# Project 1: Direct Methods in Optimization with Constraints

Iris Vukovic
Numerical Linear Algebra

November 8, 2024

Theoretical Derivations

T1: Show that the predictor steps reduces to solve a linear system with matrix $M_{KKT}$.

T2: Explain the previous derivations of the different strategies and justify under which assumptions they can be applied.

T3: Isolate $\delta_s$ from the 4th row of $M_{KKT}$ and substitute into the 3rd row. Justify that this procedure leads to a linear system with a symmetric matrix.

Here is our system

$$F(z) = \begin{pmatrix} r_L \\ r_A \\ r_C \\ r_s \end{pmatrix} = \begin{pmatrix} Gx + g - A\gamma - C\lambda \\ b - A^T x \\ s + d - C^T x \\ s_i \lambda_i \end{pmatrix}$$

where $z = (x, \gamma, \lambda, s)$.

**T1:** Upon differentiating with respect to each variable in $z$, we have

$$\frac{dr_L}{dz} = \begin{pmatrix} G & -A & -C & 0 \end{pmatrix} \tag{1}$$

$$\frac{dr_A}{dz} = \begin{pmatrix} -A^T & 0 & 0 & 0 \end{pmatrix} \tag{2}$$

$$\frac{dr_C}{dz} = \begin{pmatrix} -C^T & 0 & 0 & 0 \end{pmatrix} \tag{3}$$

$$\frac{dr_s}{dz} = \begin{pmatrix} 0 & 0 & S & \Lambda \end{pmatrix} \tag{4}$$

So our $M_{KKT}$ matrix is

$$\begin{pmatrix} G & -A & -C & 0 \\ -A^T & 0 & 0 & 0 \\ -C^T & 0 & 0 & 0 \\ 0 & 0 & S & \Lambda \end{pmatrix}$$

We want to update $z_0$ to $z_1 = z_0 + \delta_{z_0}$ to eventually reach the solution of $F(z)$, where our Newton step is represented by $\delta_{z_0} = (\delta_{x_0}, \delta_{\gamma_0}, \delta_{\lambda_0}, \delta_{s_0})$. So, using first order Taylor expansion and knowing that $M_{KKT}$ is the first differential of the system $F(z)$, we have that $F(z_1) = F(z_0 + \delta_{z_0}) = F(z_0) + M_{KKT}\delta_{z_0}$. To solve this, we set

$$F(z_0) + M_{KKT}\delta_{z_0} = 0$$

which implies

$$M_{KKT}\delta_{z_0} = -F(z_0)$$

Thus, using the predictor/Newton step, we are left with a linear system that we can solve.

**T2: Strategy 1**

This is our KKT system for $A = 0$ case:

$$\begin{pmatrix} G & -C & 0 \\ -C^T & 0 & 0 \\ 0 & S & \Lambda \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \\ \delta_s \end{pmatrix} = \begin{pmatrix} -r_L \\ -r_C \\ -r_s \end{pmatrix}$$

Thus, we have

$$\begin{aligned} G\delta_x - C\delta_\lambda &= -r_L \\ C^T\delta_x + \delta_s &= -r_C \\ S\delta_\lambda + \Lambda\delta_s &= -r_s \end{aligned}$$

Now we isolate $\delta_s$ from the third row of the KKT system and find that

$$\delta_s = \Lambda^{-1}(-r_s - S\delta_\lambda)$$

Plugging it back into the second row, we have

$$\begin{aligned} G\delta_x - C\delta_\lambda &= -r_L \\ C^T\delta_x - \Lambda^{-1}S\delta_s &= -r_C - \Lambda^{-1}r_s \end{aligned}$$

And back into matrix form, we now have:

$$\begin{pmatrix} G & -C \\ -C^T & -S\Lambda^{-1} \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_\lambda \end{pmatrix} = \begin{pmatrix} -r_L \\ -r_C - \Lambda^{-1}r_s \end{pmatrix}$$

**T2: Strategy 2**

Starting from the same system of equations,

$$\begin{aligned} G\delta_x - C\delta_\lambda &= -r_L \\ C^T\delta_x + \delta_s &= -r_C \\ S\delta_\lambda + \Lambda\delta_s &= -r_s \end{aligned}$$

This time, we isolate $\delta_s$ from the second row of the KKT system and find that

$$\delta_s = -r_C + C^T\delta_x$$

Then we plug $\delta_s$ into the third row to get

$$\begin{aligned}S\delta_\lambda + \Lambda(-r_C + C^T\delta_x) = -r_s\\ \delta_\lambda = S^{-1}(-r_s - \Lambda(-r_C + C^T\delta_x)\\ \delta_\lambda = S^{-1}(-r_s - \Lambda r_C) - S^{-1}\Lambda C^T\delta_x\end{aligned}$$

And finally, we plug $\delta_\lambda$ into the first row to obtain

$$\begin{aligned}G\delta_x - C(S^{-1}(-r_s + \Lambda r_C) - S^{-1}\Lambda C^T\delta_x) = -r_L\\ \delta_x(G + CS^{-1}\Lambda C^T) = -r_L + CS^{-1}(-r_s + \Lambda r_C)\\ \delta_x\hat{G} = -r_L - \hat{r}\end{aligned}$$

where $\hat{G} = G + CS^{-1}\Lambda C^T$ and $\hat{r} = -CS^{-1}(-r_s + \Lambda r_C)$.

To utilize these strategies, we must assume that for (1) $\Lambda$ is invertible and for (2) that $S$ is invertible. To solve the system using $LDL^T$ factorization or Cholesky factorization, $G$ must be positive definite aka having all positive eigenvalues otherwise the factorization will fail.

**T3:**
Given our full KKT system:

$$\begin{pmatrix} G & -A & -C & 0\\ -A^T & 0 & 0 & 0\\ -C^T & 0 & 0 & 0\\ 0 & 0 & S & \Lambda\end{pmatrix}\begin{pmatrix}\delta_x\\ \delta_\gamma\\ \delta_\lambda\\ \delta_s\end{pmatrix} = \begin{pmatrix}-r_L\\ -r_A\\ -r_C\\ -r_s\end{pmatrix}$$

We first isolate $\delta_s$ from the 4th row of the system

$$\begin{aligned}S\delta_\lambda + \Lambda\delta_s = -r_s\\ \delta_s = \Lambda^{-1}(-r_s - S\delta_\lambda)\end{aligned}$$

Then, we plug it into the 3rd row of the system

$$\begin{aligned}C^T\delta_x + \Lambda^{-1}(-r_s - S\delta_\lambda) = -r_C\\ -C^T\delta_x - \Lambda^{-1}S\delta_\lambda = -r_C - \Lambda^{-1}r_s\end{aligned}$$

So our system becomes

$$\begin{pmatrix} G & -A & -C\\ -A^T & 0 & 0\\ -C^T & 0 & -\Lambda^{-1}S\end{pmatrix}\begin{pmatrix}\delta_x\\ \delta_\gamma\\ \delta_\lambda\end{pmatrix} = \begin{pmatrix}-r_L\\ -r_A\\ -r_C - \Lambda^{-1}r_s\end{pmatrix}$$

Assuming that $G$, $A$, and $C$ are symmetric, we know that $S$ and $\Lambda$ are diagonal matrices which implies their symmetry. $\Lambda^{-1}$ is diagonal too, and the product of two diagonal matrices is also diagonal, thus $-\Lambda^{-1}S$ is symmetric. And so we can conclude that the reduced matrix is symmetric and we are left with a simpler linear system to solve.