^e]_j \ln ([z_2]_j [u^x - x_F + \mu^e e]_j^2) - [z_2 \cdot (u^x - x_F + \mu^e e)]_j\} \\
&- \sum_{i=1}^{m_F} \{\mu^e [w_1^e]_i \ln ([w_1]_i [s_F - \ell^s + \mu^e e]_i^2) - [w_1 \cdot (s_F - \ell^s + \mu^e e)]_i\} \\
&- \sum_{i=1}^{m_F} \{\mu^e [w_2^e]_i \ln ([w_2]_i [u^s - s_F + \mu^e e]_i^2) - [w_2 \cdot (u^s - s_F + \mu^e e)]_i\}.
\end{align*}
\]  

The residuals of the unsymmetric path-following equations may be written as

\[
 r = \begin{pmatrix}
 g - J^T y - z \\
 y - w \\
 c - s + \mu^v(y - y^e) \\
 z_1 \cdot (x - \ell^x) + \mu^x(z_1 - z_1^e) \\
 z_2 \cdot (u^x - x) + \mu^x(z_2 - z_2^e) \\
 w_1 \cdot (s - \ell^s) + \mu^x(w_1 - w_1^e) \\
 w_2 \cdot (u^s - s) + \mu^x(w_2 - w_2^e)
\end{pmatrix} = \begin{pmatrix}
 g - J^T y - z \\
 y - w \\
 \mu^v(y - \pi^v) \\
 X^1 \cdot (z_1 - \pi_1^e) \\
 X^2 \cdot (z_2 - \pi_2^e) \\
 S^1 \cdot (w_1 - \pi_1^e) \\
 S^2 \cdot (w_2 - \pi_2^e)
\end{pmatrix},
\]

with \( z = z_1 - z_2 \) and \( w = w_1 - w_2 \).
10. General case: upper and lower bounds on some of the variables

Finally, we assume that the problem has nonlinear equality constraints \( c(x) - s = 0 \), where \( s \) is the vector of slack variables. In addition, it is assumed that a subset of the components of \( x \) and \( s \) are fixed and that a subset of the other components are subject to upper and lower bounds.

10.1. Problem statement and optimality conditions

The problem of interest has the form

\[
\begin{align*}
\text{minimize} \quad & f(x) \\
\text{subject to} \quad & c(x) - s = 0, \quad L_X s = h_X, \quad \ell^s \leq L_L s, \quad L_U s \leq u^s, \\
& Ax - b = 0, \quad E_X x = b_X, \quad \ell^x \leq E_L x, \quad E_U x \leq u^x,
\end{align*}
\]

where \( c : \mathbb{R}^n \to \mathbb{R} \) and \( f : \mathbb{R}^n \to \mathbb{R} \) are twice-continuously differentiable. The quantity \( E_X \) denotes an \( n_X \times n \) matrix formed from \( n_X \) independent rows of \( I_n \), the identity matrix of order \( n \). This implies that the equality constraints \( E_X x = b_X \) fix \( n_X \) components of \( x \) at the corresponding values of \( b_X \). Similarly, \( E_L \) and \( E_U \) denote matrices formed from subsets of \( I_n \) such that \( E_L^T E_L = 0 \), \( E_U^T E_U = 0 \), i.e., a variable is either fixed or free to move, possibly bounded by an upper or lower bound. Note that a variable \( x_j \) need not be subject to a lower or upper bound, or may be bounded below and above, in which case \( e_j \) is not a row of \( E_X \), \( E_L \) or \( E_U \). Analogous definitions hold for \( L_X \), \( L_L \) and \( L_U \) as subsets of rows of \( I_m \). However, we impose the restriction that a given \( s_j \) must be either fixed or restricted by an upper or lower bound, i.e., there are no unrestricted slacks. The shifted primal-dual penalty-barrier equations can be derived without this restriction, but the derivation is beyond the scope of this note.

In addition, \( E_F \) and \( L_F \) denote rows of \( I_n \) and \( I_m \) such that \( (E_F^T \ E_U^T) \) and \( (E_L^T \ E_L^T) \) are column permutations of \( I_n \) and \( I_m \). It follows that the rows of \( E_F \) and \( E_U \) are a subset of the rows of \( E_F \), and that \( L_F \) is formed from the rows of \( L_L \) and \( L_U \). The bound constraints involving \( E_X \) and \( E_L \) are enforced explicitly. The linear constraints \( Ax - b = 0 \) are imposed using the shifted primal-dual augmented Lagrangian.

The first-order KKT conditions for problem (10.1) are

\[
\begin{align*}
g(x^*) - J(x^*)^T y^* - A^T v^* - E_X^T z_1^* - E_U^T z_2^* &= 0, \\
y^* - L_X^T w_X^* - L_L^T w_1^* + L_U^T w_2^* &= 0, \\
c(x^*) - s^* &= 0, \\
Ax^* - b &= 0, \\
E_L x^* - \ell^x &\geq 0, \\
L_L s^* - \ell^l &\geq 0, \\
z_1^* \cdot (E_L x^* - \ell^x) &= 0, \\
z_2^* \cdot (w^s - E_U x^*) &= 0, \\
w_1^* \cdot (L_L s^* - \ell^l) &= 0, \\
w_2^* \cdot (u^s - L_U s^*) &= 0,
\end{align*}
\]

where \( \ell^s \) are free variables. If we further impose the restriction that a subset of the rows of \( E_L \) and \( E_U \) are a subset of the rows of \( E_F \), and that \( L_F \) is formed from the rows of \( L_L \) and \( L_U \). The bound constraints involving \( E_X \) and \( E_F \) are enforced explicitly.
where \( y^* \) are the multipliers for the equality constraints \( c(x) - s = 0 \), and \( z_1^*, z_2^*, w_1^* \) and \( w_2^* \) may be interpreted as the Lagrange multipliers for the inequality constraints \( E_i x - \ell^i \geq 0 \), \( u^x - E_0 x \geq 0 \), \( L_i s - \ell^i \geq 0 \) and \( u^s - L_0 s \geq 0 \), respectively. The components of \( v^* \) are the multipliers for the linear equality constraints \( Ax = b \). If \( x_1 = E_i x - \ell^i \), \( x_2 = u^x - E_0 x \), \( s_1 = L_i s - \ell^i \), and \( s_2 = u^s - L_0 s \), then \( z_1^*, z_2^*, w_1^* \), and \( w_2^* \) are the Lagrange multipliers for the inequality constraints \( x_1 \geq 0 \), \( x_2 \geq 0 \), \( s_1 \geq 0 \), and \( s_2 \geq 0 \), respectively. In the derivations that follow, the vectors \( z \) and \( w \) are defined as

\[
\begin{align*}
z &= E_T^T u_1 - E_T^T z_2, \quad \text{and} \quad w = L_T^T w_1 - L_T^T w_2.
\end{align*}
\]

(10.3)

### 10.2. The path-following equations

Let \( z_1^* \) and \( z_2^* \), \( w_1^* \) and \( w_2^* \) denote nonnegative estimates of \( z_1^*, z_2^*, w_1^* \) and \( w_2^* \). Given small positive scalars \( \mu^e, \mu^s \) and \( \mu^a \), consider the perturbed optimality conditions

\[
\begin{align*}
g(x) - J(x)^T y - A^T v - E_T^T x - E_T^T z_1 + E_T^T z_2 &= 0, \\
y - L_T^T w_1 + L_T^T w_2 &= 0, \\
c(x) - s &= \mu^e (y^e - y), \\
er^x - b &= \mu^a (v^a - v), \\
e_i x - \ell^i \geq 0, \\
L_i s - \ell^i \geq 0, \\
z_1 \cdot (E_i x - \ell^i) &= \mu^a (z_1^e - z_1^a), \\
(z_2^e - z_2^a), \\
w_1 \cdot (L_i s - \ell^i) &= \mu^a (w_1^e - w_1^a), \\
w_2 \cdot (u^s - L_0 s) &= \mu^a (w_2^e - w_2^a).
\end{align*}
\]

(10.4a) - (10.4h)

Consider the primal-dual path-following equations \( F(x, s, y, w, x, z, z_1, z_2, w_1, w_2; \mu^e, \mu^s, y^e, v^e, z_1^e, z_2^e, w_1^e, w_2^e) = 0 \), with

\[
F(x, s, y, v, z_1, z_2, w_1, w_2; \mu^e, \mu^s, y^e, v^e, z_1^e, z_2^e, w_1^e, w_2^e) =
\begin{pmatrix}
g(x) - J(x)^T y - A^T v - E_T^T x - E_T^T z_1 + E_T^T z_2 \\
y - L_T^T w_1 + L_T^T w_2 \\
c(x) - s + \mu^e (y - y^e) \\
er^x - b + \mu^a (v - v^e) \\
e_i x - b &= \mu^a (v^a - v), \\
L_i s - \ell^i \geq 0, \\
z_1 \cdot (E_i x - \ell^i) &= \mu^a (z_1^e - z_1^a), \\
(z_2^e - z_2^a), \\
w_1 \cdot (L_i s - \ell^i) &= \mu^a (w_1^e - w_1^a), \\
w_2 \cdot (u^s - L_0 s) &= \mu^a (w_2^e - w_2^a).
\end{pmatrix}
\]

(10.5)
Any zero \((x, s, y, v, w_X, z_X, z_1, z_2, w_1, w_2)\) of \(F\) such that \(\ell^x < E_L, E_U x < u^x, \ell^s < L_L s, L_U < u^s, z_1 > 0, z_2 > 0, w_1 > 0,\) and \(w_2 > 0\) approximates a point satisfying the optimality conditions (10.2), with the approximation becoming increasingly accurate as the terms \(\mu^x(y - y^e), \mu^t(v - v^e), \mu^a(z_1 - z_1^f), \mu^a(z_2 - z_2^f), \mu^a(w_1 - w_1^f)\) and \(\mu^a(w_2 - w_2^f)\) approach zero. For any sequence of \(z_1^f, z_2^f, w_1^f, w_2^f, v^f\) and \(y^f\) such that \(z_1^f \rightarrow z_1^*, z_2^f \rightarrow z_2^*, w_1^f \rightarrow w_1^*, w_2^f \rightarrow w_2^*, v^f \rightarrow v^*\) and \(y^f \rightarrow y^*\), and it must hold that solutions \((x, s, y, v, z_1, z_2, w_1, w_2)\) of (10.4) must satisfy \(z_1^* \cdot (x - \ell^x) \rightarrow 0, z_2^* \cdot (u^x - x) \rightarrow 0, w_1^* \cdot (s - \ell^s) \rightarrow 0,\) and \(w_2^* \cdot (u^s - s) \rightarrow 0\). This implies that any solution \((x, s, y, v, w_X, z_X, z_1, z_2, w_1, w_2)\) of (10.4) will approximate a solution of (10.2) independently of the values of \(\mu^x, \mu^t, \mu^a\) (i.e., it is not necessary that \(\mu^x \rightarrow 0, \mu^t \rightarrow 0\) and \(\mu^a \rightarrow 0\)).

### 10.3. A shifted primal-dual penalty-barrier function

Problem (10.1) is equivalent to

\[
\begin{align*}
\text{minimize} & \quad f(x) \\
\text{subject to} & \quad c(x) - s = 0, \quad Ax - b = 0, \\
& \quad E_L x - x_1 = \ell^x, \quad L_L s - s_1 = \ell^s, \quad x_1 \geq 0, \quad s_1 \geq 0, \\
& \quad E_U x + x_2 = u^x, \quad L_U s + s_2 = u^s, \quad x_2 \geq 0, \quad s_2 \geq 0, \\
& \quad E_X x - b_X = 0, \quad L_X s - h_X = 0.
\end{align*}
\]

Consider the shifted primal-dual penalty-barrier problem

\[
\begin{align*}
\text{minimize} & \quad M(x, x_1, x_2, s, s_1, s_2, y, v, w_1, w_2; \mu^e, \mu^a, y^e, u^e, w_1^e, w_2^e) \\
\text{subject to} & \quad E_L x - x_1 = \ell^x, \quad L_L s - s_1 = \ell^s, \quad x_1 + \mu^a e > 0, \quad z_1 > 0, \quad s_1 + \mu^a e > 0, \quad w_1 > 0, \\
& \quad E_U x + x_2 = u^x, \quad L_U s + s_2 = u^s, \quad x_2 + \mu^a e > 0, \quad z_2 > 0, \quad s_2 + \mu^a e > 0, \quad w_2 > 0, \\
& \quad E_X x - b_X = 0, \quad L_X s - h_X = 0,
\end{align*}
\] (10.6)
where $M(x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2; \mu^e, \mu^y, y^e, v^e, z_1^e, z_2^e, w_1^e, w_2^e)$ is the shifted primal-dual penalty-barrier function

$$f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^e} \|c(x) - s\|^2 + \frac{1}{2\mu^e} \|c(x) - s + \mu^e (y - y^e)\|^2$$

$$- (Ax - b)^T v^e + \frac{1}{2\mu^e} \|Ax - b\|^2 + \frac{1}{2\mu^e} \|Ax - b + \mu^e (v - v^e)\|^2$$

$$- \sum_{j=1}^{n_F} \{ \mu^e [z_1^e]_j \ln ([z_1]_j [x_1 + \mu^e e]_j^2) - [z_1 \cdot (x_1 + \mu^e e)]_j \}$$

$$- \sum_{j=1}^{n_F} \{ \mu^e [z_2^e]_j \ln ([z_2]_j [x_2 + \mu^e e]_j^2) - [z_2 \cdot (x_2 + \mu^e e)]_j \}$$

$$- \sum_{i=1}^{m_F} \{ \mu^e [w_1^e]_i \ln ([w_1]_i [s_1 + \mu^e e]_i^2) - [w_1 \cdot (s_1 + \mu^e e)]_i \}$$

$$- \sum_{i=1}^{m_F} \{ \mu^e [w_2^e]_i \ln ([w_2]_i [s_2 + \mu^e e]_i^2) - [w_2 \cdot (s_2 + \mu^e e)]_i \}.$$ (10.7)

The gradient of $M$ may be defined in terms of the quantities $X_1^\mu = \text{diag}(E, x - \ell^e + \mu^e e)$, $X_2^\mu = \text{diag}(u^e - E, x + \mu^e e)$, $Z_1 = \text{diag}(z_1)$, $Z_2 = \text{diag}(z_2)$, $W_1 = \text{diag}([w_1]_i)$, $W_2 = \text{diag}([w_2]_i)$, $S_1^\mu = \text{diag}(L, s - \ell^e + \mu^e e)$ and $S_2^\mu = \text{diag}(u^e - L, s + \mu^e e)$. In particular,

$$\nabla M(x, x_1, x_2, s, s_1, s_2, y, v, z_1, z_2, w_1, w_2) =$$

$$\begin{pmatrix}
  g - A^T(2(y^e + \frac{1}{\mu^e}(Ax - b)) - v) - J^T(2(y^e - \frac{1}{\mu^e}(c - s)) - y) \\
  z_1 - 2\mu^e (X_1^\mu)^{-1} z_1^e \\
  z_2 - 2\mu^e (X_2^\mu)^{-1} z_2^e \\
  2(y^e - \frac{1}{\mu^e}(c - s)) - y \\
  w_1 - 2\mu^e (S_1^\mu)^{-1} w_1^e \\
  w_2 - 2\mu^e (S_2^\mu)^{-1} w_2^e \\
  e - s + \mu^e (y - y^e) \\
  Ax - b + \mu^e (v - v^e) \\
  x_1 + \mu^e e - \mu^e Z_1^{-1} z_1^e \\
  x_2 + \mu^e e - \mu^e Z_2^{-1} z_2^e \\
  s_1 + \mu^e e - \mu^e W_1^{-1} w_1^e \\
  s_2 + \mu^e e - \mu^e W_2^{-1} w_2^e
\end{pmatrix}.$$
The gradient may be written in several equivalent forms

\[ \nabla M = \begin{pmatrix}
  (g - A^T(2(v^e + \frac{1}{\mu^e}(Ax - b)) - v) - J^T(2(y^e - \frac{1}{\mu^e}(c - s)) - y)) \\
  z_1 - 2\mu^e (X_1^e + 1/z_1^e) \\
  z_2 - 2\mu^e (X_2^e + 1/z_2^e) \\
  2(y^e - \frac{1}{\mu^e}(c - s)) - y \\
  w_1 - 2\mu^e (S_1^e + 1/w_1^e) \\
  w_2 - 2\mu^e (S_2^e + 1/w_2^e) \\
  c - s + \mu^e(y - y^e) \\
  Ax - b + \mu^e(v - v^e) \\
  x_1 + \mu^e e - \mu^e Z_1^{-1} z_1^e \\
  x_2 + \mu^e e - \mu^e Z_2^{-1} z_2^e \\
  s_1 + \mu^e e - \mu^e W_1^{-1} w_1^e \\
  s_2 + \mu^e e - \mu^e W_2^{-1} w_2^e \\
\end{pmatrix}
\]

\[ \begin{pmatrix}
  (g - A^T(2(v^e + \frac{1}{\mu^e}(Ax - b)) - v) - J^T(2(y^e - \frac{1}{\mu^e}(c - s)) - y)) \\
  (X_1^e + 1/z_1^e)(z_1 \cdot x_1 + \mu^e z_1^e + \mu^e(z_1 - z_1^e)) \\
  (X_2^e + 1/z_2^e)(z_2 \cdot x_2 + \mu^e z_2^e + \mu^e(z_2 - z_2^e)) \\
  2(y^e - \frac{1}{\mu^e}(c - s)) - y \\
  (S_1^e + 1/w_1^e)(w_1 \cdot s_1 + \mu^e w_1^e + \mu^e(w_1 - w_1^e)) \\
  (S_2^e + 1/w_2^e)(w_2 \cdot s_2 + \mu^e w_2^e + \mu^e(w_2 - w_2^e)) \\
  c - s + \mu^e(y - y^e) \\
  Ax - b + \mu^e(v - v^e) \\
  Z_1^{-1}(z_1 \cdot x_1 + \mu^e(z_1 - z_1^e)) \\
  Z_2^{-1}(z_2 \cdot x_2 + \mu^e(z_2 - z_2^e)) \\
  W_1^{-1}(w_1 \cdot s_1 + \mu^e(w_1 - w_1^e)) \\
  W_2^{-1}(w_2 \cdot s_2 + \mu^e(w_2 - w_2^e)) \\
\end{pmatrix}
\]

\[ \begin{pmatrix}
  (g - A^T(\pi^e + (\pi^e - v)) - J^T(\pi^e + (\pi^e - y))) \\
  (\pi^e + (\pi^e - z_1)) \\
  (\pi^e + (\pi^e - z_2)) \\
  \pi^e + (\pi^e - y) \\
  (\pi^e + (\pi^e - w_1)) \\
  (\pi^e + (\pi^e - w_2)) \\
  D_v(\pi^e - y) \\
  D_a(\pi^e - v) \\
  D_1(\pi^e - z_1) \\
  D_2(\pi^e - z_2) \\
  D_1(\pi^e - w_1) \\
  D_2(\pi^e - w_2)
\end{pmatrix}
\]

where \( D_v = \mu^v I_m, D_a = \mu^a I_s, D_1 = X_1^e Z_1^{-1}, D_2 = X_2^e Z_2^{-1}, \pi_1^e = \mu^e (X_1^e + 1/z_1^e), \) and \( \pi_2^e = \mu^e (X_2^e + 1/z_2^e). \)
10.4. Derivation of the shifted primal-dual penalty-barrier direction

10.5. Computation of the shifted primal-dual penalty-barrier direction

Next we consider the solution of the path-following Newton equations (10.5). If \( v = (x, y, v, w_x, z_x, z_1, z_2, w_1, w_2) \) is a given approximate zero of \( F(v) \) such that \( \ell^x - \mu^u < E_i x, E_u x < u^s + \mu^u, \ell^y - \mu^u < L_{i} s, L_{u} s < u^s + \mu^u, z_1 > 0, z_2 > 0, w_1 > 0, \) and \( w_2 > 0 \), the Newton equations for the change in variables \( \Delta v = (\Delta x, \Delta s, \Delta y, \Delta v, \Delta w_x, \Delta z_x, \Delta z_1, \Delta z_2, \Delta w_1, \Delta w_2) \) are given by \( F'(v) \Delta v = -F(v) \), where

\[
F(v) = \begin{pmatrix}
g(x) - J(x)^T y - A^T v - z \\
y - w \\
c(x) - s + \mu^u (y - y^u) \\
A x - b + \mu^u (v - v^u) \\
L_x s - h_x \\
E_x x - b_x \\
z_1 \cdot (E_i x - \ell^x) + \mu^u (z_1 - z_1^e) \\
z_2 \cdot (u^x - E_u x) + \mu^u (z_2 - z_2^e) \\
w_1 \cdot (L_{i} s - \ell^y) + \mu^u (w_1 - w_1^e) \\
w_2 \cdot (u^s - L_{u} s) + \mu^u (w_2 - w_2^e)
\end{pmatrix}
\]

and

\[
F'(v) = \begin{pmatrix}
H & 0 & -J^T & -A^T & 0 & -E_i^T & -E_u^T & 0 & 0 \\
0 & I_m & 0 & -L_x^T & 0 & 0 & 0 & -L_{i}^T & L_{u}^T \\
-J & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & L_F & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & L_P & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-Z_1 E_i & 0 & 0 & 0 & 0 & X_1^u & 0 & 0 & 0 \\
-Z_2 E_u & 0 & 0 & 0 & 0 & X_2^u & 0 & 0 & 0 \\
0 & W_1 L_i & 0 & 0 & 0 & 0 & 0 & S_1^u & 0 \\
0 & -W_2 L_u & 0 & 0 & 0 & 0 & 0 & 0 & S_2^u
\end{pmatrix}
\]

(recall that \( z = E_x z_x + E_u z_1 - E_u z_2 \) and \( w = L_x w_x + L_D^T w_1 - L_u^T w_2 \). Any \( s \) may be written as \( s = L_P^T s_p + L_u^T s_x \), where \( L_F \) are the rows of \( I_m \) orthogonal to the rows of \( L_x \), i.e., \( L_F^T L_x = 0 \). The vectors \( s_p \) and \( s_x \) are the components of \( s \) corresponding to the “free” and “fixed” components of \( s \), respectively. The variables \( L_i s \) and \( L_u s \) form a subset of \( s_p \). Throughout, we assume that \( s \) satisfies \( L_x s - h_x = 0 \), in which case \( \Delta s_x = 0 \) and \( \Delta s \) satisfies

\[
\Delta s = L_F^T \Delta s_p + L_x^T \Delta s_x = L_F^T \Delta s_p.
\]
Similarly, any \( x \) may be written as \( x = E_F^T x_F + E_X^T x_X \), where \( x_F \) and \( x_X \) denote the components of \( x \) corresponding to the “free” and “fixed variables”, respectively. The variables \( E_L x \) and \( E_U x \) form a subset of \( x_F \). Throughout, we assume that \( x_X \) satisfies \( E_X x - b_X = 0 \), in which case \( \Delta x_X = 0 \) and \( \Delta x \) satisfies

\[
\Delta x = E_F^T \Delta x_F + E_X^T \Delta x_X = E_F^T \Delta x_F.
\]

After premultiplying the first and fifth blocks of equations of (10.9) by \( E_F \) and \( L_F \) respectively, and substituting \( \Delta x = E_F^T \Delta x_F \) and \( \Delta s = L_F^T \Delta s_F \), the equations (10.9) can be written in the reduced form \( \tilde{F}' \Delta v_F = -\tilde{F} \), where \( \Delta v_F = (\Delta x_F, \Delta s_F, \Delta y, \Delta v, \Delta z_1, \Delta z_2, \Delta w_1, \Delta w_2) \),

\[
\begin{pmatrix}
H_F & 0 & -J_F^T & -A_F^T & -E_{LF}^T & E_{UF}^T & 0 & 0 \\
0 & 0 & L_F & 0 & 0 & 0 & -L_{LF}^T & L_{UF}^T \\
J_F & -L_F^T & D_\gamma & 0 & 0 & 0 & 0 & 0 \\
A_F & 0 & 0 & D_A & 0 & 0 & 0 & 0 \\
2E_{LF} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2E_{UF} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & W_1L_{LF} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -W_2L_{UF} & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
\Delta x_F \\
\Delta s_F \\
\Delta y \\
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\end{pmatrix}
= -\begin{pmatrix}
E_F (g - J_F^T y - A_F^T v - z) \\
L_F (y - w) \\
c(x) - s + \mu^p(y - y^F) \\
A x - b + \mu^q(v - v^F) \\
2(z_1 \cdot (E_L x - \ell^X) + \mu^q(z_2 - z^F_2) \\
2(z_2 \cdot (u^X - E_U x) + \mu^p(z_2 - z^F_2) \\
w_1 \cdot (L_1 s - \ell^F) + \mu^p(w_1 - w^F_1) \\
w_2 \cdot (u^F - L_2 s) + \mu^q(w_1 - w^F_2) \\
\end{pmatrix},
\]

where \( H_F = E_A^T H E_F^T \), \( J_F = J E_F^T \), \( A_F = A E_F^T \), \( g_F = E_F g \), \( E_{LF} = E_L E_F^T \), \( E_{UF} = E_U E_F^T \), \( y_F = L_F y \), \( L_{LF} = L_L L_{LF}^T \) and \( L_{UF} = L_L L_{UF}^T \). The matrices \( J_F, A_F, E_{LF}, E_{UF} \) are the columns of \( J, A, E, E_U \) associated with the “free” components of \( x \). The matrices \( L_{LF}, L_{UF} \) are the columns of \( L_L \) associated with the “free” components of \( s \). Given the definitions (10.3), the vectors \( \Delta s \) and \( \Delta w_X \) are recovered as \( \Delta s = L_F^T \Delta s_F \) and \( \Delta w_X = [y + \Delta y - w]_X \). Similarly, \( \Delta x \) and \( \Delta z_X \) are recovered as \( \Delta x = L_F^T \Delta x_F \) and \( \Delta z_X = [g + H \Delta x - J_T^T (y + \Delta y) - z]_X \). After scaling the last four blocks of equations by (respectively) \( Z_1^{-1} \), \( Z_2^{-1} \), \( W_1^{-1} \) and \( W_2^{-1} \), collecting terms and reordering the equations and unknowns, we obtain

\[
\begin{pmatrix}
D_A & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
D_\ell^F & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & D_\ell^F & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & D_\ell^F & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & D_\ell^W & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & D_\ell^W & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & D_F & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & D_F & 0 \\
\end{pmatrix}
\begin{pmatrix}
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s_F \\
\Delta x_F \\
\Delta y \\
\end{pmatrix}
= -\begin{pmatrix}
D_A (v - \pi^A) \\
D_\ell^F (z_1 - \pi_1^A) \\
D_\ell^F (z_2 - \pi_2^A) \\
D_\ell^W (w_1 - \pi_1^W) \\
D_\ell^W (w_2 - \pi_2^W) \\
L_F (y - w) \\
D Fir (y - \pi^v) \\
E_F (g - J_F^T y - A_F^T v - z) \\
\end{pmatrix},
\]
where \( A_f = AE_f^T \) are the columns of \( A \) associated with the “free” components of \( x \), and

\[
D_y = \mu^\nu I_m, \quad \pi^y = y^e - \frac{1}{\mu^\nu}(c - s), \quad D_A = \mu^\lambda I_A, \quad \pi^A = v^e - \frac{1}{\mu^\lambda}(Ax - b),
\]

\[
D_1^w = S_1^w W_1^{-1}, \quad \pi_1^y = \mu^\beta (S_1^w)^{-1} w_1^e, \quad D_1^z = X_1^\mu Z_1^{-1}, \quad \pi_1^z = \mu^\beta (X_1^\mu)^{-1} z_1^e,
\]

\[
D_2^w = S_2^w W_2^{-1}, \quad \pi_2^y = \mu^\beta (S_2^w)^{-1} w_2^e, \quad D_2^z = X_2^\mu Z_2^{-1}, \quad \pi_2^z = \mu^\beta (X_2^\mu)^{-1} z_2^e.
\]

The diagonal matrix \( L_f(L_f^T(D_f^w)^{-1}L_v + L_f^T(D_f^w)^{-1}L_v) L_f^T \) is nonsingular if every slack is either fixed or bounded above or below. If we define \( D_w = (L_f(L_f^T(D_f^w)^{-1}L_v + L_f^T(D_f^w)^{-1}L_v) L_f^T)^{-1} \), then premultiplying the equations (10.11) by the matrix

\[
\begin{pmatrix}
I_A & 0 & I_{L_f}^e & I_{U_f}^e \\
0 & I_{L_f}^e & I_{U_f}^e & 0 \\
0 & 0 & I_{L_f}^e & 0 \\
0 & 0 & 0 & I_{L_f}^e \\
A_f^T D_A^{-1} & E_{L_f}^T(D_f^1)^{-1} & -E_{U_f}^T(D_f^2)^{-1} & 0 \\
L_f^T D_w L_{U_f}^e(D_f^1)^{-1} & -L_f^T D_w L_{U_f}^e(D_f^2)^{-1} & -L_f^T D_w L_{U_f}^e(D_f^2)^{-1} & 0 \\
L_f^T D_w L_{U_f}^e(D_f^2)^{-1} & 0 & 0 & I_m
\end{pmatrix}
\]

gives the block upper-triangular system

\[
\begin{pmatrix}
D_A & 0 & 0 & 0 & 0 & A_f & 0 \\
0 & D_1^x & 0 & 0 & 0 & E_{L_f}^x & 0 \\
0 & 0 & D_2^x & 0 & 0 & -E_{U_f}^x & 0 \\
0 & 0 & 0 & D_1^w & 0 & L_{L_f}^e & 0 \\
0 & 0 & 0 & D_2^w & -L_{U_f}^e & 0 & 0 \\
0 & 0 & 0 & 0 & D_w^{-1} & 0 & L_{U_f}^e \\
0 & 0 & 0 & 0 & 0 & J_f & D_f^x \\
0 & 0 & 0 & 0 & 0 & 0 & \hat{H}_f - J_f^T
\end{pmatrix}
\begin{pmatrix}
\Delta v \\
\Delta z_1 \\
\Delta z_2 \\
\Delta w_1 \\
\Delta w_2 \\
\Delta s_f \\
\Delta x_f \\
\Delta y
\end{pmatrix} =
\begin{pmatrix}
D_A(v - \pi^A) \\
D_1^x(z_1 - \pi_1^x) \\
D_2^x(z_2 - \pi_2^x) \\
D_1^w(w_1 - \pi_1^w) \\
D_2^w(w_2 - \pi_2^w) \\
L_f^e(y - \pi^w) \\
E_f^x(g - J_f^T y - A^T \pi^x - \pi^\xi_2) \\
E_f^x(y - \pi^w) + D_f(y - \pi^\nu)
\end{pmatrix},
\]

where \( \hat{H}_f = E_f^x(H + A^T D_A^{-1} A + E_f^T(D_f^1)^{-1} E_f + E_f^T(D_f^2)^{-1} E_f) E_f^T \), \( \hat{D}_f = D_f + L_f^T D_w L_{U_f}^e \), \( \pi^w = L_f^T \pi_1^w - L_{U_f}^e \pi_2^w \) and \( \pi^z = E_f^T \pi_1^z - E_f^T \pi_2^z \). Using block back substitution, \( \Delta x_f \) and \( \Delta y \) can be computed by solving the equations

\[
\begin{pmatrix}
H_f & - J_f^T \hat{D}_f \\
J_f & \hat{D}_f
\end{pmatrix}
\begin{pmatrix}
\Delta x_f \\
\Delta y
\end{pmatrix} =
\begin{pmatrix}
E_f^x(g - J_f^T y - A^T \pi^x - \pi^\xi_2) \\
L_f^T D_w L_{U_f}^e(y - \pi^w) + D_f(y - \pi^\nu)
\end{pmatrix}.
\]

(10.12)
The full vector $\Delta x$ is then computed as $\Delta x = E_T^T \Delta x_p$. Using the identity $\Delta s = L_T^T \Delta s_p$ in the sixth block of equations gives

$$\Delta s = -L_T^T D_w L_v(y + \Delta y - \pi^w).$$

There are several ways of computing $\Delta w_1$ and $\Delta w_2$. Instead of using the block upper-triangular system above, we use the last two blocks of equations of (10.10) to give

$$\Delta w_1 = -(S'^1)^{-1}(w_1 \cdot (L_e(s + \Delta s) - \ell^\xi + \mu^\sigma e) - \mu^\sigma w^e_1) \quad \text{and} \quad \Delta w_2 = -(S'^2)^{-1}(w_2 \cdot (u^\delta - L_v(s + \Delta s) + \mu^\sigma e) - \mu^\sigma w^e_2).$$

Similarly, using (10.10) to solve for $\Delta z_1$ and $\Delta z_2$ yields

$$\Delta z_1 = -(X'^1)^{-1}(z_1 \cdot (E_i(x + \Delta x) - \ell^\xi + \mu^\sigma e) - \mu^\sigma z^e_1) \quad \text{and} \quad \Delta z_2 = -(X'^2)^{-1}(z_2 \cdot (u^\delta - E_v(x + \Delta x) + \mu^\sigma e) - \mu^\sigma z^e_2).$$

Similarly, using the fourth and fifth block of equations of the Newton equations for a zero of (10.8) to solve for $\Delta v$ gives $\Delta v = -(v - \bar{\pi}^v)$, with $\bar{\pi}^v = v^\delta - \frac{1}{\mu^\delta}(A(x + \Delta x) - b)$. Finally, the vectors $\Delta w_x$ and $\Delta z_x$ are recovered as $\Delta w_x = [y + \Delta y - w]_x$ and $\Delta z_x = [g + H\Delta x - J^T(y + \Delta y) - z]_x$.

10.6. Summary: computations associated with the general problem

The results of the preceding section implies that the solution of the path-following equations $F'(v)\Delta v = -F(v)$ with $F$ and $F'$ given by (10.8) and (10.9) may be computed as follows. Let $x$ and $s$ be given primal variables and slack variables such that $E_x x = b_x$, $L_x s = h_x$ with $\ell^\xi - \mu^\sigma < E_l x$, $E_v x < u^\delta + \mu^\delta$, $\ell^\xi - \mu^\sigma < L_l s$, $L_v s < u^\delta + \mu^\delta$. Similarly, let $z_1$, $z_2$, $w_1$, $w_2$ and $y$ denote dual variables such that $w_1 > 0$, $w_2 > 0$, $z_1 > 0$, and $z_2 > 0$. Consider the diagonal matrices $X'^i = \text{diag}(E_i x - \ell^\xi + \mu^\delta e)$, $X^u = \text{diag}(u^\delta - E_v x + \mu^\delta e)$, $Z_1 = \text{diag}(z_1)$, $Z_2 = \text{diag}(z_2)$, $W_1 = \text{diag}([w_1]_1)$, $W_2 = \text{diag}([w_2]_1)$, $S'^1 = \text{diag}(L_l s - \ell^\xi + \mu^\delta e)$ and $S'^2 = \text{diag}(u^\delta - L_v s + \mu^\delta e)$. Given the quantities

$$D_v = \mu^\delta I_m, \quad D_s = \mu^\delta I_s, \quad (D'_x)^{-1} = (X'^1)^{-1} Z_1, \quad \pi^v = y^\delta - \frac{1}{\mu^\delta}(c - s),$$

$$D_s = \mu^\delta I_s, \quad (D'_s)^{-1} = (X'^2)^{-1} Z_2, \quad (D'^w)^{-1} = (S'^1)^{-1} W_1, \quad \pi^s = v^\delta - \frac{1}{\mu^\delta}(Ax - b),$$

$$D_w = E_T^T (D'_x)^{-1} E_l + E_T^T (D'_s)^{-1} E_v, \quad (D'^w)^{-1} = (S'^2)^{-1} W_2, \quad \pi'^w = L_T^T (D'^w)^{-1} L_v + L_T^T (D'^w)^{-1} L_v L_T^T,$$

$$\pi'^w = \mu^\delta (S'^2)^{-1} w_2^e, \quad \pi'^w = L_T^T (D'^w)^{-1} L_v L_T^T,$$

$$\pi'^w = \mu^\delta (S'^2)^{-1} w_2^e, \quad \pi'^w = L_T^T (D'^w)^{-1} L_v L_T^T.$$
10. General case: upper and lower bounds on some of the variables

Solve the KKT system

\[
\begin{pmatrix}
H_p(x, y) + A_p^T D_{\xi}^{-1} A_p + D_{\xi}^{-1} - J_p(x)^T & D_y + L_p^T D_w L_p
\end{pmatrix}
\begin{pmatrix}
\Delta x_p
\Delta y
\end{pmatrix}
= - \begin{pmatrix}
E_p(g(x) - J(x)^T y - A_p^T \pi^A - \pi^z)
L_p^T D_w L_p (y - \pi^w) + D_y (y - \pi^\gamma)
\end{pmatrix},
\tag{10.13}
\]

\[
\Delta x = E_p^T \Delta x_p, \quad \hat{x} = x + \Delta x,
\]

\[
\Delta z_1 = -(X_1^p)^{-1} (z_1 \cdot (E_p \hat{x} - \ell + \mu^e) - \mu^z Z_1^w),
\]

\[
\Delta z_2 = -(X_2^p)^{-1} (z_2 \cdot (w - E_p \hat{x} + \mu^e) - \mu^z Z_2^w),
\]

\[
\hat{y} = y + \Delta y,
\]

\[
\Delta s = -L_p^T D_w L_p (\hat{y} - \pi^w),
\]

\[
\hat{s} = s + \Delta s,
\]

\[
\Delta w_1 = - (S_1^w)^{-1} (w_1 \cdot (L_p \hat{s} - \ell + \mu^e) - \mu^w w_1^w),
\]

\[
\Delta w_2 = - (S_2^w)^{-1} (w_2 \cdot (w - L_p \hat{s} + \mu^e) - \mu^w w_2^w),
\]

\[
\hat{\pi}^A = v^e - \frac{1}{\mu^e} (A \hat{x} - b),
\]

\[
\Delta v = \hat{\pi}^A - v,
\]

\[
\hat{\nu} = \nu + \Delta \nu,
\]

\[
\Delta w = [\hat{y} - w]_x,
\]

\[
\Delta z = [g + H \Delta x - J^T \hat{y} - z]_x.
\]

As \((x, s) \to (x^*, s^*)\) it holds that \(\|D_{\varepsilon}^{-1}\|\) is bounded, but \(\|D_w\| \to \infty\) and \(\|A_p^T D_{\xi}^{-1} A_p\| \to \infty\). This implies that the matrix and right-hand side of this system goes to infinity. In the situation where \(A_p^T D_{\xi}^{-1} A_p\) is diagonal, then the KKT system can be rescaled so that the equations to be solved are bounded. If \(\hat{D}_x\) and \(\hat{D}_w\) denote diagonal matrices such that \(\hat{D}_x^2 = (A_p^T D_{\xi}^{-1} A_p)^{-1}\) and \(\hat{D}_w^2 = (L_p^T D_w L_p)^{-1}\), then \(\|\hat{D}_x\|\) and \(\|\hat{D}_w\|\) are bounded as \((x, s) \to (x^*, s^*)\). The equations (10.13) may be written in the form

\[
\begin{pmatrix}
\hat{D}_x H_p(x, y) \hat{D}_x + \hat{D}_x^2 \hat{D}_\varepsilon^{-1} + I - (\hat{D}_w J_p(x) \hat{D}_w)^T & D_y + L_p^T D_w L_p
\end{pmatrix}
\begin{pmatrix}
\Delta x_p
\Delta y
\end{pmatrix}
= - \begin{pmatrix}
\hat{D}_x E_p(g(x) - J(x)^T y - A_p^T \pi^A - \pi^z)
\hat{D}_w (L_p^T D_w L_p (y - \pi^w) + D_y (y - \pi^\gamma))
\end{pmatrix},
\tag{10.14}
\]

with \(\Delta x_p = \hat{D}_x \Delta \hat{x}_p\) and \(\Delta y = \hat{D}_w \Delta \hat{y}\). In this case, the scaled KKT matrix remains bounded if \(H(x, y)\) is bounded. Similarly, the right-hand side remains bounded if \(\|\hat{D}_w L_p^T D_w L_p (y - \pi^w)\|\) is bounded.
The associated line-search merit function (10.7) can be written as

\[ f(x) - (c(x) - s)^T y^e + \frac{1}{2\mu^e} \|c(x) - s\|^2 + \frac{1}{2\mu^e}\|c(x) - s + \mu^e(y - y^e)\|^2 \]

\[ \quad - (Ax - b)^TV^e + \frac{1}{2\mu^e}\|Ax - b\|^2 + \frac{1}{2\mu^e}\|Ax - b + \mu^e(v - v^e)\|^2 \]

\[ - \sum_{j=1}^{n_L} \left\{ \mu^e[z_1^e]_j \ln \left( \left[ \frac{z_1}{E_L} \right]_j \left[ E_L x - \ell^e + \mu^e e \right]_j \right) - \left[ z_1 \cdot (E_L x - \ell^e + \mu^e e) \right]_j \right\} \]

\[ - \sum_{j=1}^{n_U} \left\{ \mu^e[z_2^e]_j \ln \left( \left[ \frac{z_2}{E_U} \right]_j \left[ u^e - E_U x + \mu^e e \right]_j \right) - \left[ z_2 \cdot (u^e - E_U x + \mu^e e) \right]_j \right\} \]

\[ - \sum_{i=1}^{m_L} \left\{ \mu^e[w_1^e]_i \ln \left( \left[ \frac{w_1}{L_L} \right]_i \left[ L_L s - \ell^s + \mu^s e \right]_i \right) - \left[ w_1 \cdot (L_L s - \ell^s + \mu^s e) \right]_i \right\} \]

\[ - \sum_{i=1}^{m_U} \left\{ \mu^e[w_2^e]_i \ln \left( \left[ \frac{w_2}{L_U} \right]_i \left[ L_U s - \ell^u + \mu^u e \right]_i \right) - \left[ w_2 \cdot (L_U s - \ell^u + \mu^u e) \right]_i \right\} . \quad (10.15) \]

References