Matrices and Statistics with Applications
Pseudoinverses

Maya Neytcheva

SeSE, September 2020

The inverse of a nonsingular matrix

Nothing easier:
If A is a square nonsingular matrix, then A^{-1} is a matrix of the same size as A, such that

$$A^{-1}A = AA^{-1} = I.$$

Properties:

I1 $(A^{-1})^{-1} = A$

I2 $(A^{T})^{-1} = (A^{-1})^{T}$

I3 $(A^{*})^{-1} = (^{-1})^{*}$

I4 $(AB)^{-1} = B^{-1}A^{-1}$

I5 If $A\mathbf{v} = \lambda\mathbf{v}$ and $A^{-1}\mathbf{w} = \mu\mathbf{w}$ then $\mu = 1/\lambda$.

A definition of a generalized inverse

Any matrix, satisfying

$$AXA = A.$$

Example: Solvability of a linear system $A\mathbf{x} = \mathbf{b}$.

Let $\mathbf{b}$ be in the range of A, i.e., there exist a vector $\mathbf{h}$, such that $\mathbf{b} = A\mathbf{h}$.

If X is a generalized inverse of A, then $\mathbf{x} = X\mathbf{b}$.

If $AXA = A$, then $A\mathbf{x} = AX\mathbf{b} = AXA\mathbf{h} = A\mathbf{h} = \mathbf{b}$

Generalized / Pseudo- inverses

- The Moore-Penrose Pseudoinverse
- The Drazin inverse
- Weighted generalized inverses, group inverses
- The Bott-Duffin inverse (for constrained matrices)

Moore-Penrose Pseudoinverse I

The Moore-Penrose pseudoinverse A^+ is defined for any matrix and is unique. Moreover, it brings notational and conceptual clarity to the study of solutions to arbitrary systems of linear equations and linear Least Squares problems.
Consider $A \in \mathbb{R}_r^{m,n}$. The subscript r denotes the rank of A.

Moore-Penrose Pseudoinverse II

Theorem (Penrose, 1956)

Let $A \in \mathbb{R}_r^{m,n}$. Then $G = A^+$ if and only if

- P1 $AGA = A$
- P2 $GAG = G$
- P3 $(AG)^* = AG$
- P4 $(GA)^* = GA$

Furthermore, A^+ always exists and is unique.

The theorem is not constructive but gives criteria that can be checked.

Moore-Penrose Pseudoinverse III

Example:
Let $A \in \mathbb{R}_r^{m,n}$. Then, from $A = U\Sigma V^T$ we find $A^+ = V\Sigma^+ U^T$, where $\Sigma^+ = \begin{bmatrix} S^{-1} & 0 \\ 0 & 0 \end{bmatrix}$.

Moore-Penrose Pseudoinverse IV

Properties:

- $A^+ = (A^T A)^+ A^T = A^T (AA^T)^+$
- $(A^T)^+ = (A^+)^T$
- $(A^+)^+ = A$
- $(A^T A)^+ = A^+ (A^T)^+ = (A^T)^+ A^+$
- $\mathcal{R}(A^+) = \mathcal{R}(A^T) = \mathcal{R}(A^+ A) = \mathcal{R}(A^T A)$
- $\mathcal{N}(A)^+ = \mathcal{N}(AA^+) = \mathcal{N}((AA^T)^+) = \mathcal{N}(AA^T) = \mathcal{N}(A^T)$

Moore-Penrose Pseudoinverse V

For linear systems $A\mathbf{x} = \mathbf{b}$ with non-unique solutions (such as under-determined systems), the pseudoinverse may be used to construct the solution of minimum Euclidean norm $\|\mathbf{x}\|_2$ among all solutions.

If $A\mathbf{x} = \mathbf{b}$ is consistent, the vector $\mathbf{x} = A^{+}\mathbf{b}$ is a solution, and satisfies $\|\mathbf{z}\|_2 \leq \|\mathbf{x}\|_2$ for all solutions.

Uniqueness of the Moor-Penrose inverse I

Let $A \in \mathbb{R}_r^{m,n}$. Assume that there are two matrices that satisfy the conditions:

$$\begin{array}{ll} AA^+A = A & ABA = A \\ A^+AA^+ = A^+ & BAB = B \\ (AA^+)^* = AA^+ & (AB)^* = AB \\ (A^+A)^* = A^+A & (BA)^* = BA \end{array}$$

Let $M_1 = AB - AA^+ = A(B - A^+)$. By the hypothesis, M_1 is self-adjoint (since it is the difference of two self-adjoint matrices) and

$$\begin{aligned} (M_1)^2 &= (AB - AA^+)A(B - A^+) \\ &= (ABA - AA^+A)(B - A^+) = (A - A)(B - A^+)A = 0. \end{aligned}$$

Uniqueness of the Moor-Penrose inverse II

Since M_1 is self-adjoint, the fact that $M_1^2 = 0$ implies that $M_1 = 0$, since for all x one has $\|M_1 x\|^2 = (M1x, M1x) = (x, (M1)^2 x) = 0$, implying $M_1 = 0$. This showed that $AB = AA^+$.

Following the same steps we can prove that $BA = A^+ A$ (consider the self-adjoint matrix $M2 := BAA + A$ and proceed as above). Thus, $A^+ = A^+ AA^+ = A^+(AA^+) = A^+ AB = (A^+ A)B = BAB = B$, thus A^+ is unique.

The Drazin Inverse

Defined for a square matrix.
Let A be a square matrix. The index k of A is the least nonnegative integer k such that $rank(A^{k+1}) = rank(A^k)$.
The Drazin inverse of A is the unique matrix A^D which satisfies

$$A^{k+1}A^D = A^k, \quad A^D A A^D = A^D, \quad A A^D = A^D A.$$

If A is invertible with inverse A^{-1}, then $A^D = A^{-1}$.

Example: Solving systems with a singular matrix by CG

I. Ipsen, C. Meyer, The idea behind Krylov methods, *The American Mathematical Monthly*, 105 (1998)

" ... we show that the solution to a nonsingular linear system $Ax = b$ lies in a Krylov space whose dimension is the degree of the minimal polynomial of A. Therefore, if the minimal polynomial of A has low degree then the space in which a Krylov method searches for the solution is small. In this case a Krylov method has the opportunity to converge fast.

"When the matrix is singular, however, Krylov methods can fail. Even if the linear system does have a solution, it may not lie in a Krylov space. In this case we describe the class of right-hand sides for which a solution lies in a Krylov space. As it happens, there is only a single solution that lies in a Krylov space, and it can be obtained from the Drazin inverse."

Theoretical result

The following statements are equivalent:

- $Ax = b$ has a Krylov solution.
- $b \in R(A^i)$, where i is the index of the zero eigenvalue of A (the index i of an eigenvalue is the maximum size of a block, containing the eivenvalue in the Jordan canonical form).
- $A^D b$ is a solution of $Ax = b$ and it is unique.

Computing the pseudoinverse from SVD

$$A = U\Sigma V^T \;\rightarrow\; A^\dagger = V\Sigma^\dagger U^T,$$

where $A = U \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} V^T$ and $\Sigma^\dagger = \begin{bmatrix} \Sigma_1^{-1} \\ 0 \end{bmatrix}$.

Bott-Duffin inverse

Constrained generalized inverse of a square matrix: We want to solve $Ax = b$, $A(n, n)$, where x should belong to a certain subspace L of R^n.
Denote P_L to be the orthogonal projection on L. Then the constrained problem $Ax = b,\ x \in L$ has a solution if

$$AP_Lx = b$$

is solvable.
The generalized Bott-Duffin inverse is defined as

$$A^{(+)} = P_L(AP_L + P_{L^\perp})^{-1}$$

if the inverse on the right exists.