Column-wise representation of of Det of Gramian

i.e. showing $\det \mathbf{B}^\top \mathbf{B} = \gamma \mathbf{b}_i^\top \mathbf{X} \mathbf{b}_i$

Andersen Ang

Mathématique et recherche opérationnelle
UMONS, Belgium

manshun.ang@umons.ac.be Homepage: angms.science

First draft : July 26, 2017
Last update : July 25, 2019

Overview

The Gram matrix

Given a matix $\mathbf{B} \in \mathbb{R}^{m \times n}$, consider the Gramian $\mathbf{G} = \mathbf{B}^\top \mathbf{B}$:

- $\mathbf{G}$ is positive semi-definite.
  Proof: $\mathbf{x}^\top \mathbf{G} \mathbf{x} = \mathbf{x}^\top \mathbf{B}^\top \mathbf{B} \mathbf{x} = \|\mathbf{B}\mathbf{x}\|_2^2 \geq 0, \forall \mathbf{x}$.
- $\mathbf{G}$ is positive definite if $\mathbf{B}$ full rank / has linearly independent columns.
- $\mathbf{G}$ stores all the pair-wise information between columns of $\mathbf{B}$: the norm of them (on the diagnoal of $\mathbf{G}$) and the angle between them (off the diagonal of $\mathbf{G}$).

Consider $\det \mathbf{B}^\top \mathbf{B}$.

- It is the square of the volume of the parallelotope formed by the column vectors in $\mathbf{B}$.
- Determinant of a matrix equals to the product of the eigenvalues of that matrix, we have $\det \mathbf{B}^\top \mathbf{B} = \prod \lambda_i(\mathbf{B}^\top \mathbf{B})$.
- As singular value of $\mathbf{B}$ is the square root of the eigenvalue of $\mathbf{B}^\top \mathbf{B}$, we have $\det \mathbf{B}^\top \mathbf{B} = \prod \sigma_i^2(\mathbf{B})$.

Given $\mathbf{B} = [\, \mathbf{b}_1,\ \mathbf{b}_2,\ \dots\ \mathbf{b}_n \,] \in \mathbb{R}^{m \times n}$. Suppose we want to express $\det \mathbf{B}^\top \mathbf{B}$ as a function of the i^{th} column $\mathbf{b}_i$ as

$$\det \mathbf{B}^\top \mathbf{B} = f(\mathbf{b}_i) = \text{certain expression involving } \mathbf{b}_i.$$

Here is how : consider expressing matrix $\mathbf{B} = [\, \mathbf{b}_1, \mathbf{b}_2, \dots \mathbf{b}_n \,]$ as $[\, \mathbf{b}_i, \mathbf{b}_1, \dots, \mathbf{b}_{i-1}, \mathbf{b}_{i+1}, \dots, \mathbf{b}_n \,]$ using column permutation matrix Π_i. i.e. the i^{th} column $\mathbf{b}_i$ of $\mathbf{B}$ is moved to the left-most column by multiplying a permulation matrix Π_i from the right with $\mathbf{B}$:

$$\mathbf{B} = [\, \mathbf{b}_i, \mathbf{B}_i \,] \Pi_i$$

where $\mathbf{B}_i = [\, \mathbf{b}_1, \dots, \mathbf{b}_{i-1}, \mathbf{b}_{i+1}, \dots, \mathbf{b}_n \,]$ is $\mathbf{B}$ without $\mathbf{b}_i$.
Then we have

$$\mathbf{B}^\top = \big([\mathbf{b}_i, \mathbf{B}_i] \Pi_i \big)^\top = \Pi_i^\top [\mathbf{b}_i, \mathbf{B}_i]^\top = \Pi_i^\top \begin{bmatrix} \mathbf{b}_i^\top \\ \mathbf{B}_i^\top \end{bmatrix}$$

$$\det \mathbf{B}^\top \mathbf{B} = \det \left(\Pi_i^T \begin{bmatrix} \mathbf{b}_i^\top \\ \mathbf{B}_i^\top \end{bmatrix} [\mathbf{b}_i \ \mathbf{B}_i] \, \Pi_i \right) = \det \left(\Pi_i^T \begin{bmatrix} \mathbf{b}_i^\top \mathbf{b}_i & \mathbf{b}_i^\top \mathbf{B}_i \\ \mathbf{B}_i^\top \mathbf{b}_i & \mathbf{B}_i^\top \mathbf{B}_i \end{bmatrix} \Pi_i \right)$$

As Π is square matrix, by $\det \mathbf{XY} = \det \mathbf{X} \det \mathbf{Y}$ we have

$$\det \mathbf{B}^\top \mathbf{B} = \det \left(\Pi_i^\top\right) \det \begin{bmatrix} \mathbf{b}_i^\top \mathbf{b}_i & \mathbf{b}_i^\top \mathbf{B}_i \\ \mathbf{B}_i^\top \mathbf{b}_i & \mathbf{B}_i^\top \mathbf{B}_i \end{bmatrix} \det \left(\Pi_i\right)$$

As determinant of permutation matrix is $(-1)^\#$ where $\#$ is the number of column swaps, so $\det \left(\Pi_i^\top\right) \cdot \det \left(\Pi_i\right) = (-1)^{2\#} = 1$ and

$$\det B^T B = \det \begin{bmatrix} \mathbf{b}_i^\top \mathbf{b}_i & \mathbf{b}_i^\top \mathbf{B}_i \\ \mathbf{B}_i^\top \mathbf{b}_i & \mathbf{B}_i^\top \mathbf{B}_i \end{bmatrix}$$

Looking at $\det \mathbf{B}^\top \mathbf{B} = \det \begin{bmatrix} \mathbf{b}_i^\top \mathbf{b}_i & \mathbf{b}_i^\top \mathbf{B}_i \\ \mathbf{B}_i^\top \mathbf{b}_i & \mathbf{B}_i^\top \mathbf{B}_i \end{bmatrix}$, it is not hard to think of the Schur determinant identity

$$\det \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \det(\mathbf{D}) \det(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}), \;\; \mathbf{D} \text{ non-singular.}$$

Assuming $\mathbf{B}$ full rank thus $\mathbf{B}_i$ non-singular, we can use the Schur identity

$$\det \mathbf{B}^\top \mathbf{B} = \det(\mathbf{B}_i^\top \mathbf{B}_i) \det \left(\mathbf{b}_i^\top \mathbf{b}_i - \mathbf{b}_i^\top \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top \mathbf{b}_i \right)$$

Column-wise coordinate representation of $\det \mathbf{B}^\top \mathbf{B}$ ... 3/3

Note $\mathbf{b}_i^\top \mathbf{b}_i$ is a dot product and $\mathbf{b}_i^\top \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top \mathbf{b}_i$ is a quadratic form, both are scalars, we can remove the det :

$$
\begin{aligned}
\det \mathbf{B}^\top \mathbf{B} &= \det(\mathbf{B}_i^\top \mathbf{B}_i)\big(\mathbf{b}_i^\top \mathbf{b}_i - \mathbf{b}_i^\top \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top \mathbf{b}_i\big) \\
&= \det(\mathbf{B}_i^\top \mathbf{B}_i)\big(\mathbf{b}_i^\top \mathbf{I}_m \mathbf{b}_i - \mathbf{b}_i^\top \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top \mathbf{b}_i\big) \\
&= \det(\mathbf{B}_i^\top \mathbf{B}_i)\mathbf{b}_i^\top \big(\mathbf{I}_m - \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top\big)\mathbf{b}_i \\
&= \gamma \mathbf{b}_i^\top \mathbf{X}_i \mathbf{b}_i
\end{aligned}
$$

where $\gamma = \det(\mathbf{B}_i^\top \mathbf{B}_i)$, $\mathbf{X}_i = \mathbf{I}_m - \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top$ with $\mathbf{B}_i$ as the matrix B with the i^{th} column $\mathbf{b}_i$ removed.
Note that both γ and $\mathbf{X}_i$ are independent of $\mathbf{b}_i$.

Fundamental subspaces of $\mathbf{B}_i$ ... 1/3

The term $\mathbf{X} = \mathbf{I} - \mathbf{B}_i(\mathbf{B}_i^\top \mathbf{B}_i)^{-1}\mathbf{B}_i^\top$ contains an inverse $(\mathbf{B}_i^\top \mathbf{B}_i)^{-1}$, which can be bypassed as follows : consider a matrix $\mathbf{C}_i$ as the orthonormal basis of the null space of $\mathbf{B}_i^\top$ (which can be obtained from the SVD of $\mathbf{B}_i$, more on this later).

As $\mathbf{C}$ is orthonormal,

$$\mathbf{C}_i^\top \mathbf{C}_i = \mathbf{I} \tag{1}$$

As $\mathbf{C}$ is the basis of the null space of $\mathbf{B}_i^\top$, the product

$$\mathbf{C}_i^\top \mathbf{B}_i = \mathbf{0} \tag{2}$$

From rank-null theorem:
- the dimension of null space of $\mathbf{B}_i^\top$ is $m - r + 1$
- the dimension of column space of $\mathbf{B}_i$ is $r - 1$

together they form a base of the m-dimensional space, any vector $\mathbf{p}$ in this space can be written as a linear combination of null$(\mathbf{B}_i^\top)$ and col$(\mathbf{B}_i)$ as

$$\mathbf{p} = \mathbf{C}_i \mathbf{x} + \mathbf{B}_i \mathbf{y}$$

where $\mathbf{x}$ and $\mathbf{y}$ are coefficients.

Fundamental subspaces of B_i ... 2/3

The expression $\mathbf{p} = \mathbf{C}_i\mathbf{x} + \mathbf{B}_i\mathbf{y}$ works for all vector in the m-dimensional space, so it works for vector $\mathbf{p} = \mathbf{b}_i$:

$$\mathbf{b}_i = \mathbf{C}_i\mathbf{x} + \mathbf{B}_i\mathbf{y} \tag{3}$$

Thus

$$\begin{aligned} \mathbf{b}_i^\top\mathbf{b}_i &= (\mathbf{C}_i\mathbf{x} + \mathbf{B}_i\mathbf{y})^\top(\mathbf{C}_i\mathbf{x} + \mathbf{B}_i\mathbf{y}) \\ &= (\mathbf{x}^\top\mathbf{C}_i^\top + \mathbf{y}^\top\mathbf{B}_i^\top)(\mathbf{C}_i\mathbf{x} + \mathbf{B}_i\mathbf{y}) \\ &\overset{(1,2)}{=} \mathbf{x}^\top\mathbf{x} + \mathbf{y}^\top\mathbf{B}_i^\top\mathbf{B}_i\mathbf{y} \end{aligned} \tag{4}$$

$$\begin{aligned} \mathbf{b}_i^\top\mathbf{B}_i(\mathbf{B}_i^\top\mathbf{B}_i)^{-1}\mathbf{B}_i^\top\mathbf{b}_i &= (\mathbf{C}_i\mathbf{x} + \mathbf{B}_i\mathbf{y})^\top\mathbf{B}_i(\mathbf{B}_i^\top\mathbf{B}_i)^{-1}\mathbf{B}_i^\top(\mathbf{C}_i\mathbf{x} + \mathbf{B}_i\mathbf{y}) \\ &\overset{(1,2)}{=} \mathbf{y}\mathbf{B}_i^\top\mathbf{B}_i(\mathbf{B}_i^\top\mathbf{B}_i)^{-1}\mathbf{B}_i^\top\mathbf{B}_i\mathbf{y} \\ &= \mathbf{y}\mathbf{B}_i^\top\mathbf{B}_i\mathbf{y} \end{aligned} \tag{5}$$

Put (5) into (4)

$$\mathbf{b}_i^\top\mathbf{b}_i = \mathbf{x}^\top\mathbf{x} + \mathbf{b}_i^\top\mathbf{B}_i(\mathbf{B}_i^\top\mathbf{B}_i)^{-1}\mathbf{B}_i^\top\mathbf{b}_i \tag{6}$$

Fundamental subspaces of B_i ... 3/3

Now we have

$$\begin{aligned}
\mathbf{C}_i^\top \mathbf{C}_i &= \mathbf{I} & (1)\\
\mathbf{C}_i^\top \mathbf{B}_i &= \mathbf{0} & (2)\\
\mathbf{b}_i &= \mathbf{C}_i \mathbf{x} + \mathbf{B}_i \mathbf{y} & (3)\\
\mathbf{b}_i^\top \mathbf{b}_i &= \mathbf{x}^\top \mathbf{x} + \mathbf{b}_i^\top \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top \mathbf{b}_i & (6)
\end{aligned}$$

so

$$\begin{aligned}
\mathbf{b}_i^\top \mathbf{C}_i \mathbf{C}_i^\top \mathbf{b}_i &= (\mathbf{C}_i \mathbf{x} + \mathbf{B}_i \mathbf{y})^\top \mathbf{C}_i \mathbf{C}_i^\top (\mathbf{C}_i \mathbf{x} + \mathbf{B}_i \mathbf{y}) x^\top \mathbf{x} && \text{by (3)}\\
&= \mathbf{x}^\top \mathbf{x} && \text{by (1,2)}\\
&= \mathbf{b}_i^\top \mathbf{b}_i - \mathbf{b}_i^\top \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top \mathbf{b}_i && \text{by (6)}\\
&= \mathbf{b}_i^\top (\mathbf{I} - \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top) \mathbf{b}_i
\end{aligned}$$

Hence the term $\mathbf{X} = \mathbf{I} - \mathbf{B}_i (\mathbf{B}_i^\top \mathbf{B}_i)^{-1} \mathbf{B}_i^\top$ can be computed as $\mathbf{X} = \mathbf{C}_i \mathbf{C}_i^\top$, where $\mathbf{C}_i$ is the orthonormal basis of the null space $\mathbf{B}_i^\top$.

Computing the orthonormal basis of null($\mathbf{B}_i^\top$)

The matrix $\mathbf{C}_i$ (an orthonormal basis of null($\mathbf{B}_i^\top$)) can be computed by SVD on $\mathbf{B}_i$: consider the (full) SVD of the matrix $\mathbf{B}_i \in \mathbb{R}^{m \times n}$

$$\mathbf{B}_i = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^\top$$

- $\mathbf{U} \in \mathbb{R}^{m \times m}$ is an orthogonal matrix called the left singular vector matrix
- $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_{\min(m,n)}) \in \mathbb{R}^{m \times n}$ is the singular value matrix
- $\mathbf{V} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix called the right singular vector matrix

As $\boldsymbol{\Sigma}$ has at most $\min(m, n)$ number of $\sigma_{>0}$ on the main diagonal, the remaining singular values are zero.

Let $\mathbf{U} = [\mathbf{U}_{>0} \ \mathbf{U}_{=0}]$, where the sub-matrix $\mathbf{U}_{=0}$ contains the columns of $\mathbf{U}$ corresponding to singular values equal to zero.

$\mathbf{U}_{=0}$ is the matrix $\mathbf{C}$ we want.

Last page - summary

- Column-wise coordinate representation of $\det(\mathbf{B}^\top \mathbf{B})$

$$\det \mathbf{B}^\top \mathbf{B} = \gamma \mathbf{b}_i^\top \mathbf{X} \mathbf{b}_i$$

  where
  $\gamma = \det(\mathbf{B}_i^\top \mathbf{B}_i)$
  $\mathbf{X} = \mathbf{I} - \mathbf{B}_i(\mathbf{B}_i^\top \mathbf{B}_i)^{-1}\mathbf{B}_i^\top$
  $\mathbf{B}_i$ = matrix $\mathbf{B}$ with the i^{th} column $\mathbf{b}_i$ removed
- $\mathbf{X} = \mathbf{C}_i \mathbf{C}_i^\top$, where $\mathbf{C}_i$ is the orthonormal basis of null$(\mathbf{B}_i^\top)$
- $\mathbf{C}_i$ can be computed as $\mathbf{U}_{=0}$ from the full SVD of $\mathbf{B}_i$

End of document