# Derivation of an inequality of logdet

$\log\det(A^TA+\delta I) \le \log\det(B^TB+\delta I) + \mathrm{Tr}\big((B^TB+\delta I)^{-1}A^TA\big) +$ constant

Andersen Ang

Mathématique et de Recherche opérationnelle
Faculté polytechnique de Mons
UMONS
Mons, Belgium

*email: manshun.ang@umons.ac.be*
homepage: *angms.science*

First draft : January 30, 2018
Current draft : February 3, 2018

# An inequality of logdet of Gram matrix

Given a matrix $A \in \mathbb{R}^{m \times n}$, a fix small constant $\delta > 0$, we have the following inequality

$$\log\det(A^T A + \delta I) \le \log\det(B^T B + \delta I) + \mathsf{Tr}\big((B^T B + \delta I)^{-1} A^T A\big) + \text{constant},$$

for any matrix $B \in \mathbb{R}^{m \times n}$.

We are going to derive this inequality.

# The proof of the inequality of logdet of Gramian

Given a matrix $A \in \mathbb{R}^{m \times n}$, a fix small constant $\delta > 0$, the matrix $A^T A$ is the Gramian of $A$. Gramian is always symmetric positive semidefinite regardless of $A$.

Gramian $A^T A$ has non-negative eigenvalues. Adding $\delta I$ to the $A^T A$ gives a positive definite matrix $A^T A + \delta I$ with positive eigenvalues.

The reason why adding $\delta I$ to $A^T A$ : let $Q = A^T A + \delta I$. As determinant of a matrix is the product of eigenvalues of that matrix, we have $\det Q = \prod \lambda_i$. So $\log \det Q = \sum \log \lambda_i$, the sum of log of eigenvalue of $Q$. As log only works for positive values, so it is why $\delta I$ is added to $A^T A$.

It can be shown that, the matrix function $\log \det Q$ with symmetric positive definite matrix $Q$, is a concave function with respect to $Q$ (not with respect to $A$ !). For the proof, see Page 8 of this slide.

# The proof of the inequality of logdet of Gram matrix

As $f(Q) = \log\det Q$ is a concave function, we can upper bound it by its first order Taylor approximation :

$$f(Q) \leq f(S) + \langle \nabla f(S), Q - S \rangle ,$$

where $\nabla f(Q)$ is $Q^{-T} = \left(Q^{-1}\right)^T$ and $\langle \, , \, \rangle$ is matrix inner product with the expression $\langle X, Y \rangle = \mathsf{Tr}(X^{-T}Y)$. So we have

$$\begin{aligned} \log\det Q &\leq \log\det S + \mathsf{Tr}\big((S^{-T})^T(Q - S)\big) \\ &= \log\det S + \mathsf{Tr}\big(S^{-1}(Q - S)\big) \end{aligned}$$

Distribute $S^{-1}$ to $Q - S$ :

$$\begin{aligned} \log\det Q &\leq \log\det S + \mathsf{Tr}\big(S^{-1}Q - I_n\big) \\ &= \log\det S + \mathsf{Tr}\big(S^{-1}Q\big) - n \end{aligned}$$

Put $Q = A^TA + \delta I$ and $S = B^TB + \delta I$, we have

$$\log\det(A^TA + \delta I) \leq \log\det(B^TB + \delta I) + \mathsf{Tr}\big((B^TB + \delta I)^{-1}(A^TA + \delta I)\big) - n$$

# The proof of the inequality of logdet of Gram matrix

Focus on the term $\text{Tr}\big((B^TB+\delta I)^{-1}(A^TA+\delta I)\big)$

$$\text{Tr}\big((B^TB+\delta I)^{-1}(A^TA+\delta I)\big) = \text{Tr}\big((B^TB+\delta I)^{-1}A^TA\big)+\text{Tr}\big((B^TB+\delta I)^{-1}\delta I\big)$$

Assume the SVD of the matrix $B$ is $U\Sigma V^T$, then $B^TB = V\Sigma^2V^T$. Based on the property of SVD, we have $V^T = V^{-1}$ and

$$\begin{aligned}
\text{Tr}\big((B^TB+\delta I)^{-1}\delta I\big) &= \delta\text{Tr}\big((V\Sigma^2V^T+\delta I_n)^{-1}I\big) \\
&= \delta\text{Tr}\big((V\Sigma^2V^T+\delta I_n)^{-1}\big) \\
(I_n = VV^T) \quad &= \delta\text{Tr}\big((V\Sigma^2V^T+\delta VV^T)^{-1}\big) \\
\text{(factorization)} \quad &= \delta\text{Tr}\big((V(\Sigma^2+\delta I_n)V^T)^{-1}\big) \\
\text{(take inverse)} \quad &= \delta\text{Tr}\big(V(\Sigma^2+\delta I_n)^{-1}V^T\big) \\
\text{(trace is invariant to } V) \quad &= \delta\text{Tr}\big((\Sigma^2+\delta I_n)^{-1}\big) \\
&= \sum_{i=1}^{n}\frac{\delta}{\sigma_i^2+\delta}
\end{aligned}$$

Note. Recall that $(\Sigma^2+\delta I_n)^{-1} \neq (\Sigma^{-2}+\delta^{-1}I_n)$.

Note2. As $\delta > 0$ and $\sigma_i \geq 0$, $\dfrac{\delta}{\sigma_i^2+\delta} = \dfrac{1}{\frac{\sigma_i^2}{\delta}+1} \leq 1 \implies \sum_{i=1}^{n}\dfrac{\delta}{\sigma_i^2+\delta} \leq n$

# The proof of the inequality of logdet of Gram matrix

So, we have

$$\log\det(A^TA+\delta I) \le \log\det(B^TB+\delta I) + \mathsf{Tr}\big((B^TB+\delta I)^{-1}A^TA\big) + \sum_{i=1}^{n}\frac{\delta}{\sigma_i^2+\delta} - n$$

Note : if $A = B$ and SVD of the matrix $B$ is $U\Sigma V^T$,

$$\mathsf{Tr}\big((B^TB+\delta I)^{-1}B^TB\big) = \mathsf{Tr}\big((\Sigma^2+\delta I_n)^{-1}\Sigma^2\big) = \sum_{i=1}^{n}\frac{\sigma_i^2}{\sigma_i^2+\delta}$$

Thus

$$\mathsf{Tr}\big((B^TB+\delta I)^{-1}A^TA\big) + \sum_{i=1}^{n}\frac{\delta}{\sigma_i^2+\delta} - n = 0$$

and the equality is established.

# The inequality as a convexification of logdet of Gram

The inequality we have

$$\log\det(A^TA+\delta I) \leq \log\det(B^TB+\delta I)+\mathsf{Tr}\big((B^TB+\delta I)^{-1}A^TA\big)+\sum_{i=1}^{n}\frac{\delta}{\sigma_i^2+\delta}-n$$

If we ignoring all the terms that are independent of $A$, we have

$$\log\det(A^TA+\delta I) \leq \mathsf{Tr}\big((B^TB+\delta I)^{-1}A^TA\big) + \text{constants}$$

Let $(B^TB+\delta I)^{-1} = D$, we have

$$\begin{aligned}\log\det(A^TA+\delta I) &\leq \mathsf{Tr}\big(DA^TA\big) + \text{constants} \\ &= \|A\|_D^2 + \text{constants}\end{aligned}$$

That is, logdet is upper bounded by a weighted F-norm. As norm is convex with respect to matrix $A$ while logdet is not[1], the inequality can be used as a **convex relaxation** of logdet.

[1] logdet of a symmetric positive definite matrix is concave, but logdet of $A^TA+\delta I$ w.r.t. $A$ is not concave nor convex

# Last page - summary

Given a matrix $A \in \mathbb{R}^{m\times n}$, a fix small constant $\delta > 0$, we have the following inequality for any matrix $B \in \mathbb{R}^{m\times n}$ :

$$\log\det(A^TA + \delta I) \leq \log\det(B^TB + \delta I) + \mathsf{Tr}\big((B^TB + \delta I)^{-1}A^TA\big) + \mathsf{c},$$

Let $\sigma_i$ be the singular value of $B$, the constant $c$ is

$$c = \sum_{i=1}^{n} \frac{\delta}{\sigma_i^2 + \delta} - n \leq 0$$

Equality is established when $B = A$.

The inequality is a convex relaxation of $\log\det$ of Grammian.

End of document