Supplementary Material
Analytic expressions for $\nabla h_X(M)$

Abstract

This document contains the main technical arguments for the computation of the gradient expressions $\nabla h_X(M)$ introduced as Propositions 1–8 in the main article. Further details on similar derivations can be found in (Couillet et al., 2018).

1. Derivation of the gradient for the examples provided in the article

Our main concern is to find an analytical expression of the gradient defined as:

$$\nabla h_X(M) = \frac{-\hat{\delta}(M,X)}{\pi \imath p} \oint_{\hat{\Gamma}} g\left(-m_{\tilde{\mu}_p}(z;M)\right) \operatorname{sym}\left(\hat{C}(M^{-1}\hat{C} - zI_p)^{-2}\right) dz. \tag{1}$$

Equation (1) can be written as:

$$\nabla h_X(M) = \frac{-\hat{\delta}(M,X)}{\pi \imath p} \operatorname{sym}\left(\hat{C}U\left(\oint_{\hat{\Gamma}} g\left(-m_{\tilde{\mu}_p}(z;M)\right)(\Lambda - zI_p)^{-2} dz\right) U^{\mathsf{T}}\right). \tag{2}$$

where $M^{-1}\hat{C} = U\Lambda U^{\mathsf{T}}$ in its spectral decomposition. Our main focus is on the diagonal matrix

$$A \equiv \frac{1}{2\imath\pi} \oint_{\hat{\Gamma}} g\left(-m_{\tilde{\mu}_p}(z;M)\right)(\Lambda - zI_p)^{-2} dz$$

and particularly on its k-th diagonal element

$$A_{kk} = \frac{1}{2\imath\pi} \oint_{\hat{\Gamma}} \frac{g\left(-m_{\tilde{\mu}_p}(z;M)\right)}{(\lambda_k - z)^2} dz \tag{3}$$

with $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$.

To solve (3), we use elementary properties of the rational function $m_{\tilde{\mu}_p}(z;M)$ that we recall is defined as

$$m_{\tilde{\mu}_p}(z;M) = \frac{c}{p} \sum_{i=1}^{p} \frac{1}{\lambda_i - z} + \frac{1-c}{z}.$$

Remarking that the poles of $m_{\tilde{\mu}_p}(z;M)$ are $\{\lambda_i\}_{i=1}^p$ and 0, and that $\{\xi_i\}_{i=1}^p$ are the zeros of $m_{\tilde{\mu}_p}(z;M)$ (see (Couillet et al., 2018) for details), we have :

$$m_{\tilde{\mu}_p}(z;M) = \frac{\prod_{i=1}^p (z - \xi_i)}{z \prod_{i=1}^p (z - \lambda_i)}.$$

With these ingredients, we can evaluate A_{kk} for various functions f (recall that $g(z) = f(1/z)$).

1.1. Case $f(t) = t$

For this case, $g(t) = \frac{1}{t}$, and thus the integrand of A_{kk} is

$$I = \frac{z \prod_{\substack{i=1 \\ i \neq k}}^{p} (z - \lambda_i)}{(z - \lambda_k) \prod_{i=1}^{p} (z - \xi_i)}.$$

Under this rational form, A_{kk} is easy to evaluate since it only requires to evaluate the residue for each pole of I:

- the first order pole λ_k for which the residue R_1 is given by

$$\begin{aligned} R_1 &= \lim_{z \to \lambda_k} \frac{z \prod_{\substack{i=1 \\ i \neq k}}^{p} (z - \lambda_i)}{\prod_{i=1}^{p} (z - \xi_i)} \\ &= \lim_{z \to \lambda_k} \frac{1}{(z - \lambda_k) m_{\tilde{\mu}_p}(z)} \\ &= -\frac{p}{c} \end{aligned}$$

- and the first order poles ξ_j, $j \in \{1, \ldots, p\}$ for which the residue R_2 is given by:

$$\begin{aligned} R_2 &= \sum_{j=1}^{p} \lim_{z \to \xi_j} \frac{z \prod_{i=1}^{p} (z - \lambda_i)}{(z - \lambda_k)^2 \prod_{\substack{i=1 \\ i \neq j}}^{p} (z - \xi_i)} \\ &= \sum_{j=1}^{p} \frac{1}{(\xi_j - \lambda_k)^2} \lim_{z \to \xi_j} \frac{z - \xi_j}{m_{\tilde{\mu}_p}(z)} \\ &= \sum_{j=1}^{p} \frac{1}{(\xi_j - \lambda_k)^2 m'_{\tilde{\mu}_p}(\xi_j)} \end{aligned}$$

  Putting the $p + 1$ residues together then yields:

$$A_{kk} = -\frac{p}{c} + \sum_{j=1}^{p} \frac{1}{(\xi_j - \lambda_k)^2 m'_{\tilde{\mu}_p}(\xi_j)}.$$

1.2. Case $f(t) = \log(t)$

For this case, $g(t) = -\log(t)$ and therefore the integrand of A_{kk} becomes

$$I = -\frac{\log \left(\frac{\prod_{i=1}^{p} (z - \xi_i)}{z \prod_{i=1}^{p} (z - \lambda_i)} \right)}{(\lambda_k - z)^2}$$

Elementary functional analysis allows us to find the discontinuities of this multi-valued function (the z's for which the argument of the logarithm function is negative). This set of points, or branch cuts, are exactly the segments $[\xi_i, \lambda_i]$, $i = 1, \ldots, p$. These segments lie inside the integration contour Γ, that needs be modified for proper integration; the new contour, denoted Γ_n is depicted in Figure 1.

Under Γ_n, A_{kk} is the sum of several integrals, subdivided in four types:

- integrals over the circles surrounding $\{\xi_j\}_{j=1}^{p}$ which, thanks to the variable change $z = \xi_j + \epsilon \, \mathrm{e}^{\imath \theta}$, reduce to

$$\lim_{\epsilon \to 0} -\int_{\epsilon}^{2\pi - \epsilon} \frac{\log \left(\epsilon \, \mathrm{e}^{\imath \theta} \frac{\prod_{\substack{i=1 \\ i \neq j}}^{p} (\xi_j + \epsilon \, \mathrm{e}^{\imath \theta} - \xi_i)}{(\xi_j + \epsilon \, \mathrm{e}^{\imath \theta}) \prod_{i=1}^{p} (\xi_j + \epsilon \, \mathrm{e}^{\imath \theta} - \lambda_i)} \right)}{(\lambda_k - \xi_j - \epsilon \, \mathrm{e}^{\imath \theta})^2} \imath \epsilon \, \mathrm{e}^{\imath \theta} = 0$$

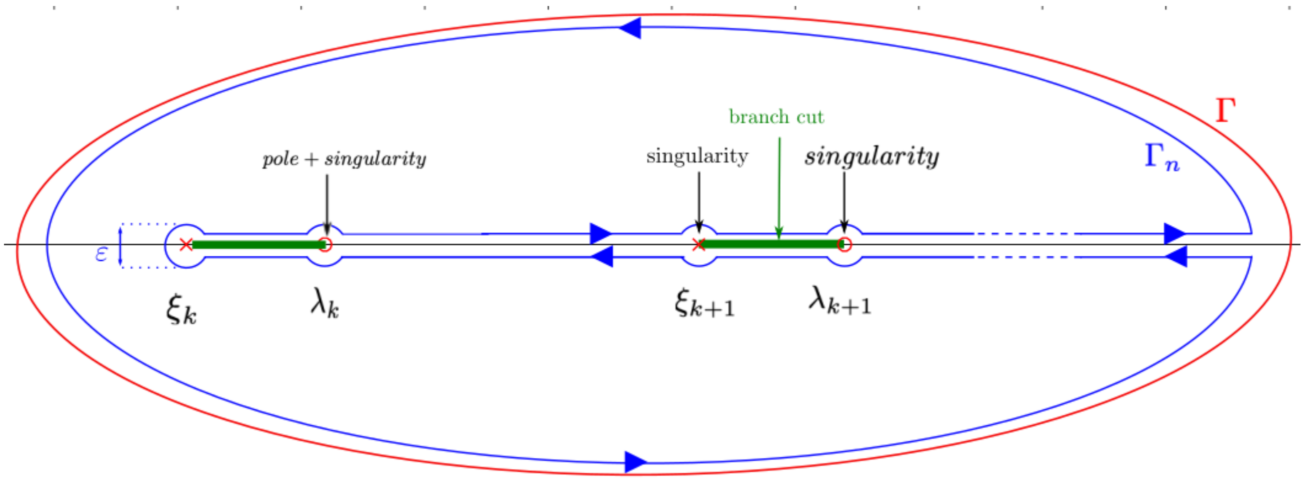

Figure 1. Contour deformation

- integrals over the circles surrounding $\{\lambda_i\}_{\substack{i=1\\i\neq k}}^{p}$ which are null following the same line of reasoning as for ξ_i.
- real integrals over the segments $[\xi_i, \lambda_i]$ which can be computed by remarking that the $\log$ function has a discontinuity of $2\imath\pi$ at the branch cut.

$$\begin{aligned}
&\frac{1}{2\imath\pi}\sum_{j=1}^{p}\int_{\xi_j+\epsilon}^{\lambda_j-\epsilon}\frac{\log\left(|m_{\tilde{\mu}_p}|\right)+\imath\pi-\log\left(|m_{\tilde{\mu}_p}|\right)+\imath\pi}{(\lambda_k-z)^2}\\
&=\sum_{j=1}^{p}\int_{\xi_j+\epsilon}^{\lambda_j-\epsilon}\frac{1}{(\lambda_k-z)^2}\\
&=\sum_{j=1}^{p}\lim_{z\to\lambda_j}\frac{1}{(\lambda_k-z)}-\frac{1}{(\lambda_k-\xi_j)}\\
&=\sum_{\substack{j=1\\j\neq k}}^{p}\frac{1}{(\lambda_k-\lambda_j)}-\sum_{j=1}^{p}\frac{1}{(\lambda_k-\xi_j)}+\lim_{z\to\lambda_k}\frac{1}{(\lambda_k-z)}
\end{aligned}$$

- the integral over the circle surrounding λ_k computed by remarking that λ_k is a second order pole

$$\begin{aligned}
&\lim_{z\to\lambda_k}\frac{\partial}{\partial z}\left(\log\left(-m_{\tilde{\mu}_p}(z;M)\right)\right)\\
&=\lim_{z\to\lambda_k}\sum_{j=1}^{p}\frac{1}{z-\xi_j}-\frac{1}{z}-\sum_{j=1}^{p}\frac{1}{z-\lambda_j}\\
&=\sum_{j=1}^{p}\frac{1}{\lambda_k-\xi_j}-\frac{1}{\lambda_k}-\sum_{\substack{j=1\\j\neq k}}^{p}\frac{1}{\lambda_k-\lambda_j}-\lim_{z\to\lambda_k}\frac{1}{z-\lambda_k}
\end{aligned}$$

where the second line is obtained remarking that:

$$\frac{m'_{\tilde{\mu}_p}(z;M)}{m_{\tilde{\mu}_p}(z;M)}=\sum_{j=1}^{p}\frac{1}{z-\xi_j}-\frac{1}{z}-\sum_{j=1}^{p}\frac{1}{z-\lambda_j}.$$

Combining these integrals then yields to the solution of the integral:

$$A_{kk} = -\frac{1}{\lambda_k}.$$

1.3. Case $f(t) = \log(1 + st)$

For this case, the integrand of A_{kk} can be derived similarly as in the case of the logarithm by noting that the argument of the logarithm $(1 - s/m_{\tilde{\mu}_p}(z))$ is a polynomial for which the poles are λ_i and 0. The zeros are in number $p + 1$ and denoted κ_i, $i = 0, \ldots, p$ with $\kappa_0 < 0 < \kappa_1 < \ldots < \kappa_p$; in particular, only $\kappa_1, \ldots, \kappa_p$ are inside the integration contour (see (Couillet et al., 2018) for details). Therefore, the integrand is written similarly as for the log function as:

$$I = -\frac{\log\left(\frac{\prod_{i=0}^{p}(z-\kappa_i)}{z\prod_{i=1}^{p}(z-\lambda_i)}\right)}{(\lambda_k - z)^2}.$$

The integration contour can be deformed as for the $\log$ function. Using similar integration techniques, the calculus then yields to the solution derived in the article.

1.4. Case $f(t) = \log^2(t)$

Here $g(t) = f(t) = \log^2(t)$ and the integrand for this case is simply

$$I = -\frac{\log^2\left(\frac{\prod_{i=1}^{p}(z-\xi_i)}{z\prod_{i=1}^{p}(z-\lambda_i)}\right)}{(\lambda_k - z)^2}.$$

Again, we use here exactly the same line of work performed on the $\log(t)$ and $\log(1 + st)$ functions. Technical difficulties however arise when addressing the real integrals which involve products of logarithms and rational functions. These difficulties are mostly cumbersome calculus which are addressed similar to (Couillet et al., 2018).

The treatment of the complex integrals resulting from the estimation of C^{-1} is performed similarly.

References

Couillet, R., Tiomoko, M., Zozor, S., and Moisan, E. Random matrix-improved estimation of covariance matrix distances. *arXiv preprint arXiv:1810.04534*, 2018.