Interior point method using slacks

Albin Fredriksson

February 2025

1 Introduction

We consider the optimization problem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 0.5x'Hx + c'x \\ \text{subject to} \quad & l_A \leq Ax \leq u_A \\ & l_x \leq x \leq u_x \end{aligned}$$

where it is allowed that $(l_A)_i = (u_A)_i$ (equality constraints), $(l_A)_i = -\infty$ or $(u_A)_i = \infty$ (constraint unbounded from below/above), and $(l_x)_i = -\infty$ and/or $(u_x)_i = \infty$ (variable unbounded from below/above).

Add slacks to the constraints:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 0.5x'Hx + c'x \\ \text{subject to} \quad & Ax - s_A = 0 \\ & l_A \leq s_A \leq u_A \\ & l_x \leq x \leq u_x \end{aligned}$$

Add slacks to the bounds:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & 0.5x'Hx + c'x \\ \text{subject to} \quad & Ax - s = 0 \\ & s - g = l_A \\ & s + t = u_A \\ & x - y = l_x \\ & x + z = u_x \\ & g, t, y, z \geq 0 \end{aligned}$$

Now, the only inequalities are the nonnegativity constraints for the slack variables. We handle these inequalities by log-barriers:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & 0.5x'Hx + c'x - \mu e' \log(g) - \mu e' \log(t) - \mu_e' \log(y) - \mu e' \log(w) \\ \text{subject to} & Ax - s = 0 \\ & s - g = l_A \\ & s + t = u_A \\ & x - y = l_x \\ & x + z = u_x \end{array}$$

The Lagrangian of this problem is:

$$\begin{aligned} L(x, s, g, t, y, z, \lambda_A, \lambda_g, \lambda_t, \lambda_y, \lambda_z) = {} & 0.5x'Hx + c'x \\ & - \mu e' \log(g) - \mu e' \log(t) - \mu_e' \log(y) - \mu e' \log(z) \\ & + \lambda_A'(Ax - s) \\ & - \lambda_g'(s - g - l_A) + \lambda_t'(s + t - u_A) \\ & - \lambda_y'(x - y - l_x) + \lambda_z'(x + z - u_x) \end{aligned}$$

And the first-order optimality conditions are the following:

$$\begin{aligned} \nabla_x L &= Hx + c + A'\lambda_A - \lambda_y + \lambda_z = 0 \\ \nabla_{\lambda_A} L &= Ax - s = 0 \\ -\nabla_{\lambda_g} L &= s - g - l_A = 0 \\ \nabla_{\lambda_t} L &= s + t - u_A = 0 \\ -\nabla_{\lambda_y} L &= x - y - l_x = 0 \\ \nabla_{\lambda_z} L &= x + z - u_x = 0 \\ \nabla_s L &= -\lambda_A - \lambda_g + \lambda_t = 0 \\ \nabla_g L &= -\mu G^{-1} e + \lambda_g = 0 \\ \nabla_t L &= -\mu T^{-1} e + \lambda_t = 0 \\ \nabla_y L &= -\mu Y^{-1} e + \lambda_y = 0 \\ \nabla_z L &= -\mu Z^{-1} e + \lambda_z = 0 \end{aligned}$$

Rearrangement of the first-order optimality conditions and taking a step gives us:

$$
\begin{aligned}
H(x+\Delta x) + A'(\lambda_A + \Delta\lambda_A) - \lambda_y - \Delta\lambda_y + \lambda_z + \Delta\lambda_z &= -c \\
A(x+\Delta x) - s - \Delta s &= 0 \\
s + \Delta s - g - \Delta g - l_A &= 0 \\
s + \Delta s + t + \Delta t - u_A &= 0 \\
x + \Delta x - y - \Delta y - l_x &= 0 \\
x + \Delta x + z + \Delta z - u_x &= 0 \\
-\lambda_A - \Delta\lambda_A - \lambda_g - \Delta\lambda_g + \lambda_t + \Delta\lambda_t &= 0 \\
(G+\Delta G)(\Lambda_g + \Delta\Lambda_g)e &= \mu e \\
(T+\Delta T)(\Lambda_t + \Delta\Lambda_t)e &= \mu e \\
(Y+\Delta Y)(\Lambda_y + \Delta\Lambda_y)e &= \mu e \\
(Z+\Delta Z)(\Lambda_z + \Delta\Lambda_z)e &= \mu e
\end{aligned}
$$

We rearrange this to get the deltas on left and drop the nonlinear terms in delta:

$$
\begin{aligned}
H\Delta x + A'\Delta\lambda_A - \Delta\lambda_y + \Delta\lambda_z &= -c - Hx - A'\lambda_A + \lambda_y - \lambda_z \\
A\Delta x - \Delta s &= -Ax + s \\
\Delta s - \Delta g &= -s + g + l_A \\
\Delta s + \Delta t &= -s - t + u_A \\
\Delta x - \Delta y &= -x + y + l_x \\
\Delta x + \Delta z &= -x - z + u_x \\
-\Delta\lambda_A - \Delta\lambda_g + \Delta\lambda_t &= \lambda_A + \lambda_g - \lambda_t \\
\Lambda_g \Delta G e + G\Delta\Lambda_g e &= \mu e - G\Lambda_g e \\
\Lambda_t \Delta T e + T\Delta\Lambda_t e &= \mu e - T\Lambda_t e \\
\Lambda_y \Delta Y e + Y\Delta\Lambda_y e &= \mu e - Y\Lambda_y e \\
\Lambda_z \Delta Z e + Z\Delta\Lambda_z e &= \mu e - Z\Lambda_z e
\end{aligned}
$$

We can formulate these equations (for computing the Newton step) on matrix form:

$$
\begin{pmatrix}
H & A' & 0 & 0 & -I & I & 0 & 0 & 0 & 0 & 0 \\
A & 0 & 0 & 0 & 0 & 0 & -I & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I & -I & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & I & 0 & I & 0 & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -I & 0 \\
I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\
0 & -I & -I & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & G & 0 & 0 & 0 & 0 & \Lambda_g & 0 & 0 & 0 \\
0 & 0 & 0 & T & 0 & 0 & 0 & 0 & \Lambda_t & 0 & 0 \\
0 & 0 & 0 & 0 & Y & 0 & 0 & 0 & 0 & \Lambda_y & 0 \\
0 & 0 & 0 & 0 & 0 & Z & 0 & 0 & 0 & 0 & \Lambda_z
\end{pmatrix}
\begin{pmatrix}
\Delta x \\ \Delta\lambda_A \\ \Delta\lambda_g \\ \Delta\lambda_t \\ \Delta\lambda_y \\ \Delta\lambda_z \\ \Delta s \\ \Delta g \\ \Delta t \\ \Delta y \\ \Delta z
\end{pmatrix}
= -
\begin{pmatrix}
Hx + A'\lambda_A - \lambda_y + \lambda_z + c \\
Ax - s \\
s - g - l_A \\
s + t - u_A \\
x - y - l_x \\
x + z - u_x \\
-\lambda_A - \lambda_g + \lambda_t \\
G\Lambda_g e - \mu e \\
T\Lambda_t e - \mu e \\
Y\Lambda_y e - \mu e \\
Z\Lambda_z e - \mu e
\end{pmatrix}
= -
\begin{pmatrix}
r_x \\ r_{\lambda_A} \\ r_{\lambda_g} \\ r_{\lambda_t} \\ r_{\lambda_y} \\ r_{\lambda_z} \\ r_s \\ r_g \\ r_t \\ r_y \\ r_z
\end{pmatrix}
$$

Utilizing that $Z\Delta\lambda_z + \Lambda_z\Delta z = -r_z$, we get

$$\Delta z = -\Lambda_z^{-1} r_z - \Lambda_z^{-1} Z \Delta\lambda_z$$

and, similarly, for y, t, and g. We add row eleven multiplied by $-\Lambda_z^{-1}$ to row six etc.

$$\begin{pmatrix} H & A' & 0 & 0 & -I & I & 0 \\ A & 0 & 0 & 0 & 0 & 0 & -I \\ 0 & 0 & \Lambda_g^{-1}G & 0 & 0 & 0 & I \\ 0 & 0 & 0 & -\Lambda_t^{-1}T & 0 & 0 & I \\ I & 0 & 0 & 0 & \Lambda_y^{-1}Y & 0 & 0 \\ I & 0 & 0 & 0 & 0 & -\Lambda_z^{-1}Z & 0 \\ 0 & -I & -I & I & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda_A \\ \Delta\lambda_g \\ \Delta\lambda_t \\ \Delta\lambda_y \\ \Delta\lambda_z \\ \Delta s \end{pmatrix} = - \begin{pmatrix} r_x \\ r_{\lambda_A} \\ r_{\lambda_g} + \Lambda_g^{-1} r_g \\ r_{\lambda_t} - \Lambda_t^{-1} r_t \\ r_{\lambda_y} + \Lambda_y^{-1} r_y \\ r_{\lambda_z} - \Lambda_z^{-1} r_z \\ r_s \end{pmatrix}$$

We now use that

$$\Delta\lambda_y = -Y^{-1}\Lambda_y \Delta x - Y^{-1}\Lambda_y r_{\lambda_y} - Y^{-1} r_y$$

and

$$\Delta\lambda_z = Z^{-1}\Lambda_z \Delta x + Z^{-1}\Lambda_z r_{\lambda_z} - Z^{-1} r_z$$

to remove the fifth and sixth rows and the $\Delta\lambda_y$ and $\Delta\lambda_z$ columns, i.e., add row five multiplied by $Y^{-1}\Lambda_y$ and row six multiplied by $Z^{-1}\Lambda_z$ to row one to get:

$$\begin{pmatrix} H + Y^{-1}\Lambda_y + Z^{-1}\Lambda_z & A' & 0 & 0 & 0 \\ A & 0 & 0 & 0 & -I \\ 0 & 0 & \Lambda_g^{-1}G & 0 & I \\ 0 & 0 & 0 & -\Lambda_t^{-1}T & I \\ 0 & -I & -I & I & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda_A \\ \Delta\lambda_g \\ \Delta\lambda_t \\ \Delta s \end{pmatrix} = - \begin{pmatrix} r_x + Y_y^{-1} r_{\lambda_y} + Y^{-1} r_y + Z^{-1}\Lambda_z r_{\lambda_z} - Z^{-1} r_z \\ r_{\lambda_A} \\ r_{\lambda_g} + \Lambda_g^{-1} r_g \\ r_{\lambda_t} - \Lambda_t^{-1} r_t \\ r_s \end{pmatrix}$$

Now, we use that

$$\Delta\lambda_g = -G^{-1}\Lambda_g \Delta s - G^{-1}\Lambda_g r_{\lambda_g} - G^{-1} r_g$$

and

$$\Delta\lambda_t = T^{-1}\Lambda_t \Delta s + T^{-1}\Lambda_t r_{\lambda_t} - T^{-1} r_t$$

i.e., add row three multiplied by $G^{-1}\Lambda_g$ and four multiplied by $T^{-1}\Lambda_t$ to row five to get

$$\begin{pmatrix} H + Y^{-1}\Lambda_y + Z^{-1}\Lambda_z & A' & 0 \\ A & 0 & -I \\ 0 & -I & G^{-1}\Lambda_g + T^{-1}\Lambda_t \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda_A \\ \Delta s \end{pmatrix} = - \begin{pmatrix} r_x + Y_y^{-1} r_{\lambda_y} + Y^{-1} r_y + Z^{-1}\Lambda_z r_{\lambda_z} - Z^{-1} r_z \\ r_{\lambda_A} \\ r_s + G^{-1}\Lambda_g r_{\lambda_g} + G^{-1} r_g + T^{-1}\Lambda_t r_{\lambda_t} - T^{-1} r_t \end{pmatrix}$$

Finally, we add the third row multiplied by $(G^{-1}\Lambda_g + T^{-1}\Lambda_t)^{-1}$ to the second row, i.e., use the fact that

$$-\Delta\lambda_A + (G^{-1}\Lambda_g + T^{-1}\Lambda_t)\Delta s = -(r_s + G^{-1}\Lambda_g r_{\lambda_g} + G^{-1}r_g + T^{-1}\Lambda_t r_{\lambda_t} - T^{-1}r_t)$$

to arrive at

$$\begin{pmatrix} H + Y^{-1}\Lambda_y + Z^{-1}\Lambda_z & A' \\ A & -(G^{-1}\Lambda_g + T^{-1}\Lambda_t)^{-1} \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda_A \end{pmatrix} = -\begin{pmatrix} r_x + Y^{-1}\Lambda_y r_{\lambda_y} + Y^{-1}r_y + Z^{-1}\Lambda_z r_{\lambda_z} - Z^{-1}r_z \\ r_{\lambda_A} + (G^{-1}\Lambda_g + T^{-1}\Lambda_t)^{-1}(r_s + G^{-1}\Lambda_g r_{\lambda_g} + G^{-1}r_g + T^{-1}\Lambda_t r_{\lambda_t} - T^{-1}r_t) \end{pmatrix}$$

which is used to find Δx and $\Delta\lambda_A$. Then these can be used to get:

$$\begin{aligned}
\Delta s &= (G^{-1}\Lambda_g + T^{-1}\Lambda_t)^{-1}(\Delta\lambda_A - (r_s + G^{-1}\Lambda_g r_{\lambda_g} + G^{-1}r_g + T^{-1}\Lambda_t r_{\lambda_t} - T^{-1}r_t)) \\
\Delta\lambda_g &= -G^{-1}\Lambda_g\Delta s - G^{-1}\Lambda_g r_{\lambda_g} - G^{-1}r_g \\
\Delta\lambda_t &= T^{-1}\Lambda_t\Delta s + T^{-1}\Lambda_t r_{\lambda_t} - T^{-1}r_t \\
\Delta\lambda_y &= -Y^{-1}\Lambda_y\Delta x - Y^{-1}\Lambda_y r_{\lambda_y} - Y^{-1}r_y \\
\Delta\lambda_z &= Z^{-1}\Lambda_z\Delta x + Z^{-1}\Lambda_z r_{\lambda_z} - Z^{-1}r_z \\
\Delta g &= -\Lambda_g^{-1}r_g - \Lambda_g^{-1}G\Delta\lambda_g \\
\Delta t &= -\Lambda_t^{-1}r_t - \Lambda_t^{-1}T\Delta\lambda_t \\
\Delta y &= -\Lambda_y^{-1}r_y - \Lambda_y^{-1}Y\Delta\lambda_y \\
\Delta z &= -\Lambda_z^{-1}r_z - \Lambda_z^{-1}Z\Delta\lambda_z
\end{aligned}$$