Gradient and Hessian of the SPARROW Function

Christian Steffens

Consider the objective function

$$f(\boldsymbol{s}) = \mathrm{Tr}(\boldsymbol{Q}^{-1}\hat{\boldsymbol{R}}) + \boldsymbol{1}^{\mathsf{T}}\boldsymbol{s}, \tag{1}$$

where $\boldsymbol{Q} = \boldsymbol{A}\,\mathrm{diag}(\boldsymbol{s})\boldsymbol{A}^{\mathsf{H}} + \lambda\boldsymbol{I} \in \mathbb{C}^{M\times M}$, with $\boldsymbol{s} = [s_1, \dots, s_K]^{\mathsf{T}} \in \mathbb{R}_+^K$, $\boldsymbol{A} = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_K] \in \mathbb{C}^{M\times K}$ and $\boldsymbol{1}$ is a vector of ones. Using the elementwise derivatives

$$\frac{\partial \boldsymbol{Q}}{\partial s_k} = \frac{\partial}{\partial s_k}\sum_{i=1}^{K} s_i\boldsymbol{a}_i\boldsymbol{a}_i^{\mathsf{H}} + \lambda\boldsymbol{I} = \boldsymbol{a}_k\boldsymbol{a}_k^{\mathsf{H}} \tag{2}$$

$$\frac{\partial \boldsymbol{Q}^{-1}}{\partial s_k} = -\,\boldsymbol{Q}^{-1}\frac{\partial \boldsymbol{Q}}{\partial s_k}\boldsymbol{Q}^{-1} = -\boldsymbol{Q}^{-1}\boldsymbol{a}_k\boldsymbol{a}_k^{\mathsf{H}}\boldsymbol{Q}^{-1} \tag{3}$$

on the function (1), we obtain the elementwise derivatives

$$\frac{\partial f(\boldsymbol{s})}{\partial s_k} = 1 - \mathrm{Tr}(\boldsymbol{Q}^{-1}\boldsymbol{a}_k\boldsymbol{a}_k^{\mathsf{H}}\boldsymbol{Q}^{-1}\hat{\boldsymbol{R}}) = 1 - \boldsymbol{a}_k^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{Q}^{-1}\boldsymbol{a}_k, \tag{4}$$

which are summarized as the gradient

$$\frac{\partial f(\boldsymbol{s})}{\partial \boldsymbol{s}} = \boldsymbol{1} - \mathrm{vecd}\big(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{Q}^{-1}\boldsymbol{A}\big), \tag{5}$$

where $\mathrm{vecd}(\boldsymbol{X})$ denotes the vector containing the elements on the main diagonal of matrix $\boldsymbol{X}$. Using the product rule and (3), the elementwise second order derivative of (4) is given as

$$\begin{aligned}
\frac{\partial^2 f(\boldsymbol{s})}{\partial s_k \partial s_l} &= \boldsymbol{a}_k^{\mathsf{H}}\frac{\partial \boldsymbol{Q}^{-1}}{\partial s_l}\boldsymbol{R}\boldsymbol{Q}^{-1}\boldsymbol{a}_k + \boldsymbol{a}_k^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{R}\frac{\partial \boldsymbol{Q}^{-1}}{\partial s_l}\boldsymbol{a}_k \\
&= \boldsymbol{a}_k^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{a}_l\boldsymbol{a}_l^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{Q}^{-1}\boldsymbol{a}_k + \boldsymbol{a}_k^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{Q}^{-1}\boldsymbol{a}_l\boldsymbol{a}_l^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{a}_k \\
&= 2\mathrm{Re}\Big\{(\boldsymbol{a}_k^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{a}_l)\cdot(\boldsymbol{a}_l^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{Q}^{-1}\boldsymbol{a}_k)\Big\}
\end{aligned} \tag{6}$$

which can be written in compact matrix notation as

$$\frac{\partial^2 f(\boldsymbol{s})}{\partial \boldsymbol{s}\,\partial \boldsymbol{s}^{\mathsf{T}}} = 2\,\mathrm{Re}\Big\{(\boldsymbol{A}^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{A})^{\mathsf{T}} \odot (\boldsymbol{A}^{\mathsf{H}}\boldsymbol{Q}^{-1}\boldsymbol{R}\boldsymbol{Q}^{-1}\boldsymbol{A})\Big\}, \tag{7}$$

forming the Hessian matrix of (1), with $\odot$ denoting the Hadamard product, i.e., elementwise multiplication. From the Schur product theorem it can be concluded that the Hessian matrix in (7) is positive semidefinite, since for $s_1, \dots, s_K \geq 0$ it holds that $\boldsymbol{Q} \succeq \boldsymbol{0}$. In other words, the SPARROW formulation in (1) is convex for nonnegative $s_1, \dots, s_K \geq 0$.