Solutions to Part B of Problem Sheet 5

Solution (5.4) The last block of rows reads as

$$\boldsymbol{S}\Delta\boldsymbol{x} + \boldsymbol{X}\Delta\boldsymbol{s} = -\boldsymbol{X}\boldsymbol{S}\boldsymbol{e} + \sigma\mu\boldsymbol{e},$$

so that multiplying by $\boldsymbol{X}^{-1}$ (the diagonal matrix $\boldsymbol{X}$ is non-singular, since we are in $\mathcal{F}^{\circ}$) and solving for $\Delta\boldsymbol{s}$, we get

$$\Delta\boldsymbol{s} = -\boldsymbol{S}\boldsymbol{e} - \boldsymbol{X}^{-1}\boldsymbol{S}\Delta\boldsymbol{x} + \sigma\mu\boldsymbol{X}^{-1}\boldsymbol{e}.$$

Substituting $\Delta\boldsymbol{s}$ into the first block of rows, we get

$$\boldsymbol{A}^{\top}\Delta\boldsymbol{y} + \boldsymbol{X}^{-1}\boldsymbol{S}\Delta\boldsymbol{x} = \boldsymbol{S}\boldsymbol{e} - \sigma\mu\boldsymbol{X}^{-1}\boldsymbol{e}.$$

Using $\boldsymbol{D} = \boldsymbol{S}^{-1/2}\boldsymbol{X}^{1/2}$, we get the new system

$$\begin{pmatrix} \boldsymbol{0} & \boldsymbol{A} \\ \boldsymbol{A}^{\top} & -\boldsymbol{D}^{-2} \end{pmatrix} \begin{pmatrix} \Delta\boldsymbol{y} \\ \Delta\boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{s} - \sigma\mu\boldsymbol{X}^{-1}\boldsymbol{e} \end{pmatrix} \tag{1}$$

$$\Delta\boldsymbol{s} = -\boldsymbol{S}\boldsymbol{e} - \boldsymbol{X}^{-1}\boldsymbol{S}\Delta\boldsymbol{x} + \sigma\mu\boldsymbol{X}^{-1}\boldsymbol{e}. \tag{2}$$

Multiplying $\boldsymbol{A}\boldsymbol{D}^2$ to the second block of rows, we get

$$\boldsymbol{A}\boldsymbol{D}^2\boldsymbol{A}^{\top}\Delta\boldsymbol{y} - \boldsymbol{A}\Delta\boldsymbol{x} = \boldsymbol{A}\boldsymbol{D}^2\boldsymbol{s} - \sigma\mu\boldsymbol{A}\boldsymbol{D}^2\boldsymbol{X}^{-1}\boldsymbol{e},$$

which in view of $\boldsymbol{A}\Delta\boldsymbol{x} = \boldsymbol{0}$, $\boldsymbol{D}^2\boldsymbol{X}^{-1} = \boldsymbol{S}^{-1}$ and $\boldsymbol{D}^2\boldsymbol{s} = \boldsymbol{x}$ simplifies to

$$(\boldsymbol{A}\boldsymbol{D}^2\boldsymbol{A}^{\top})\Delta\boldsymbol{y} = \boldsymbol{b} - \sigma\mu\boldsymbol{A}\boldsymbol{S}^{-1}\boldsymbol{e}.$$

This gives a system of equations for recovering $\Delta\boldsymbol{y}$, with a symmetric coefficient matrix $\boldsymbol{A}\boldsymbol{D}^2\boldsymbol{A}^{\top}$. From the second block of (1) we also get

$$\boldsymbol{X}^{-1}\boldsymbol{S}\Delta\boldsymbol{x} - \sigma\mu\boldsymbol{X}^{-1}\boldsymbol{e} = \boldsymbol{A}^{\top}\Delta\boldsymbol{y} - \boldsymbol{s},$$

and substituting this into the expression for $\Delta\boldsymbol{s}$ in (2) above gives

$$\Delta\boldsymbol{s} = -\boldsymbol{A}^{\top}\Delta\boldsymbol{y}.$$

Finally, we can obtain $\Delta\boldsymbol{x}$ from $\Delta\boldsymbol{s}$ via (2),

$$\Delta\boldsymbol{x} = -\boldsymbol{x} - \boldsymbol{S}^{-1}\boldsymbol{X}\Delta\boldsymbol{s} + \sigma\mu\boldsymbol{S}^{-1}\boldsymbol{e}.$$

This is the last of the three equations. The benefit is that one only has to solve one system of equations with symmetric coefficient matrix $\boldsymbol{A}\boldsymbol{D}^2\boldsymbol{A}^{\top}$, which can be done efficiently using, for example, Cholesky factorization or other methods. Once $\Delta\boldsymbol{y}$ is found, one can compute the other parts $\Delta\boldsymbol{x}$ and $\Delta\boldsymbol{s}$ easily.