Code name: $\boxed{\textbf{SD}}$
Argument type: PSD matrix

1 Stein divergence

Stein loss introduced in 1961[1] is a square matrix symmetric divergence. Let D denote the number of rows/colums of positive definite matrices $\Sigma_1 \succ 0$ and $\Sigma_2 \succ 0$.

Then Stein loss[2] also called S-divergence[3] or Jensen-Bregman LogDet[4] is defined by:

$$\mathrm{S}\left(\Sigma_1;\Sigma_2\right) = \mathrm{tr}(\Sigma_1^{-1}\Sigma_2) - \log\det(\Sigma_1^{-1}\Sigma_2) - D = \mathrm{S}\left(\Sigma_2;\Sigma_1\right) \tag{1}$$

It is a matrix Jensen-Bregman divergence for the convex matrix generator $F(X) = -\log\det(X)$, hence its name Jensen-Bregman LogDet divergence[5].

2 Properties

- The square root $\sqrt{\mathrm{S}\left(\Sigma_1;\Sigma_2\right)}$ of Stein divergence is a metric.
- Stein divergence is invariant to affine transformations: $\mathrm{S}\left(A\Sigma_1 B; A\Sigma_2 B\right) = \mathrm{S}\left(\Sigma_1;\Sigma_2\right)$, for $A, B \in \mathrm{GL}(D)$.
- Stein divergence is invariant to inverse transformations: $\mathrm{S}\left(\Sigma_1^{-1};\Sigma_2^{-1}\right) = \mathrm{S}\left(\Sigma_1;\Sigma_2\right)$.
- Stein divergence may not be convex[6]
- The matrix gradient $\nabla_X \mathrm{S}\left(X;Y\right) = (X+Y)^{-1} - \frac{1}{2}X^{-1}$

[1] William James and Charles Stein. "Estimation with quadratic loss". In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability.* Vol. 1. 1961. 1961, pp. 361–379.

[2] James and Stein, "Estimation with quadratic loss".

[3] Suvrit Sra. "Positive definite matrices and the S-divergence". In: *arXiv preprint arXiv:1110.1773* (2011).

[4] Anoop Cherian et al. "Jensen-Bregman LogDet divergence with application to efficient similarity search for covariance matrices". In: *IEEE transactions on pattern analysis and machine intelligence* 35.9 (2013), pp. 2161–2174.

[5] Cherian et al., "Jensen-Bregman LogDet divergence with application to efficient similarity search for covariance matrices".

[6] Frank Nielsen and Richard Nock. "Skew Jensen-Bregman Voronoi diagrams". In: *Transactions on Computational Science XIV*. Springer, 2011, pp. 102–128; Cherian et al., "Jensen-Bregman LogDet divergence with application to efficient similarity search for covariance matrices".