Math 5601 Independent Study Project

Jacob Hauck

1 Introduction

2 Theory

Definition 1. (Index slice) *Let $m \leq n$, where m and n are integers. Define the* ***index slice from*** m ***to*** n *by the sequence*

$$m : n = \{i\}_{i=m}^{n}. \tag{1}$$

Definition 2. (Submatrix) *Let $A \in \mathbf{R}^{m\times n}$ be a matrix. Let $I = \{I_i\}_{i=1}^{r}$ be a sequence of distinct row indices of A, and let $J = \{J_j\}_{j=1}^{c}$ be a sequence of distinct column indices of A. The* ***submatrix of*** A ***with rows*** I ***and columns*** J *is the matrix $A(I, J) \in \mathbf{R}^{r\times c}$ with entries*

$$[A(I, J)]_{ij} = A_{I_i J_j}. \tag{2}$$

If the special symbol $:$ is used as row indices or column indices, it means the entire sequence $1 : m$ or $1 : n$.

If I or J is a single integer i or j instead of a sequence, we take this to mean $I = i : i$ or $J = j : j$, the sequence consisting of that one integer.

Definition 3. (Skeleton decomposition) *Let $A \in \mathbf{R}^{m\times n}$, and let $B = A(I, J) \in \mathbf{R}^{r\times r}$ be a nonsingular, square submatrix of A. Then the* ***skeleton decomposition of*** A ***with core*** $B = A(I, J)$ *is given by*

$$\mathscr{S}_B = A(:, J)A(I, J)^{-1}A(I, :) \in \mathbf{R}^{m\times n}. \tag{3}$$

Theorem 1. *Let $A \in \mathbf{R}^{m\times n}$. If $B = A(I, J) \in \mathbf{R}^{r\times r}$ is a square submatrix of A with rank r, then*

$$A(I, J) = \mathscr{S}_B(I, J). \tag{4}$$

Proof. Let $C = \{b_{ij}\} = A(I, J)^{-1}$. Then, for $i, j \in 1 : r$,

$$\begin{aligned}
[\mathrm{RHS}(4)]_{ij} &= \sum_{k=1}^{r}\sum_{\ell=1}^{r} A_{I_i J_k} C_{k\ell} A_{I_\ell J_j} = \sum_{k=1}^{r}\sum_{\ell=1}^{r} A(I, J)_{ik} C_{k\ell} A(I, J)_{\ell j} && (5)\\
&= [A(I, J)A(I, J)^{-1}A(I, J)]_{ij} = A_{I_i J_j} = A(I, J)_{ij} && (6)\\
&= [\mathrm{LHS}(4)]_{ij}, && (7)
\end{aligned}$$

which completes the proof. □

Definition 4. (Standard basis) *Let $e_j \in \mathbf{R}^n$ denote the jth standard basis vector in $\mathbf{R}^n$.*

Theorem 2. (Exact skeleton decomposition) *Let $A \in \mathbf{R}^{m\times n}$ be a matrix with rank r. If $B = A(I,J) \in \mathbf{R}^{r\times r}$ is a square submatrix of A with rank r, then*

$$A = \mathscr{S}_B. \tag{8}$$

Proof. The columns of A at indices J (that is, $\{A(:,J_j)\}_{j=1}^r$) are linearly independent because

$$\sum_{j=1}^r \alpha_j A(:,J_j) = 0 \implies \sum_{j=1}^r \alpha_j A(I,J_j) = 0 \implies \alpha_j = 0, \quad j \in 1:r \tag{9}$$

because the columns $\{A(I,J_j)\}_{j=1}^r$ of $A(I,J)$ must be linearly independent by the fact that $A(I,J)$ has rank r.

Thus, since A has rank r, every other column of A must be a linear combination of the columns at indices J. That is, there exists $\{\alpha_{\ell j}\}$ for $j \in 1:r$ and $\ell \in 1:n$ such that

$$A(:,\ell) = \sum_{j=1}^r \alpha_{\ell j} A(:,J_j). \tag{10}$$

Define $\varphi : \mathbf{R}^r \to \operatorname{span}\{A(:,J_j) \mid j \in 1:r\}$ by $\varphi(e_j) = A(:,J_j)$. Clearly, φ is linear and onto. By the linear independence of $\{A(:,J_j)\}$, φ maps an r-dimensional space onto an r-dimensional space, so φ must also be one-to-one. Thus, φ is invertible, with $\varphi^{-1}(A(:,J_j)) = e_j$ for $j \in 1:r$.

Let $x \in \mathbf{R}^n$. Viewing A and $A(I,J)$ as linear mappings defined by matrix-vector multiplication, we have

$$\begin{aligned}
(A(I,J)\circ\varphi^{-1}\circ A)(x) &= (A(I,J)\circ\varphi^{-1})\left(\sum_{\ell=1}^n A(:,\ell)x_\ell\right) = \sum_{\ell=1}^n x_\ell A(I,J)\varphi^{-1}(A(:,\ell)) && (11)\\
&= \sum_{\ell=1}^n x_\ell A(I,J)\varphi^{-1}\left(\sum_{j=1}^r \alpha_{\ell j} A(:,J_j)\right) && (12)\\
&= \sum_{\ell=1}^n x_\ell A(I,J)\sum_{j=1}^r \alpha_{\ell j} e_j = \sum_{\ell=1}^n x_\ell \sum_{j=1}^r \alpha_{\ell j} A(I,J_j) && (13)\\
&= \sum_{\ell=1}^n x_\ell \left(\sum_{j=1}^r \alpha_{\ell j} A(:,J_j)\right)(I,:) = \sum_{\ell=1}^n A(I,\ell)x_\ell && (14)\\
&= A(I,:)x. && (15)
\end{aligned}$$

Since x was arbitrary, and $A(I,J)$ and φ^{-1} are invertible, it follows that

$$A = \varphi \circ A(I,J)^{-1} \circ A(I,:) \tag{16}$$

as a linear map.

For any $x \in \mathbf{R}^n$, we can write $A(I,J)^{-1}A(I,:)x$ as a linear combination of $\{e_j\}_{j=1}^r$; that is, there exists $\{\beta_j\}$ such that

$$A(I,J)^{-1}A(I,:)x = \sum_{j=1}^{r} \beta_j e_j. \tag{17}$$

Then

$$Ax = \varphi\left(\sum_{j=1}^{r} \beta_j e_j\right) = \sum_{j=1}^{r} \beta_j A(:,J_j) = A(:,J)\sum_{j=1}^{r} \beta_j e_j = A(:,J)A(I,J)^{-1}A(I,:)x. \tag{18}$$

Since x was arbitrary, (8) follows. □

Definition 5. (Chebyshev Norm) *If $A \in \mathbf{R}^{m\times n}$, define the **Chebyshev norm** of A by*

$$\|A\|_\infty = \max_{i,j} |A_{ij}|. \tag{19}$$

Definition 6. (Volume) *Let $A \in \mathbf{R}^{r\times r}$ be a square matrix. Then the **volume** of A is defined to be*

$$\mathcal{V}(A) = |\det(A)|. \tag{20}$$

Definition 7. (Maximum volume submatrix) *Let $A \in \mathbf{R}^{m\times n}$. A submatrix $A_\blacksquare = A(I,J) \in \mathbf{R}^{r\times r}$ of A is a **rank-r maximum volume submatrix** of A if*

$$\mathcal{V}(A_\blacksquare) = \max\Big\{\mathcal{V}(A(I',J')) \mid A(I',J') \in \mathbf{R}^{r\times r} \text{ is a submatrix of } A\Big\}. \tag{21}$$

We will typically denote maximum volume submatrices of A by $A_\blacksquare$.

Theorem 3. (Approximate skeleton decomposition) *Let $A \in \mathbf{R}^{m\times n}$ be a matrix with singular values $\{\sigma_i\}$. If $A_\blacksquare = A(I,J)$ is a rank-r maximum volume submatrix of A, then*

$$\left\|A - \mathscr{S}_{A_\blacksquare}\right\|_\infty \le (r+1)\sigma_{r+1}, \tag{22}$$

where $\sigma_{\max\{m,n\}+1} = 0$ by convention.

Proof. □

Definition 8. (Dominant submatrix of a tall matrix) *Let $A \in \mathbf{R}^{m\times r}$ have rank r (which means that $m \ge r$). A nonsingular, square submatrix $A_\square = A(I,:) \in \mathbf{R}^{r\times r}$ of A is a **dominant submatrix** of A if*

$$\left\|AA_\square^{-1}\right\|_\infty \le 1. \tag{23}$$

We will typically denote dominant submatrices of A by $A_\square$.

Lemma 1. *Let $A \in \mathbf{R}^{m\times r}$, and let $B \in \mathbf{R}^{r\times r}$. Then for any row indices I of A,*

$$(AB)(I,:) = A(I,:)B. \tag{24}$$

Proof. Observe that

$$[(AB)(I,:)]_{ij} = (AB)_{I_i j} = \sum_{k=1}^{r} A_{I_i k} B_{kj} = \sum_{k=1}^{r} A(I,:)_{ik} B_{kj} = [A(I,:)B]_{ij}, \quad i \in 1:n,\ j \in 1:r, \tag{25}$$

so $(AB)(I,:) = A(I,:)B$. $\square$

Lemma 2. *Let $A \in \mathbf{R}^{m\times r}$, and let $B \in \mathbf{R}^{r\times r}$ be nonsingular. If $A(I,:), A(I',:) \in \mathbf{R}^{r\times r}$ are square submatrices of A, and $A(I',:)$ is nonsingular, then $(AB)(I',:)$ is nonsingular, and*

$$\frac{\mathcal{V}(A(I,:))}{\mathcal{V}(A(I',:))} = \frac{\mathcal{V}((AB)(I,:))}{\mathcal{V}((AB)(I',:))}. \tag{26}$$

Proof. By Lemma 1, $(AB)(I,:) = A(I,:)B$, and $(AB)(I',:) = A(I',:)B$. Thus,

$$\det((AB)(I',:)) = \det(A(I',:)B) = \det(A(I',:))\det(B) \neq 0. \tag{27}$$

Similarly, $\det((AB)(I,:)) = \det(A(I,:))\det(B)$. Therefore,

$$\frac{\det((AB)(I,:))}{\det((AB)(I',:))} = \frac{\det(A(I,:))\det(B)}{\det(A(I',:))\det(B)} = \frac{\det(A(I,:))}{\det(A(I',:))}. \tag{28}$$

Taking the absolute value on both sides gives (26). $\square$

Theorem 4. (Maximum volume submatrices are dominant) *Let $A \in \mathbf{R}^{m\times r}$ have rank r, and let $A_{\blacksquare} \in \mathbf{R}^{r\times r}$ be a maximum volume submatrix of A. Then $A_{\blacksquare}$ is a dominant submatrix of A.*

Proof. Since the rank of A is r, there must be a set of r linearly independent rows of A, say at indices I'. Then $A(I',:)$ is nonsingular, and $\mathcal{V}(A(I',:)) > 0$. This implies that $\mathcal{V}(A_{\blacksquare}) \geq \mathcal{V}(A(I',:)) > 0$.

Since $\mathcal{V}(A_{\blacksquare}) > 0$, it follows that $A_{\blacksquare}$ is invertible. Define $B = AA_{\blacksquare}^{-1}$. There is some row index sequence I such that $A_{\blacksquare} = A(I,:)$. By Lemma 2, $A_{\blacksquare}$ has maximal volume in A if and only if $B(I,:)$ has maximal volume in B, as multiplication by the invertible matrix $A_{\blacksquare}^{-1}$ preserves the ratios of $r \times r$ submatrix volumes.

Furthermore, $B(I,:)$ is the identity matrix $I_{r\times r}$ because, by Lemma 1,

$$B(I,:) = \left(AA_{\blacksquare}^{-1}\right)(I,:) = A(I,:)A_{\blacksquare}^{-1} = A_{\blacksquare}A_{\blacksquare}^{-1} = I_{r\times r}. \tag{29}$$

Thus, $B(I,:)$ is dominant in B if and only if $\left\|BB(I,:)^{-1}\right\|_{\infty} = \|B\|_{\infty} = \left\|AA_{\blacksquare}^{-1}\right\|_{\infty} \leq 1$, that is, if and only if $A_{\blacksquare}$ is dominant in A.

We now prove the claim by contradiction. Suppose that $A_{\blacksquare}$ is not dominant in A. Then $B(I,:)$ is not dominant in B; that is, there exists $k \in 1:m$ and $j \in 1:r$ such that $|B_{kj}| > 1$.

Let $I'_i = I_i$ if $i \neq j$, and let $I'_j = k$. Then every row of $B(I', :)$ is a row of $I_{r \times r}$ except for the jth row, which is the kth row of B. That is,

$$B(I', :) = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ & & B(k, :)\ (j\text{th row}) & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}. \tag{30}$$

Expanding by cofactors and expanding on the jth row of $B(I', :)$ last shows that

$$|\det(B(I', :))| = |B_{kj}| > 1. \tag{31}$$

This means that $\mathcal{V}(B(I', :)) > 1 = \mathcal{V}(B(I, :))$, so $B(I, :)$ is not maximal in B. Then $A_{\blacksquare}$ is not maximal in A, which is a contradiction.

Hence, $A_{\blacksquare}$ is dominant in A. $\square$

Lemma 3. (Hadamard's Inequality) *Let $A \in \mathbf{R}^{m \times m}$. Then*

$$\mathcal{V}(A) \leq \prod_{i=1}^{m} \|A(:, i)\|_2, \tag{32}$$

where $\|\cdot\|_2$ is the Euclidean vector norm.

Proof. CITE $\square$

Theorem 5. (Approximation by dominant submatrix) *Let $A \in \mathbf{R}^{m \times r}$ have rank r, and let $A_{\blacksquare}$ be a maximum volume submatrix of A. Then*

$$\mathcal{V}(A_{\square}) \geq r^{-\frac{r}{2}} \mathcal{V}(A_{\blacksquare}) \tag{33}$$

for all dominant submatrices $A_{\square}$ of A. The inequality is sharp.

Proof. Let $A_{\square}$ be a dominant submatrix of A, and let $B = AA_{\square}^{-1}$. By definition, $\|B\|_{\infty} \leq 1$. Thus, if we take r rows of B at indices I, then $\|B(I, :)\|_{\infty} \leq 1$ as well, which implies that $\|B(I, j)\|_2 \leq \sqrt{r}$. By Hadamard's inequality (Lemma 3), then,

$$\mathcal{V}(B(I, :)) \leq \prod_{j=1}^{r} \|B(I, j)\|_2 \leq r^{\frac{r}{2}}, \tag{34}$$

with equality holding if $\{B(I, j)\}_{j=1}^{r}$ forms an orthogonal set.

In particular, choose I such that $A_{\blacksquare} = A(I, :)$. By Lemma 1, we have

$$r^{\frac{r}{2}} \geq \mathcal{V}(B(I, :)) = |\det(B(I, :))| = |\det(A(I, :))\det(A_{\square}^{-1})| = \frac{\mathcal{V}(A_{\blacksquare})}{\mathcal{V}(A_{\square})}. \tag{35}$$

Then (33) follows.

If we choose $A = (1, 1)^T$, then the maximum volume submatrix of A is $A_{\blacksquare} = [1]$, with volume 1. If we set $A_{\square} = A_{\blacksquare}$ and note that $r = 1$ for this choice of A, we see that $\mathcal{V}(A_{\square}) = 1 = r^{-\frac{r}{2}} \mathcal{V}(A_{\blacksquare})$. $\square$

3 The `maxvol` Algorithm

4 Implementation in `NumPy`

5 Experiments

6 Conclusion