Sufficiency condition for a constrained minimum

The first order Taylor series expansion of the objective function and the equality constraints gives:

$$\begin{aligned} \delta f &= f - f^* = \nabla f^T \delta \mathbf{x} = \nabla_{\mathbf{s}} f^T \delta \mathbf{s} + \nabla_{\mathbf{d}} f^T \delta \mathbf{d} \\ \delta \mathbf{h} &= \mathbf{h} - \mathbf{h}^* = \nabla \mathbf{h}^T \delta \mathbf{x} = \nabla_{\mathbf{s}} \mathbf{h}^T \delta \mathbf{s} + \nabla_{\mathbf{d}} \mathbf{h}^T \delta \mathbf{d} \end{aligned} \tag{1}$$

Since $\delta \mathbf{h} = \mathbf{0}$ in order to satisfy feasibility, we get from the second of Eq. (1):

$$\delta \mathbf{s} = -\left[\nabla_{\mathbf{s}} \mathbf{h}^T\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T \delta \mathbf{d} \tag{2}$$

The substitution of $\delta \mathbf{s}$ of Eq. (2) into the first of Eq. (1) yields:

$$\begin{aligned} \delta f &= \nabla_{\mathbf{s}} f^T \delta \mathbf{s} + \nabla_{\mathbf{d}} f^T \delta \mathbf{d} = -\nabla_{\mathbf{s}} f^T \left[\nabla_{\mathbf{s}} \mathbf{h}^T\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T \delta \mathbf{d} + \nabla_{\mathbf{d}} f^T \delta \mathbf{d} \\ &= \left\{-\nabla_{\mathbf{s}} f^T \left[\nabla_{\mathbf{s}} \mathbf{h}^T\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T + \nabla_{\mathbf{d}} f^T\right\} \delta \mathbf{d} \\ &= \left\{\left(\frac{\partial f}{\partial \mathbf{s}}\right)^T \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T + \left(\frac{\partial f}{\partial \mathbf{d}}\right)^T\right\} \delta \mathbf{d} \end{aligned} \tag{3}$$

The last line of Eq. (3) uses a different notation to show the gradients for the convenience of further manipulations. Note that this change of notation also changes the way Eq. (2) is written, which is shown below.

$$\delta \mathbf{s} = -\left[\nabla_{\mathbf{s}} \mathbf{h}^T\right]^{-1} \nabla_{\mathbf{d}} \mathbf{h}^T \delta \mathbf{d} = -\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T \delta \mathbf{d} \tag{2'}$$

We now interpret δf as another "reduced function" δz wherein z depends only on independent (or decision) variables $\mathbf{d}$. From the last line of Eq. (3), the "reduced gradient" is given by

$$\begin{aligned} \delta f = \delta z &= \left\{\left(\frac{\partial f}{\partial \mathbf{s}}\right)^T \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T + \left(\frac{\partial f}{\partial \mathbf{d}}\right)^T\right\} \delta \mathbf{d} = \left(\frac{dz}{d\mathbf{d}}\right)^T \delta \mathbf{d} \\ \Rightarrow \left(\frac{dz}{d\mathbf{d}}\right)^T &= \left(\frac{\partial f}{\partial \mathbf{s}}\right)^T \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T + \left(\frac{\partial f}{\partial \mathbf{d}}\right)^T \\ \Rightarrow \left(\frac{dz}{d\mathbf{d}}\right) &= \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\left(\frac{\partial f}{\partial \mathbf{s}}\right) + \left(\frac{\partial f}{\partial \mathbf{d}}\right) \end{aligned} \tag{4}$$

Minimization of z is unconstrained. So, the sufficient condition is that its Hessian is positive definite. In the new notation we adopted, the Hessian can be written as

$$\mathbf{H}_{\mathbf{d}}(z) = \frac{d}{d\mathbf{d}}\left(\frac{dz}{d\mathbf{d}}\right) = \frac{d^2 z}{d\mathbf{d}^2} = \frac{d}{d\mathbf{d}}\left\{\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\left(\frac{\partial f}{\partial \mathbf{s}}\right) + \left(\frac{\partial f}{\partial \mathbf{d}}\right)\right\} \tag{5}$$

It can be expanded using the chain rule as follows.

$$\begin{aligned}\frac{d^2 z}{d\mathbf{d}^2} &= \frac{d}{d\mathbf{d}}\left\{\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\left(\frac{\partial f}{\partial \mathbf{s}}\right)\right\} + \frac{d}{d\mathbf{d}}\left(\frac{\partial f}{\partial \mathbf{d}}\right) \\ &= \frac{d}{d\mathbf{d}}\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\left(\frac{\partial f}{\partial \mathbf{s}}\right) + \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\left\{\frac{d}{d\mathbf{d}}\left(\frac{\partial f}{\partial \mathbf{s}}\right)\right\}^T + \frac{\partial^2 f}{\partial \mathbf{d}^2} + \frac{d\mathbf{s}}{d\mathbf{d}}\frac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}} \\ &= \frac{d^2\mathbf{s}}{d\mathbf{d}^2}\left(\frac{\partial f}{\partial \mathbf{s}}\right) + \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\frac{\partial^2 f}{\partial \mathbf{s}\partial \mathbf{d}} + \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)\frac{\partial^2 f}{\partial \mathbf{s}^2}\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T + \frac{\partial^2 f}{\partial \mathbf{d}^2} + \frac{d\mathbf{s}}{d\mathbf{d}}\frac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}}\end{aligned} \tag{6}$$

The above expression can be re-arranged compactly as shown below.

$$\frac{d^2 z}{d\mathbf{d}^2} = \begin{bmatrix}\mathbf{I} & \frac{d\mathbf{s}}{d\mathbf{d}}\end{bmatrix}\begin{bmatrix}\frac{\partial^2 f}{\partial \mathbf{d}^2} & \frac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}} \\ \frac{\partial^2 f}{\partial \mathbf{s}\partial \mathbf{d}} & \frac{\partial^2 f}{\partial \mathbf{s}^2}\end{bmatrix}\begin{bmatrix}\mathbf{I} \\ \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\end{bmatrix} + \left(\frac{d^2\mathbf{s}}{d\mathbf{d}^2}\right)\left(\frac{\partial f}{\partial \mathbf{s}}\right) \tag{6'}$$

Similarly, the Hessian of the constraints can be written as

$$\frac{d^2 \mathbf{h}}{d\mathbf{d}^2} = \begin{bmatrix}\mathbf{I} & \frac{d\mathbf{s}}{d\mathbf{d}}\end{bmatrix}\begin{bmatrix}\frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}^2} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}\partial \mathbf{s}} \\ \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}\partial \mathbf{d}} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}^2}\end{bmatrix}\begin{bmatrix}\mathbf{I} \\ \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\end{bmatrix} + \left(\frac{d^2\mathbf{s}}{d\mathbf{d}^2}\right)\left(\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right) \tag{7}$$

For feasibility, even this second derivative (Hessian) of the constraints must be zero. This enables us to solve for the only unknown quantity in Eqs. (6') and (7), viz., $\frac{d^2\mathbf{s}}{d\mathbf{d}^2}$.

$$\left(\frac{d^2\mathbf{s}}{d\mathbf{d}^2}\right) = -\begin{bmatrix}\mathbf{I} & \frac{d\mathbf{s}}{d\mathbf{d}}\end{bmatrix}\begin{bmatrix}\frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}^2} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}\partial \mathbf{s}} \\ \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}\partial \mathbf{d}} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}^2}\end{bmatrix}\begin{bmatrix}\mathbf{I} \\ \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\end{bmatrix}\left(\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right)^{-1} \tag{8}$$

By substituting $\frac{d^2\mathbf{s}}{d\mathbf{d}^2}$ from Eq. (8) into Eq. (6'), we get

$$\frac{d^2 z}{d\mathbf{d}^2} = \begin{bmatrix}\mathbf{I} & \frac{d\mathbf{s}}{d\mathbf{d}}\end{bmatrix}\begin{bmatrix}\frac{\partial^2 f}{\partial \mathbf{d}^2} & \frac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}} \\ \frac{\partial^2 f}{\partial \mathbf{s}\partial \mathbf{d}} & \frac{\partial^2 f}{\partial \mathbf{s}^2}\end{bmatrix}\begin{bmatrix}\mathbf{I} \\ \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\end{bmatrix} - \begin{bmatrix}\mathbf{I} & \frac{d\mathbf{s}}{d\mathbf{d}}\end{bmatrix}\begin{bmatrix}\frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}^2} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{d}\partial \mathbf{s}} \\ \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}\partial \mathbf{d}} & \frac{\partial^2 \mathbf{h}}{\partial \mathbf{s}^2}\end{bmatrix}\begin{bmatrix}\mathbf{I} \\ \left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\end{bmatrix}\left(\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right)^{-1}\left(\frac{\partial f}{\partial \mathbf{s}}\right) \tag{9}$$

Noting that

$$-\left(\frac{\partial \mathbf{h}}{\partial \mathbf{s}}\right)^{-1}\left(\frac{\partial f}{\partial \mathbf{s}}\right)=\boldsymbol{\lambda} \quad \text{and} \quad L=f+\boldsymbol{\lambda}^T\mathbf{h} \tag{10}$$

Eq. (9) can be re-written as

$$\frac{d^2 z}{d\mathbf{d}^2}=\begin{bmatrix}\mathbf{I} & \dfrac{d\mathbf{s}}{d\mathbf{d}}\end{bmatrix}\begin{bmatrix}\dfrac{\partial^2 f}{\partial \mathbf{d}^2}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{d}^2}\boldsymbol{\lambda} & \dfrac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{d}\partial \mathbf{s}}\boldsymbol{\lambda} \\ \dfrac{\partial^2 f}{\partial \mathbf{s}\partial \mathbf{d}}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{s}\partial \mathbf{d}}\boldsymbol{\lambda} & \dfrac{\partial^2 f}{\partial \mathbf{s}^2}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{s}^2}\boldsymbol{\lambda}\end{bmatrix}\begin{bmatrix}\mathbf{I} \\ \left(\dfrac{d\mathbf{s}}{d\mathbf{d}}\right)^T\end{bmatrix} \tag{9'}$$

As we noted earlier, for sufficiency of the minimum of z, we need positive definiteness of $\frac{d^2 z}{d\mathbf{d}^2}$, i.e.,

$$\delta\mathbf{d}^T\frac{d^2 z}{d\mathbf{d}^2}\delta\mathbf{d}>0 \quad \text{for any } \delta\mathbf{d} \tag{11}$$

By expanding $\frac{d^2 z}{d\mathbf{d}^2}$ using Eq. (9') and using $\left(\frac{d\mathbf{s}}{d\mathbf{d}}\right)^T\delta\mathbf{d}=\delta\mathbf{s}$ in Eq. (10), we arrive at the following result.

$$\begin{bmatrix}\delta\mathbf{d}^T & \delta\mathbf{s}^T\end{bmatrix}\begin{bmatrix}\dfrac{\partial^2 f}{\partial \mathbf{d}^2}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{d}^2}\boldsymbol{\lambda} & \dfrac{\partial^2 f}{\partial \mathbf{d}\partial \mathbf{s}}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{d}\partial \mathbf{s}}\boldsymbol{\lambda} \\ \dfrac{\partial^2 f}{\partial \mathbf{s}\partial \mathbf{d}}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{s}\partial \mathbf{d}}\boldsymbol{\lambda} & \dfrac{\partial^2 f}{\partial \mathbf{s}^2}+\dfrac{\partial^2 \mathbf{h}}{\partial \mathbf{s}^2}\boldsymbol{\lambda}\end{bmatrix}\begin{bmatrix}\delta\mathbf{d} \\ \delta\mathbf{s}\end{bmatrix}>0 \quad \text{for any } \delta\mathbf{d} \tag{12}$$

From the definition of the Lagrangian (the second of Eq. (10), we can see that Eq. (12) is the same as

$$\begin{bmatrix}\delta\mathbf{d}^T & \delta\mathbf{s}^T\end{bmatrix}\begin{bmatrix}\dfrac{\partial^2 L}{\partial \mathbf{d}^2} & \dfrac{\partial^2 L}{\partial \mathbf{d}\partial \mathbf{s}} \\ \dfrac{\partial^2 L}{\partial \mathbf{s}\partial \mathbf{d}} & \dfrac{\partial^2 L}{\partial \mathbf{s}^2}\end{bmatrix}\begin{bmatrix}\delta\mathbf{d} \\ \delta\mathbf{s}\end{bmatrix}>0 \quad \text{for any } \delta\mathbf{d} \tag{13}$$

Noting that $\delta\mathbf{x}^T=\begin{bmatrix}\delta\mathbf{d}^T & \delta\mathbf{s}^T\end{bmatrix}$ and that the dependent (or solution) variable part has to saitisfy the feasibility condition, we can re-write Eq. (14) in an enhanced form as

$$\delta\mathbf{x}^T\begin{bmatrix}\dfrac{\partial^2 L}{\partial \mathbf{d}^2} & \dfrac{\partial^2 L}{\partial \mathbf{d}\partial \mathbf{s}} \\ \dfrac{\partial^2 L}{\partial \mathbf{s}\partial \mathbf{d}} & \dfrac{\partial^2 L}{\partial \mathbf{s}^2}\end{bmatrix}\delta\mathbf{x}>0 \quad \text{for any } \delta\mathbf{x} \text{ satisfying } \nabla\mathbf{h}^T\delta\mathbf{x}=\mathbf{0} \tag{14}$$

It is important to note that the above condition is less demanding on requiring the Hessian of the Lagrangian to be positive definite.