Gradient-based ICA-CPA

Lieven De Lathauwer, Maarten De Vos

February 1, 2012

Recently, we proposed a combination of Independent Component Analysis and Parallel Factor Analysis, which we called ICA-CPA. The computation was based on an ELSCS optimization framework. An alternative way of updating the costfunction (18) can be based on gradient descent. In this technical report, we provide the formulas explicitly:
For the cost function ϕ in (18) we have

$$\begin{aligned}\phi(\mathbf{A}+\mathbf{\Delta A},\mathbf{B}+\mathbf{\Delta B},\mathbf{D}+\mathbf{\Delta D}) = \phi(\mathbf{A},\mathbf{B},\mathbf{D}) &+ < \mathbf{A}^*, \nabla_{\mathbf{A}}\phi > + < \mathbf{A}, \nabla_{\mathbf{A}^*}\phi > \\ &+ < \mathbf{B}^*, \nabla_{\mathbf{B}}\phi > + < \mathbf{B}, \nabla_{\mathbf{B}^*}\phi > + < \mathbf{C}^*, \nabla_{\mathbf{C}}\phi > + < \mathbf{C}, \nabla_{\mathbf{C}^*}\phi > \end{aligned} \quad (1)$$

up to first order terms, in which $\nabla_{\mathbf{A}^*}\phi$, $\nabla_{\mathbf{B}^*}\phi$, $\nabla_{\mathbf{C}^*}\phi$ are the partial complex gradients of ϕ [**?**]. Due to the symmetry of ϕ, we have that

$$\nabla_{\mathbf{A}}\phi = (\nabla_{\mathbf{A}^*}\phi)^*, \nabla_{\mathbf{B}}\phi = (\nabla_{\mathbf{B}^*}\phi)^* \text{and} \nabla_{\mathbf{C}}\phi = (\nabla_{\mathbf{C}^*}\phi)^*.$$

Computation of the expansion (1) yields:

$$\begin{aligned}\nabla_{\mathbf{A}^*}\phi &= \mathbf{A}\cdot(\mathbf{B}^*\odot\mathbf{A}\odot\mathbf{B}\odot\mathbf{C}^*)^H\cdot(\mathbf{B}^*\odot\mathbf{A}\odot\mathbf{B}\odot\mathbf{C}^*) + \mathbf{A}\cdot(\mathbf{A}\odot\mathbf{B}\odot\mathbf{B}^*\odot\mathbf{C})^H\cdot(\mathbf{A}\odot\mathbf{B}\odot\mathbf{B}^*\odot\mathbf{C}) \\ &\quad -\mathbf{U}_{[1;2,3,4,5]}\cdot(\mathbf{B}^*\odot\mathbf{A}\odot\mathbf{B}\odot\mathbf{C}^*) - (\mathbf{U}_{[3;1,2,4,5]})^*\cdot(\mathbf{A}\odot\mathbf{B}\odot\mathbf{B}^*\odot\mathbf{C}) \quad (2)\\ \nabla_{\mathbf{B}^*}\phi &= \mathbf{B}\cdot(\mathbf{A}^*\odot\mathbf{A}\odot\mathbf{B}\odot\mathbf{C}^*)^H\cdot(\mathbf{A}^*\odot\mathbf{A}\odot\mathbf{B}\odot\mathbf{C}^*) + \mathbf{B}\cdot(\mathbf{A}\odot\mathbf{B}\odot\mathbf{A}^*\odot\mathbf{C})^H\cdot(\mathbf{A}\odot\mathbf{B}\odot\mathbf{A}^*\odot\mathbf{C}) \\ &\quad -\mathbf{U}_{[2;1,3,4,5]}\cdot(\mathbf{A}^*\odot\mathbf{A}\odot\mathbf{B}\odot\mathbf{C}^*) - (\mathbf{U}_{[4;1,2,3,5]})^*\cdot(\mathbf{A}\odot\mathbf{B}\odot\mathbf{A}^*\odot\mathbf{C}) \quad (3)\\ \nabla_{\mathbf{C}^*}\phi &= \mathbf{C}\cdot(\mathbf{A}^*\odot\mathbf{B}^*\odot\mathbf{A}\odot\mathbf{B})^H\cdot(\mathbf{A}^*\odot\mathbf{B}^*\odot\mathbf{A}\odot\mathbf{B}) \\ &\quad -\mathbf{U}_{[5;1,2,3,4]}\cdot(\mathbf{A}^*\odot\mathbf{B}^*\odot\mathbf{A}\odot\mathbf{B}). \quad (4)\end{aligned}$$

Expressions (2)–(4) can be efficiently computed as follows:

$$\begin{aligned}\nabla_{\mathbf{A}^*}\phi &= 2\mathbf{A}\cdot[(\mathbf{A}^H\cdot\mathbf{A})*(\mathbf{B}^H\cdot\mathbf{B})*(\mathbf{B}^H\cdot\mathbf{B})*(\mathbf{C}^H\cdot\mathbf{C})] \\ &\quad -\mathcal{U}\cdot_2\mathbf{B}^H\cdot_3\mathbf{A}^T\cdot_4\mathbf{B}^T\cdot_5\mathbf{C}^H - (\mathcal{U}_{[3,2,1,4,5]})^*\cdot_2\mathbf{B}^T\cdot_3\mathbf{A}^T\cdot_4\mathbf{B}^H\cdot_5\mathbf{C}^T \\ \nabla_{\mathbf{B}^*}\phi &= 2\mathbf{B}\cdot[(\mathbf{A}^H\cdot\mathbf{A})*(\mathbf{A}^H\cdot\mathbf{A})*(\mathbf{B}^H\cdot\mathbf{B})*(\mathbf{C}^H\cdot\mathbf{C})] \\ &\quad -(\mathcal{U}_{[2,1,3,4,5]})\cdot_2\mathbf{A}^H\cdot_3\mathbf{A}^T\cdot_4\mathbf{B}^T\cdot_5\mathbf{C}^H - (\mathcal{U}_{[4,2,3,1,5]})^*\cdot_2\mathbf{B}^T\cdot_3\mathbf{A}^H\cdot_4\mathbf{A}^T\cdot_5\mathbf{C}^T \\ \nabla_{\mathbf{C}^*}\phi &= \mathbf{C}\cdot[(\mathbf{A}^H\cdot\mathbf{A})*(\mathbf{A}^H\cdot\mathbf{A})*(\mathbf{B}^H\cdot\mathbf{B})*(\mathbf{B}^H\cdot\mathbf{B})] \\ &\quad -(\mathcal{U}_{[5,2,3,4,1]})\cdot_2\mathbf{B}^H\cdot_3\mathbf{A}^T\cdot_4\mathbf{B}^T\cdot_5\mathbf{A}^H\end{aligned}$$

References

[1] A. Hjorungnes and D. Gesbert. Complex-Valued Matrix Differentiation: Techniques and Key Results. *IEEE Trans. Signal Process.*, 55:2740–2746, 2007.