Column-wise representation of $\det X^T X$

Andersen Ang

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

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

Created : July 26, 2017
Last update : February 16, 2018

Overview

1. $\det B^T B$
2. The column-wise coordinate representation of $\det B^T B$
3. Computation of the column-wise coordinate representation of $\det B^T B$
4. Summary

The Gram matrix

Given a matix $B \in \mathbb{R}^{m \times n}$.
Consider the Gram matrix $G = B^T B$:

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

Consider $\det B^T B$.

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

Column-wise coordinate representation of $\det B^T B$ ... 1/4

Given $B = [\, b_1 \ b_2 \ ... \ b_n \,] \in \mathbb{R}^{m \times n}$. Suppose we want to express $\det B^T B$ as a function of the i^{th} column b_i as

$$\det B^T B = f(b_i) = \text{some expressions}$$

Here is how: consider expressing matrix $B = [\, b_1 \ b_2 \ ... \ b_n \,]$ as $[\, b_i \ b_1 \ ... \ b_{i-1} \ b_{i+1} \ ... \ b_n \,]$ with the help of column permutation matrix Π_i. i.e. the i^{th} column b_i of B is moved to the left of the matrix B by mutiplying a permulation matrix Π_i from the right with B :

$$B = [\, b_i \ B_i \,]\Pi_i$$

where $B_i = [\, b_1 \ ... \ b_{i-1} \ b_{i+1} \ ... \ b_n \,]$ is the matrix B without b_i.
Thus we have

$$B^T = ([b_i \ B_i]\Pi_i)^T = \Pi_i^T [b_i \ B_i]^T = \Pi_i^T \begin{pmatrix} b_i^T \\ B_i^T \end{pmatrix}$$

Column-wise coordinate representation of $\det B^T B$ ... 2/4

With $B = [b_i \ B_i]\Pi_i$, $\det B^T B$ becomes

$$\det B^T B = \det \left(\Pi_i^T \begin{pmatrix} b_i^T \\ B_i^T \end{pmatrix} (b_i \ B_i) \, \Pi_i \right) = \det \left(\Pi_i^T \begin{pmatrix} b_i^T b_i & b_i^T B_i \\ B_i^T b_i & B_i^T B_i \end{pmatrix} \Pi_i \right)$$

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

$$\det B^T B = \det \Pi_i^T \det \begin{pmatrix} b_i^T b_i & b_i^T B_i \\ B_i^T b_i & B_i^T B_i \end{pmatrix} \det \Pi_i$$

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

$$\det B^T B = \det \begin{pmatrix} b_i^T b_i & b_i^T B_i \\ B_i^T b_i & B_i^T B_i \end{pmatrix}$$

By looking at $\det B^T B = \det \begin{pmatrix} b_i^T b_i & b_i^T B_i \\ B_i^T b_i & B_i^T B_i \end{pmatrix}$, it is not hard to think of the Schur determinant identity

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(D) \det(A - BD^{-1}C)$$

if D is non-singular.

Assuming B has full rank and hence B_i is non-singular, we have

$$\det B^T B = \det(B_i^T B_i) \det(b_i^T b_i - b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i)$$

Note that $b_i^T b_i$ is a vector dot product and $b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i$ is a vector-matrix quadratic form, so both of them are scalars, we have

$$\det B^T B = \det(B_i^T B_i)(b_i^T b_i - b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i)$$

Column-wise coordinate representation of $\det B^T B$ ... 4/4

As $b_i^T b_i = b_i^T I b_i$ so

$$\begin{aligned}
\det B^T B &= \det(B_i^T B_i)(b_i^T I b_i - b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i) \\
&= \det(B_i^T B_i)(b_i^T I - b_i^T B_i (B_i^T B_i)^{-1} B_i^T) b_i \\
&= \det(B_i^T B_i) b_i^T (I - B_i (B_i^T B_i)^{-1} B_i^T) b_i
\end{aligned}$$

So given $B \in \mathbb{R}^{m \times n}$, the expression of $\det B^T B$ as a function of the i^{th} column b_i is

$$\det B^T B = f(b_i) = \gamma b_i^T X b_i$$

where $\gamma = \det(B_i^T B_i)$ and $X = I - B_i (B_i^T B_i)^{-1} B_i^T$ with $B_i = [b_1 \ b_2 \ ...b_{i-1} \ b_{i+1} \ ... \ b_n]$ (matrix B with the i^{th} column b_i removed). Both γ and X are independent of b_i.

Fundamental subspaces of B_i ... 1/3

The term $X = I - B_i(B_i^T B_i)^{-1} B_i^T$ contains an inverse $(B_i^T B_i)^{-1}$. The following discusses how to bypass computing inverse: consider a matrix C_i as the orthonormal basis of the null space of B_i^T (which can be obtained from the SVD of B_i, more on this later).

As C is orthonormal, the Gramian $C_i^T C_i = I$ (1)
As C is the basis of the null space of B_i^T, the product $C_i^T B_i = 0$ (2)

From rank-null theorem:
- the dimension of null space of B_i^T is $m - r + 1$
- the dimension of column space of B_i is $r - 1$

together they form a base of the m-dimensional space, any vector p in this space can be written as a linear combination of null(B_i^T) and col(B_i) as

$$p = C_i x + B_i y$$

where x and y are coefficients.

Fundamental subspaces of B_i ... 2/3

The expression $p = C_i x + B_i y$ works for all vector in the m-dimensional space, so it works for vector $p = b_i$:

$$b_i = C_i x + B_i y \tag{3}$$

Thus we have

$$\begin{aligned} b_i^T b_i &= (C_i x + B_i y)^T (C_i x + B_i y) \\ &= (x^T C_i^T + y^T B_i^T)(C_i x + B_i y) \\ \text{by } (1),(2) &= x^T x + y^T B_i^T B_i y \end{aligned} \tag{4}$$

and

$$\begin{aligned} b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i &= (C_i x + B_i y)^T B_i (B_i^T B_i)^{-1} B_i^T (C_i x + B_i y) \\ \text{by } (1),(2) &= y B_i^T B_i (B_i^T B_i)^{-1} B_i^T B_i y \\ &= y B_i^T B_i y \end{aligned} \tag{5}$$

Put (5) into (4) we have

$$b_i^T b_i = x^T x + b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i \tag{6}$$

Fundamental subspaces of B_i ... 3/3

Now we have

$$
\begin{aligned}
C_i^T C_i &= I && (1)\\
C_i^T B_i &= 0 && (2)\\
b_i &= C_i x + B_i y && (3)\\
b_i^T b_i &= x^T x + b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i && (6)
\end{aligned}
$$

so

$$
\begin{aligned}
b_i^T C_i C_i^T b_i = &\ (C_i x + B_i y)^T C_i C_i^T (C_i x + B_i y) x^T x && \text{by (3)}\\
= &\ x^T x && \text{by (1,2)}\\
= &\ b_i^T b_i - b_i^T B_i (B_i^T B_i)^{-1} B_i^T b_i && \text{by (6)}\\
= &\ b_i^T (I - B_i (B_i^T B_i)^{-1} B_i^T) b_i
\end{aligned}
$$

Hence the term $X = I - B_i (B_i^T B_i)^{-1} B_i^T$ can be computed as $X = C_i C_i^T$, where C_i is the orthonormal basis of the null space B_i^T.

Computing the orthonormal basis of null(B_i^T)

The matrix C_i (an orthonormal basis of null(B_i^T)) can be computed by SVD on B_i : consider the (full) SVD of the matrix $B_i \in \mathbb{R}^{m \times n}$

$$B_i = U \Sigma V^T$$

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

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

Let $U = [U_{>0} \; U_{=0}]$, where the submatrix $U_{=0}$ contains the columns of U corresponding to singular values equal to zero.

$U_{=0}$ is the matrix C we want.

Last page - summary

- Column-wise coordinate representation of $\det(B^T B)$

$$\det B^T B = \gamma b_i^T X b_i$$

  where $\gamma = \det(B_i^T B_i)$
  $X = I - B_i(B_i^T B_i)^{-1} B_i^T$
  $B_i =$ matrix B with the i^{th} column b_i removed
- $X = C_i C_i^T$, where C_i is the orthonormal basis of null(B_i)
- C_i can be computed as $U_{=0}$ from the full SVD of B_i

End of document