Lecture 4: Partitioned Matrices and Determinants

Elementary row operations

Recall the **elementary operations** on the rows of a matrix, equivalent to premultiplying by an **elementary matrix** E:

(1) multiplying row i by a **nonzero** scalar α, denoted by $E^i(\alpha)$,

(2) adding β times row j to row i, denoted by $E^{ij}(\beta)$ (here β is any scalar), and

(3) interchanging rows i and j, denoted by E^{ij}, (here $i \neq j$),

called **elementary row operations** of types 1,2 and 3 resp.

Illustrations for $m = 4$:

$$E^2(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; E^{42}(\beta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \beta & 0 & 1 \end{bmatrix}, \; E^{13} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Q. Calculate the determinants of these matrices.

Determinants

Definition (**Cullen and Gale**). The **determinant** is the function $\det : \mathbb{C}^{n\times n} \to \mathbb{C}$ such that

(a) $\det\left(E^i(\alpha)\right) = \alpha$, for all $\alpha \in \mathbb{C}$, $i \in \overline{1,n}$, and

(b) $\det(AB) = \det(A)\det(B)$, for all $A, B \in \mathbb{C}^{n\times n}$.

Q. Prove that this definition is equivalent to the one you know.

Q. Explain why $\det E^{ij} = -1$, or equivalently, why

$$\det\ E^{ij}\, A = -\det\ A\ ,\ \forall\, A.$$

The Binet–Cauchy formula. If $A \in \mathbb{C}^{k\times n}$, $B \in \mathbb{C}^{n\times k}$ then

$$\det\,(AB) = \sum_{I\in Q_{k,n}} \det\ A_{I*}\ \det\ B_{*I}\ .$$

Here $Q_{k,n}$ is the **set of increasing sequences of** k **elements** from $\overline{1,n}$, for example: $Q_{2,3} = \{\{1,2\},\{1,3\},\{2,3\}\}$.

Bordered matrices

Theorem (Blattner). Let $A \in \mathbb{C}_r^{m \times n}$ and let the matrices U and V satisfy

(a) $U \in \mathbb{C}_{(m-r)}^{m \times (m-r)}$ and the columns of U are a basis for $N(A^*)$.

(b) $V \in \mathbb{C}_{(n-r)}^{n \times (n-r)}$ and the columns of U are a basis for $N(A)$.

Then the matrix

$$\begin{bmatrix} A & U \\ V^* & O \end{bmatrix} , \tag{1}$$

is nonsingular and its inverse is

$$\begin{bmatrix} A^\dagger & V^{*\dagger} \\ U^\dagger & O \end{bmatrix} . \tag{2}$$

The Cramer rule

Given a matrix A and a vector $\mathbf{b}$, $A[j \leftarrow \mathbf{b}]$ denotes the matrix obtained from A by replacing the j_{th}–column by $\mathbf{b}$.

Theorem (Cramer). Let $A \in \mathbb{C}^{n \times n}$ be nonsingular. Then for any $\mathbf{b} \in \mathbb{C}^n$, the solution $\mathbf{x} = [x_j]$ of

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

is given by

$$x_j = \frac{\det A[j \leftarrow \mathbf{b}]}{\det A}\ , \ j \in \overline{1, n}\ .$$

Proof (Robinson). Write $A\mathbf{x} = \mathbf{b}$ as

$$A\, I_n[j \leftarrow \mathbf{x}] = A[j \leftarrow \mathbf{b}]\ , \quad j \in \overline{1, n}\ ,$$

and take determinants

$$\det A\ \det I_n[j \leftarrow \mathbf{x}] = \det A[j \leftarrow \mathbf{b}]\ .$$

Bordered matrices (cont'd)

Proof. Recall: $A \in \mathbb{C}_r^{m\times n}$, and the matrices U and V satisfy

(a) $U \in \mathbb{C}_{(m-r)}^{m\times(m-r)}$ and the columns of U are a basis for $N(A^*)$.

(b) $V \in \mathbb{C}_{(n-r)}^{n\times(n-r)}$ and the columns of U are a basis for $N(A)$.

$$\begin{bmatrix} A & U \\ V^* & O \end{bmatrix} \begin{bmatrix} A^\dagger & V^{*\dagger} \\ U^\dagger & O \end{bmatrix} = \begin{bmatrix} AA^\dagger + UU^\dagger & AV^{*\dagger} \\ V^*A^\dagger & V^*V^{*\dagger} \end{bmatrix} .$$

(1) $R(U) = N(A^*) = R(A)^\perp \implies AA^\dagger + UU^\dagger = I_n$

(2) $V^*A^\dagger = V^*A^\dagger AA^\dagger = V^*A^*A^{\dagger *}A^\dagger = (AV)^*A^{\dagger *}A^\dagger = O$

(3) $AV^{*\dagger} = AV^{\dagger *} = A(V^\dagger VV^\dagger)^* = A(V^\dagger V^{\dagger *}V^*)^*$

$\qquad = AVV^\dagger V^{\dagger *} = O$,

(4) $V^*V^{*\dagger} = I_{n-r}$.

A special case

Corollary. Let $A \in \mathbb{C}_r^{m\times n}$ and let the matrices $U \in \mathbb{C}^{m\times(m-r)}$ and $V \in \mathbb{C}^{n\times(n-r)}$ satisfy

$$AV = O\ , \quad V^*V = I_{n-r}\ , \quad A^*U = O\ , \quad \text{and} \quad U^*U = I_{m-r}\ .$$

Then the matrix

$$\begin{bmatrix} A & U \\ V^* & O \end{bmatrix} ,$$

is nonsingular and its inverse is

$$\begin{bmatrix} A^\dagger & V \\ U^* & O \end{bmatrix} .$$

MNLSS

Corollary. Let $A \in \mathbb{C}_r^{m\times n}$ and let the matrices U and V satisfy

(a) $U \in \mathbb{C}_{(m-r)}^{m\times(m-r)}$ and the columns of U are a basis for $N(A^*)$.

(b) $V \in \mathbb{C}_{(n-r)}^{n\times(n-r)}$ and the columns of U are a basis for $N(A)$.

Consider the linear equation

$$A\mathbf{x} = \mathbf{b} \,. \tag{1}$$

The solution $\mathbf{x}, \mathbf{y}$ of

$$\begin{bmatrix} A & U \\ V^* & O \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix} \,. \tag{2}$$

satisfies

$\mathbf{x} = A^\dagger \mathbf{b}$, the **minimal–norm least squares solution** of (1) ,

$U\mathbf{y} = P_{N(A^*)}\mathbf{b}$, the **residual** of (1) .

MNLSS

Corollary. Let $A \in \mathbb{C}_r^{m\times n}$ and let the matrices U and V satisfy

(a) $U \in \mathbb{C}_{(m-r)}^{m\times(m-r)}$ and the columns of U are a basis for $N(A^*)$.

(b) $V \in \mathbb{C}_{(n-r)}^{n\times(n-r)}$ and the columns of U are a basis for $N(A)$.

Consider the linear equation

$$A\mathbf{x} = \mathbf{b}\ . \tag{1}$$

The **minimal–norm least–squares solution** $\mathbf{x} = [x_j]$ of (1) is given by

$$x_j = \frac{\det \begin{bmatrix} A[j \leftarrow \mathbf{b}] & U \\ V^*[j \leftarrow \mathbf{0}] & O \end{bmatrix}}{\det \begin{bmatrix} A & U \\ V^* & O \end{bmatrix}}\ , \quad j \in \overline{1,n}\ .$$

Greville's method

Let $A \in \mathbb{C}^{m\times n}$, and let $A_k = A[*, \overline{1,k}] \in \mathbb{C}^{m\times k}$ be partitioned as

$$A_k = \begin{bmatrix} A_{k-1} & \mathbf{a}_k \end{bmatrix} \tag{1}$$

Let the vectors $\mathbf{d}_k$ and $\mathbf{c}_k$ be defined by

$$\mathbf{d}_k := A_{k-1}^{\dagger}\mathbf{a}_k \tag{2}$$

$$\mathbf{c}_k := \mathbf{a}_k - A_{k-1}\mathbf{d}_k = P_{N(A_{k-1}^*)}\mathbf{a}_k \tag{3}$$

Theorem (Greville).

$$\begin{bmatrix} A_{k-1} & \mathbf{a}_k \end{bmatrix}^{\dagger} = \begin{bmatrix} A_{k-1}^{\dagger} - \mathbf{d}_k\mathbf{b}_k^* \\ \mathbf{b}_k^* \end{bmatrix} , \tag{4}$$

where

$$\mathbf{b}_k^* = \mathbf{c}_k^{\dagger} \quad \text{if } \mathbf{c}_k \neq \mathbf{0} ,$$

$$\mathbf{b}_k^* = (1 + \mathbf{d}_k^*\mathbf{d}_k)^{-1}\mathbf{d}_k^* A_{k-1}^{\dagger} \quad \text{if } \mathbf{c}_k = \mathbf{0} .$$

Schur complement

Let A be partitioned as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} , \quad A_{11} \text{ nonsingular} ,$$

and consider the homogeneous equation

$$A_{11}\mathbf{x}_1 + A_{12}\mathbf{x}_2 = \mathbf{0}$$
$$A_{21}\mathbf{x}_1 + A_{22}\mathbf{x}_2 = \mathbf{0}$$

Eliminating $\mathbf{x}_1$ we get the equation for $\mathbf{x}_2$

$$(A_{22} - A_{21}A_{11}^{-1}A_{12})\mathbf{x}_2 = \mathbf{0}$$

The **Schur complement** of A_{11} in A, denoted A/A_{11}, is

$$A/A_{11} := A_{22} - A_{21}A_{11}^{-1}A_{12} .$$

Schur complement (cont'd)

Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} , \quad A_{11} \text{ nonsingular} ,$$

$$A/A_{11} := A_{22} - A_{21} A_{11}^{-1} A_{12} \ .$$

(a) If A is square, its **determinant** is

$$\det A = \det A_{11} \det(A/A_{11}) \ .$$

(b) **The quotient property**. If A_{11} is further partitioned as

$$A_{11} = \begin{bmatrix} E & F \\ G & H \end{bmatrix} , E \text{ nonsingular} , \ \text{ then } A/A_{11} = (A/E)/(A_{11}/E) \ .$$

(c) $\qquad \operatorname{rank} A = \operatorname{rank} A_{11} \iff A/A_{11} = O.$

Schur complement (cont'd)

Let the equation $A\mathbf{x} = \mathbf{b}$ be partitioned as

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix} \tag{1}$$

Then (1) is consistent if and only if

$$(A/A_{11})\mathbf{x}_2 = \mathbf{b}_2 - A_{21}A_{11}^{-1}\mathbf{b}_1 \tag{2a}$$

is consistent, in which case a solution is completed by

$$\mathbf{x}_1 = A_{11}^{-1}(\mathbf{b}_1 - A_{12}\mathbf{x}_2)\ . \tag{2b}$$

Proof. Eliminate $\mathbf{x}_1 = A_{11}^{-1}(\mathbf{b}_1 - A_{12}\,\mathbf{x}_2)$ from the top of (1), and substitute in the bottom to get,

$$(A_{22} - A_{21}A_{11}^{-1}A_{12})\,\mathbf{x}_2 = \mathbf{b}_2 - A_{21}A_{11}^{-1}\mathbf{b}_1$$

Basic solutions

Let $A \in \mathbb{C}_n^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$, and consider the equation

$$A\,\mathbf{x} = \mathbf{b} \tag{1}$$

Let $\mathcal{I}(A)$ be then index set of **maximal full–rank (nonsingular) submatrices**

$$\mathcal{I}(A) = \{I \in Q_{n,m} : \operatorname{rank} A[I,*] = n\}$$

For each $I \in \mathcal{I}(A)$, the $I_{\mathbf{th}}$ **basic solution** of (1) is the vector

$$\mathbf{x}_I = A[I,*]^{-1}\mathbf{b}[I],$$

the solution of the subsystem

$$A[I,*]\mathbf{x} = \mathbf{b}[I]\ .$$

There are at most $\binom{m}{n}$ basic solutions.

LSS

Let $A \in \mathbb{C}_n^{m \times n}$, $\mathbf{b} \in \mathbb{C}^m$. Then the LSS $\mathbf{x}^*$ of the equation

$$A\,\mathbf{x} = \mathbf{b} \tag{1}$$

is unique,

$$\mathbf{x}^* = A^\dagger\,\mathbf{b}$$

and is a **convex combination** of the **basic solutions**

$$\mathbf{x}^* = \sum_{I \in \mathcal{I}(A)} \lambda_I \, A[I,*]^{-1} \mathbf{b}[I] \; ,$$

$$\sum_{I \in \mathcal{I}(A)} \lambda_I = 1 \; , \; \lambda_I \geq 0 \; , \; \forall \, I \; .$$

Not surprising, but the real surprise is that **the weights** λ_I are **proportional to the squares of the determinants** of $A[I,*]$,

$$\lambda_I \sim \det{}^2 A[I,*]$$

The Moore–Penrose inverse and basic inverses

Theorem. Let $A \in \mathbb{C}_n^{m\times n}$. Then

$$A^\dagger = \sum_{I\in\mathcal{I}(A)} \lambda_I \, A\widehat{[I,*]}^{-1} \,,\ \lambda_I = \frac{\det^2 A[I,*]}{\sum\limits_{K\in\mathcal{I}(A)} \det^2 A[K,*]}$$

Theorem. Let $A \in \mathbb{C}_r^{m\times n}$, and let $\mathcal{N}(A)$ be the index set of maximal nonsingular submatrices,

$$\mathcal{N}(A) := \{(I,J) : \operatorname{rank} A[I,J] = r\}$$

Then $A^\dagger$ is a convex combination,

$$A^\dagger = \sum_{(I,J)\in\mathcal{N}(A)} \lambda_{IJ} \, A\widehat{[I,J]}^{-1}$$

$$\lambda_{IJ} = \frac{\det^2 A[I,J]}{\sum\limits_{(K,L)\in\mathcal{N}(A)} \det^2 A[K,L]}$$