Matrix Properties

Go to: Introduction, Notation, Index

Adjoint or Adjugate

The adjoint of **A**, **ADJ**(**A**) is the transpose of the matrix formed by taking the cofactor of each element of **A**.

- **ADJ**(**A**) **A** = det(**A**) **I**
  - If det(**A**) != 0, then $\mathbf{A}^{-1}$ = **ADJ**(**A**) / det(**A**) but this is a numerically and computationally poor way of calculating the inverse.
- $\mathbf{ADJ}(\mathbf{A}^T)=\mathbf{ADJ}(\mathbf{A})^T$
- $\mathbf{ADJ}(\mathbf{A}^H)=\mathbf{ADJ}(\mathbf{A})^H$

Characteristic Equation

The *characteristic equation* of a matrix $\mathbf{A}_{[n\#n]}$ is $|t\mathbf{I}-\mathbf{A}| = 0$. It is a polynomial equation in t.

The properties of the characteristic equation are described in the section on eigenvalues.

Characteristic Matrix

The *characteristic matrix* of $\mathbf{A}_{[n\#n]}$ is (t**I**-**A**) and is a function of the scalar t.

The properties of the characteristic matrix are described in the section on eigenvalues.

Characteristic Polynomial

The *characteristic polynomial*, p(t), of a matrix $\mathbf{A}_{[n\#n]}$ is $p(t) = |t\mathbf{I} - \mathbf{A}|$.

The properties of the characteristic polynomial are described in the section on eigenvalues.

Cofactor

The *cofactor* of a minor of **A**:$n\#n$ is equal to the product of (i) the determinant of the submatrix consisting of all the rows and columns that are not in the minor and (ii) -1 raised to the power of the sum of all the row and column indices that are in the minor.

- The cofactor of the element $a(i,j)$ equals -1^{i+j} det(**B**) where **B** is the matrix formed by deleting row i and column j from **A**.

See Minor, Adjoint

Compound Matrix

The k^{th} compound matrix of $\mathbf{A}_{[m\#n]}$ is the $m!(k!(m-k)!)^{-1}\#n!(k!(n-k)!)^{-1}$ matrix formed from the determinants of all k#k submatrices of **A** arranged with the submatrix index sets in lexicographic order. Within this section, we denote this matrix by $\mathbf{C}_k(\mathbf{A})$.

- $\mathbf{C}_1(\mathbf{A}) = \mathbf{A}$
- $\mathbf{C}_n(\mathbf{A}_{[n\#n]}) = \det(\mathbf{A})$
- $\mathbf{C}_k(\mathbf{AB}) = \mathbf{C}_k(\mathbf{A})\mathbf{C}_k(\mathbf{B})$
- $\mathbf{C}_k(a\mathbf{X}) = a^k\mathbf{C}_k(\mathbf{X})$
- $\mathbf{C}_k(\mathbf{I}) = \mathbf{I}$
- $\mathbf{C}_k(\mathbf{A}^H) = \mathbf{C}_k(\mathbf{A})^H$
- $\mathbf{C}_k(\mathbf{A}^T) = \mathbf{C}_k(\mathbf{A})^T$
- $\mathbf{C}_k(\mathbf{A}^{-1}) = \mathbf{C}_k(\mathbf{A})^{-1}$

Condition Number

The *condition number* of a matrix is its largest singular value divided by its smallest singular value.

- If $\mathbf{Ax}=\mathbf{y}$ and $\mathbf{A}(\mathbf{x}+\mathbf{p})=\mathbf{y}+\mathbf{q}$ then $||\mathbf{p}||/||\mathbf{x}|| <= k\ ||\mathbf{q}||/||\mathbf{y}||$ where k is the condition number of $\mathbf{A}$. Thus it provides a sensitivity bound for the solution of a linear equation.
- If $\mathbf{A}_{[2\#2]}$ is hermitian positive definite then its condition number, r, satisfies $4 <= \mathrm{tr}(\mathbf{A})^2/\det(\mathbf{A}) = (r+1)^2/r$. This expression is symmetric between r and r^{-1} and is monotonically increasing for $r>1$. It therefore provides an easy way to check on the range of r.

Conjugate Transpose

$\mathbf{X}=\mathbf{Y}^H$ is the Hermitian transpose or Conjugate transpose of $\mathbf{Y}$ iff $x_{i,j}=y_{j,i}{}^C$.

See Hermitian Transpose.

Constructibility

The pair of matrices $\{\mathbf{A}_{[n\#n]}, \mathbf{C}_{[m\#n]}\}$ are *constructible* iff $\{\mathbf{A}^H, \mathbf{C}^H\}$ are controllable.

- If $\{\mathbf{A}, \mathbf{C}\}$ are observable then they are constructible.
- If $\det(\mathbf{A})!=0$ and $\{\mathbf{A}, \mathbf{C}\}$ are constructible then they are observable.
- If $\{\mathbf{A}, \mathbf{C}\}$ are constructible then they are detectable.

Controllability

The pair of matrices $\{\mathbf{A}_{[n\#n]}, \mathbf{B}_{[n\#m]}\}$ are *controllable* iff any of the following equivalent conditions are true

1. There exists a $\mathbf{G}_{[mn\#n]}$ such that $\mathbf{A}^n = \mathbf{CG}$ where $\mathbf{C} = [\mathbf{B}\ \mathbf{AB}\ \mathbf{A}^2\mathbf{B}\ \ldots\ \mathbf{A}^{n-1}\mathbf{B}]_{[n\#mn]}$ is the *controllability matrix*.
2. If $\mathbf{x}^T\mathbf{A}^r\mathbf{B} = \mathbf{0}$ for $0<=r<n$ then $\mathbf{x}^T\mathbf{A}^n = \mathbf{0}$.
3. If $\mathbf{x}^T\mathbf{B} = \mathbf{0}$ and $\mathbf{x}^T\mathbf{A} = k\mathbf{x}^T$ then either $k=0$ or else $\mathbf{x} = \mathbf{0}$.

- If $\{\mathbf{A}, \mathbf{B}\}$ are reachable then they are controllable.
- If $\det(\mathbf{A})!=0$ and $\{\mathbf{A}, \mathbf{B}\}$ are controllable then they are reachable.
- If $\{\mathbf{A}, \mathbf{B}\}$ are controllable then they are stabilizable.

- {**DIAG(a)**, **b**} are controllable iff all non-zero elements of **a** are distinct and all the corresponding elements of **b** are non-zero.

Definiteness

A Hermitian square matrix **A** is

- *positive definite* if $\mathbf{x}^H\mathbf{A}\mathbf{x} > 0$ for all non-zero **x**.
- *positive semi-definite* or *non-negative definite* if $\mathbf{x}^H\mathbf{A}\mathbf{x} >= 0$ for all non-zero **x**.
- *indefinite* if $\mathbf{x}^H\mathbf{A}\mathbf{x}$ is > 0 for some **x** and < 0 for some other **x**.

This definition only applies to Hermitian and real-symmetric matrices; if **A** is non-real and non-Hermitian then $\mathbf{x}^H\mathbf{A}\mathbf{x}$ is complex for some values of **x** and so the concept of definiteness does not make sense. Some authors also call a real non-symmetric matrix positive definite if $\mathbf{x}^H\mathbf{A}\mathbf{x} > 0$ for all non-zero real **x**; this is true iff its symmetric part is positive definite (see below).

- A (not necessarily symmetric) real matrix **A** satisfies $\mathbf{x}^H\mathbf{A}\mathbf{x} > 0$ for all non-zero real **x** iff its symmetric part $\mathbf{B}=(\mathbf{A}+\mathbf{A}^T)/2$ is positive definite. Indeed $\mathbf{x}^T\mathbf{A}\mathbf{x}= \mathbf{x}^T\mathbf{B}\mathbf{x}$ for all **x**.
- The following are equivalent
  - **A** is Hermitian and +ve semidefinite
  - $\mathbf{A}=\mathbf{B}^H\mathbf{B}$ for some **B** (not necessarily square)
  - $\mathbf{A}=\mathbf{C}^2$ for some Hermitian **C**.
  - $\mathbf{D}^H\mathbf{A}\mathbf{D}$ is Hermitian and +ve semidefinite for any **D**
- If **A** is +ve definite then $\mathbf{A}^{-1}$ exists and is +ve definite.
- If **A** is +ve semidefinite, then for any integer k>0 there exists a unique +ve semidefinite **B** with $\mathbf{A}=\mathbf{B}^k$. This **B** also satisifes:
  - **AB=BA**
  - **B**=p(**A**) for some polynomial p()
  - rank(**B**) = rank(**A**)
  - if **A** is real then so is **B**.
- **A** is +ve definite iff all its eigenvalues are > 0.
  - If **A** is +ve definite then det(**A**) > 0 and tr(**A**) > 0.
  - A Hermitian matrix $\mathbf{A}_{[2\#2]}$ is +ve definite iff det(**A**) >0 and tr(**A**) > 0.
- The columns of $\mathbf{B}_{[m\#n]}$ are linearly independent iff $\mathbf{B}^H\mathbf{B}$ is +ve definite.
- If **S** is +ve semidefinite, then $|\mathbf{a}^H\mathbf{S}\mathbf{b}|^2 <= \mathbf{a}^H\mathbf{S}\mathbf{a}\times\mathbf{b}^H\mathbf{S}\mathbf{b}$ for any **a**, **b** [3.6]
  - $|s_{i,j}| <= \mathrm{sqrt}(s_{i,i}s_{j,j})$ [3.6]
- If **A** and **B** are positive semidefinite, then **A+B** is positive semidefinite
- If **B** is +ve definite and **A** is +ve semidefinite then:
  - $\mathbf{B}^{-1}\mathbf{A}$ is diagonalizable and has non-negative eigenvalues [3.7]
  - $\mathrm{tr}(\mathbf{B}^{-1}\mathbf{A}) = 0$ iff **A=0**
  - **A+B** is positive definite

Detectability

The pair of matrices $\{\mathbf{A}_{[n\#n]}, \mathbf{C}_{[m\#n]}\}$ are *detectable* iff $\{\mathbf{A}^H, \mathbf{C}^H\}$ are stabilizable.

If {**A**, **C**} are observable or constructible then they are detectable..

Determinant

For an $n\#n$ matrix **A**, det(**A**) is a scalar number defined by det(**A**)=**sgn**(**PERM**(n))'***prod**(**A**(1:n,**PERM**(n)))

This is the sum of $n!$ terms each involving the product of n matrix elements of which exactly one comes from each row and each column. Each term is multiplied by the signature (+1 or -1) of the column-order permutation !. See the notation section for definitions of **sgn**(), **prod**() and **PERM**().

The determinant is important because **INV(A)** exists iff det(**A**) != 0.

Geometric Interpretation

The determinant of a matrix equals the +area of the +parallelogram that has the matrix columns as n of its sides. If a vector space is transformed by multiplying by a matrix **A**, then all +areas will be multiplied by det(**A**).

Properties of Determinants

- det($\mathbf{A}^T$) = det(**A**)
- det($\mathbf{A}^H$) = conj(det(**A**))
- det(c**A**) = c^n det(**A**)
- det($\mathbf{A}^k$) = (det(**A**))^k , k must be positive if det(**A**)=0.
- Interchanging any pair of columns of a matrix multiplies its determinant by -1(likewise rows).
- Multiplying any column of a matrix by c multiplies its determinant by c (likewise rows).
- Adding any multiple of one column onto another column leaves the determinant unaltered (likewise rows).
- det(**A**) != 0 iff **INV(A)** exists.
- [**A**,**B**:$n\#m$; m>=n]: If **Q** = **CHOOSE**(m,n). and **d**(k) = det(**A**(:,**Q**(k,:)) det(**B**(:,**Q**(k,:)) for k=1:rows(**Q**) then det($\mathbf{AB}^T$) = sum(**d**). This is the Binet-Cauchy theorem.
- Suppose that for some r, **P** = **CHOOSE**(n,r) and **Q** = **CHOOSE**(n,n-r) with the rows of **Q** ordered so that **P**(k,:) and **Q**(k,:) have no elements in common. If we define **D**(m,k) = $(-1)^{\text{sum}([\mathbf{P}(m,:)\ \mathbf{P}(k,:)])}$ det(**A**(**P**(m,:)^T,**P**(k,:)) det(**A**(**Q**(m,:)^T,**Q**(k,:)) for m,k=1:rows(**P**) then det(**A**) = sum(**D**(m,**:**)) = sum(**D**(:,k)) for any k or m. This is the Laplace expansion theorem.
    - If we set k=r=1 then **P**(m,:)=[m] and we obtain the familiar expansion by the first column: **d**(m)=$(-1)^{m+1}$ **A**(m,1) det(**A**([1:m-1 m+1:n]^T,2:n)) and det(**A**)=sum(**d**).
- det(**A**) = 0 iff the columns of **A** are linearly dependent (likewise rows).
    - det(**A**) = 0 if two columns are identical (likewise rows).
    - det(**A**) = 0 if any column consists entirely of zeros (likewise rows).
- If **A** = [$\mathbf{a}_1$ $\mathbf{a}_2$ ... $\mathbf{a}_n$] then |det(**A**)| <= prod(||$\mathbf{a}_i$||) with equality iff the $\mathbf{a}_i$ are mutually orthogonal where ||**a**|| is the Euclidean norm; this is the *Hadamard inequality*.
    - If |$a_{i,j}$|<=B for all i,j then |det(**A**)| <= $n^{0.5n}B^n$
    - [**A** +ve semidefinite]: det(**A**) <= prod(**diag**(**A**))
- [**A**:3#3]: If **A** = [**a b c**] then det(**A**) = det([**a b c**]) = $\mathbf{a}^T$ **SKEW**(**b**) **c** = $\mathbf{b}^T$ **SKEW**(**c**) **a** = $\mathbf{c}^T$ **SKEW**(**a**) **b**

Determinants of simple matrices

- det([a b; c d]) = ad - bc
- det([**a b c**]) = $a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_2c_1b_3 + a_3b_1c_2 - a_3c_1b_2$
- The determinant of a diagonal or triangular matrix is the product of its diagonal elements.
- The determinant of a unitary matrix has an absolute value of 1.
    - The determinant of an orthogonal matrix is +1 or -1.
- The determinant of a permutation matrix equals the signature of the column permutation.

Determinants of sums and products

- [**A**,**B**:$n\#n$]:det(**AB**) = det(**A**) det(**B**)

- [A,B:$m\#n$]:det($\mathbf{I} + \mathbf{A}^T\mathbf{B}$) = det($\mathbf{I} + \mathbf{A}\mathbf{B}^T$) = det($\mathbf{I} + \mathbf{B}^T\mathbf{A}$) = det($\mathbf{I} + \mathbf{B}\mathbf{A}^T$) [3.2]
- [A:$n\#n$]:det($\mathbf{A}+\mathbf{x}\mathbf{y}^T$) = ($1+\mathbf{y}^T\mathbf{A}^{-1}\mathbf{x}$) det($\mathbf{A}$) [3.4]
  - det($\mathbf{I}+\mathbf{x}\mathbf{y}^T$) = $1+\mathbf{y}^T\mathbf{x}$ = $1+\mathbf{x}^T\mathbf{y}$ [3.3]
  - det($k\mathbf{I}+\mathbf{x}\mathbf{y}^T$) = $k^n+k^{n-1}\mathbf{y}^T\mathbf{x}$ = $k^n+k^{n-1}\mathbf{x}^T\mathbf{y}$
- [A,B: n#n, symmetric, +ve semidefinite]:
  - $(\det(\mathbf{A}+\mathbf{B}))^{1/n} >= (\det(\mathbf{A}))^{1/n} + (\det(\mathbf{B}))^{1/n}$; this is the *Minkowski determinant inequality*.
    - If 0<=k<=1, then $(\det(k\mathbf{A}+(1-k)\mathbf{B}))^{1/n} >= k(\det(\mathbf{A}))^{1/n} + (1-k)(\det(\mathbf{B}))^{1/n}$
  - If 0<=k<=1, then det($k\mathbf{A}+(1-k)\mathbf{B}$) >= $(\det(\mathbf{A}))^k (\det(\mathbf{B}))^{1-k}$
    - det($\mathbf{A}+\mathbf{B}$) >= sqrt(det(4$\mathbf{AB}$))
  - For any integer m>0, $n(\det(\mathbf{A})\det(\mathbf{B}))^{m/n}$ <= tr($\mathbf{A}^m\mathbf{B}^m$)

Determinants of block matrices/a>

In this section we have $\mathbf{A}_{[m\#m]}$, $\mathbf{B}_{[m\#n]}$, $\mathbf{C}_{[n\#m]}$ and $\mathbf{D}_{[n\#n]}$.

- det([**A**, **B**; **C**, **D**]) = det([**D**, **C**; **B**, **A**]) = det(**A**)*det($\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}$) = det(**D**)*det($\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}$) [3.1]
  - det([a, $\mathbf{b}^T$; **c**, **D**]) = (a - $\mathbf{b}^T\mathbf{D}^{-1}\mathbf{c}$)det(**D**)
- det([**I**, **B**; **C**, **I**]) = det($\mathbf{I}_{[m\#m]}$-**BC**) = det($\mathbf{I}_{[n\#n]}$-**CB**)
- det([**A**, **B**; **0**, **D**]) = det([**A**, **0**; **C**, **D**]) = det(**A**) det(**D**)
  - det([a, $\mathbf{b}^T$; **0**, **D**]) = det([a, **0**; **c**, **D**]) = a det(**D**)
- For the special case when m=n (i.e. **A**, **B**, **C**, **D** all $n\#n$):
  - det([**A**, **B**; **C**, **0**]) = -det($\mathbf{B}\mathbf{C}^T$)
  - **[AB=BA]**: det([**A**, **B**; **C**, **D**]) **=** det(**DA-CB**)
  - **[AC=CA]**: det([**A**, **B**; **C**, **D**]) **=** det(**AD-CB**)
  - **[BD=DB]**: det([**A**, **B**; **C**, **D**]) **=** det(**DA-BC**)
  - **[CD=DC]**: det([**A**, **B**; **C**, **D**]) **=** det(**AD-BC**)

See also Grammian, Schur Complement

Displacement Rank

The displacement rank of $\mathbf{X}_{[m\#n]}$ is given by dis_rank(**X**) = rank($\mathbf{X} - \mathbf{Z}\mathbf{X}\mathbf{Z}^T$) where the **Z** are shift matrices of size $m\#m$ and $n\#n$ respectively.

- dis_rank(**X**+**Y**) <= dis_rank(**X**) + dis_rank(**Y**)
- dis_rank(**XY**) <= dis_rank(**X**) + dis_rank(**Y**)
- dis_rank($\mathbf{X}^{-1}$)=dis_rank(**JXJ**) where **J** is the exchange matrix.
- [**X**: Toeplitz] dis_rank(**X**) = 2 unless **X** is upper or lower triangular in which case dis_rank(**X**)=1 unless **X** = **0** , in which case dis_rank(**X**)=0.
  - [$\mathbf{X}_{[n\#n]}$: Toeplitz] If $a = \mathbf{X}_{1,1}$ and $b = \mathbf{X}^2_{1,1}$, then the characteristic polynomial of $\mathbf{X} - \mathbf{Z}\mathbf{X}\mathbf{Z}^T$ is $(t^2 - at + a^2-b)\, t^{n-2}$

Eigenvalues

The eigenvalues of **A** are the roots of its characteristic equation: $|t\mathbf{I}-\mathbf{A}| = 0$.

The properties of the eigenvalues are described in the section on eigenvalues.

Field of Values

The *field of values* of a square matrix **A** is the set of complex numbers $\mathbf{x}^H\mathbf{Ax}$ for all **x** with $||\mathbf{x}||=1$.

- The field of values is a closed convex set.
- The field of values contains the convex hull of the eigenvalues of **A**.
- If **A** is normal then the field of values equals the convex hull of its eigenvalues.
  - [n<5] $\mathbf{A}_{[n\#n]}$ is normal iff its field of values is the convex hull of its eigenvalues.
- **A** is hermitian iff its field of values is a real interval.
- If **A** and **B** are unitarily similar, they have the same field of values.

Generalized Inverse

A *generalized inverse* of $\mathbf{X}$:*m#n* is any matrix, $\mathbf{X}^\#$:*n#m* satisfying $\mathbf{XX}^\#\mathbf{X}=\mathbf{X}$. Note that if X is singular or non-square, then $\mathbf{X}^\#$ is not unique. This is also called a *weak generalized inverse* to distinguish it from the pseudoinverse.

- If **X** is square and non-singular, $\mathbf{X}^\#$ is unique and equal to $\mathbf{X}^{-1}$.
- $(\mathbf{X}^\#)^H$ is a generalized inverse of $\mathbf{X}^H$.
- [k!=0] $\mathbf{X}^\#/k$ is a generalized inverse of $k\mathbf{X}$.
- [**A**,**B** non-singular] $\mathbf{B}^{-1}\mathbf{X}^\#\mathbf{A}^{-1}$ is a generalized inverse of **AXB**
- rank($\mathbf{X}^\#$) >= rank(**X**).
- rank(**X**)=rank($\mathbf{X}^\#$) iff **X** is also the generalized inverse of $\mathbf{X}^\#$ (i.e. $\mathbf{X}^\#\mathbf{XX}^\#=\mathbf{X}^\#$.).
- $\mathbf{XX}^\#$ and $\mathbf{X}^\#\mathbf{X}$ are idempotent and have the same rank as **X**.
  - $\mathbf{I}$-$\mathbf{XX}^\#$ and I-$\mathbf{X}^\#\mathbf{X}$ are also idempotent.
- If **Ax-b** has any solutions, then $\mathbf{x}=\mathbf{A}^\#\mathbf{b}$ is a solution.
- If $\mathbf{AA}^\#$ is hermitian, a value of **x** that minimizes $||\mathbf{Ax}-\mathbf{b}||$ is given by $\mathbf{x}=\mathbf{A}^\#\mathbf{b}$. With this value of **x**, the error **Ax-b** is orthogonal to the columns of **A**. If we define the projection matrix $\mathbf{P}=\mathbf{AA}^\#$, then **Ax=Pb** and **Ax-b=-(I-P)b.**
- If **X**:*m#n* has rank *r*, we can find **A**:*n#n-r*, **B**:*n#r* and **C**:*m#m-r* whose columns form bases for the null space of **X**, the range of $\mathbf{X}^+\mathbf{X}$ and the null space of $\mathbf{X}^H$ respectively.
  - The set of generalized inverses of **X** is precisely given by $\mathbf{X}^\#=\mathbf{X}^++\mathbf{AY}+\mathbf{BZC}^H$ for arbitrary **Y**:*n-r#m* and **Z**:*r#m-r* where $\mathbf{X}^+$ is the pseudoinverse.
  - For a given choice of **A**, **B** and **C**, each $\mathbf{X}^\#$ corresponds to a unique **Y** and **Z**.
  - $\mathbf{XX}^\#$ is hermitian iff **Z**=0.
- If **X**:*m#n* has rank *r*, we can find **A**:*n#n-r*, **F**:*n#r* and **C**:*m#m-r* whose columns form bases for the null space of **X**, the range of $\mathbf{X}^+$ and the null space of $\mathbf{X}^H$ respectively. We can also find **G**:*m#r* such that $\mathbf{X}^+=\mathbf{FG}^H$.
  - The set of generalized inverses $\mathbf{X}^\#$ of **X,** for which **X** is also the generalised inverse of $\mathbf{X}^\#$ is precisely given by $\mathbf{X}^\#=(\mathbf{F}+\mathbf{AV})(\mathbf{G}+\mathbf{CW})^H$ for arbitrary **V**:*n-r#r* and **W**:*m-r#r*.
  - For a given choice of **A**, **C**, **F** and **G** each $\mathbf{X}^\#$ corresponds to a unique **V** and **W**.

See also: Pseudoinverse

Gram Matrix

The *gram matrix* of **X, GRAM**(**X**), is the matrix $\mathbf{X}^H\mathbf{X}$.

- **GRAM**(**X**) is positive semi-definite hermitian.
- det(**GRAM**(**X**)) = 0 iff a principal minor of **GRAM**(**X**) is zero.
- rank(**GRAM**(**X**)) = rank(**X**)

- trace(**GRAM**(**X**)) = $||\mathbf{X}||_F^2$, the squared Frobenius matrix norm.
- **y** is an eigenvector of $\mathbf{X}^H\mathbf{X}$ iff **Xy** is an eigenvector of $\mathbf{X}\mathbf{X}^H$. The corresponding eigenvalue is the same in both cases.

If **X** is *m#n*, the elements of **GRAM**(**X**) are the n^2 possible inner products between pairs of its columns. We can form such a matrix from *n* vectors in any vector space having an inner product.

See also: Grammian

Grammian

The *grammian* of a matrix **X,** gram(**X**), equals det(**GRAM**(**X**)) = det($\mathbf{X}^H\mathbf{X}$).

- gram(**X**) is real and >= 0.
- gram(**X**) > 0 iff the columns of **X** are linearly independent, i.e. iff **Xy** = 0 implies **y** = 0
  - [$\mathbf{X}_{m\#n}$]: gram(**X**)=0 if $m<n$.
- gram(**X**) = 0 iff a principal minor of **GRAM**(**X**) is zero.
- [$\mathbf{X}_{n\#n}$]: gram(**X**) = gram($\mathbf{X}^H$) = $|\det(\mathbf{X})|^2$
- gram(**x**) = $\mathbf{x}^H\mathbf{x}$
- gram([**X Y**]) = gram([**Y X**]) = gram(**X**)*det($\mathbf{Y}^H\mathbf{Y}-\mathbf{Y}^H\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{Y}$) = gram(**X**)*det($\mathbf{Y}^H(\mathbf{I}-\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H)\mathbf{Y}$)
  - gram([**X y**]) = gram([**y X**]) = gram(**X**)*$\mathbf{y}^H\mathbf{y}-\mathbf{y}^H\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{y}$ = gram(**X**)*$\mathbf{y}^H(\mathbf{I}-\mathbf{X}(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H)\mathbf{y}$
- gram([**X y**]) = gram(**X**) $||\mathbf{X}\mathbf{X}^\#\mathbf{y} - \mathbf{y}||^2$ where $\mathbf{X}^\#$ is the generalized inverse so that $||\mathbf{X}\mathbf{X}^\#\mathbf{y} - \mathbf{y}||$ equals the distance between **y** and its orthogonal projection onto the space spanned by the columns of **X**.
- gram([**X Y**]) <= gram(**X**) gram(**Y**); this is the *generalised Hadamard inequality*.
  - gram([**X Y**]) = gram(**X**) gram(**Y**) iff either $\mathbf{X}^H\mathbf{Y} = \mathbf{0}$ or gram(**X**) gram(**Y**) = 0
  - If **X** = [$\mathbf{x}_1\ \mathbf{x}_2 \ldots \mathbf{x}_n$] then gram(**X**) <= prod($||\mathbf{x}_i||^2$) = prod(diag($\mathbf{X}^H\mathbf{X}$))
    - [$\mathbf{X}_{n\#n}$]: $|\det(\mathbf{X})|^2$<= prod($||\mathbf{x}_i||^2$) = prod(diag($\mathbf{X}^H\mathbf{X}$)); this is the *Hadamard inequality*.

Geometric Interpretation

The grammian of $\mathbf{X}_{m\#n}$ is the squared "volume" of the *n*-dimensional parallelepiped spanned by the columns of **X**.

See also: Gram Matrix

Hermitian Transpose or Conjugate Transpose

$\mathbf{X}=\mathbf{Y}^H$ is the Hermitian transpose or Conjugate transpose of **Y** iff $x(i,j)$=conj($y(j,i)$).

Inertia

The inertia of an *m#m* square matrix is the triple (p,n,z) where $p+n+z=m$ and *p*, *n* and *z* are respectively the number of eigenvalues, counting multiplicities, with positive, negative and zero real parts.

Inverse

B is a *left inverse* of **A** if **BA=I**. **B** is a *right inverse* of **A** if **AB=I**.

If $\mathbf{BA}=\mathbf{AB}=\mathbf{I}$ then $\mathbf{B}$ is the *inverse* of $\mathbf{A}$ and we write $\mathbf{B}=\mathbf{A}^{-1}$.

- **[A:n#n]** $\mathbf{AB}=\mathbf{I}$ iff $\mathbf{BA}=\mathbf{I}$, hence *inverse*, *left inverse* and *right inverse* are all equivalent for square matrices.
- **[A,B:n#n]** $(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$
- **[A:m#n]** $\mathbf{A}$ has a left inverse iff rank($\mathbf{A}$)=n and a right inverse iff rank($\mathbf{A}$)=m.
- **[A:n#m, B:m#n]** $\mathbf{AB}=\mathbf{I}$ implies that $n<=m$ and that rank($\mathbf{A}$)=rank($\mathbf{B}$)=n.

Inverse of Block Matrices

- $[\mathbf{A}, \mathbf{B}; \mathbf{C}, \mathbf{D}]^{-1} = [\mathbf{Q}^{-1}, -\mathbf{Q}^{-1}\mathbf{B}\mathbf{D}^{-1}; -\mathbf{D}^{-1}\mathbf{C}\mathbf{Q}^{-1}, \mathbf{D}^{-1}(\mathbf{I}+\mathbf{C}\mathbf{Q}^{-1}\mathbf{B}\mathbf{D}^{-1})]$ where $\mathbf{Q} = (\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})$ is the *Schur Complement* of $\mathbf{D}$ [3.5]
  $= [\mathbf{A}^{-1}(\mathbf{I}+\mathbf{B}\mathbf{P}^{-1}\mathbf{C}\mathbf{A}^{-1}), -\mathbf{A}^{-1}\mathbf{B}\mathbf{P}^{-1}; -\mathbf{P}^{-1}\mathbf{C}\mathbf{A}^{-1}, \mathbf{P}^{-1}]$ where $\mathbf{P} = (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})$ is the *Schur Complement* of $\mathbf{A}$ [3.5]
  $=[\mathbf{I}, -\mathbf{A}^{-1}\mathbf{B}; -\mathbf{D}^{-1}\mathbf{C}, \mathbf{I}]\, \mathbf{DIAG}((\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}, (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1})$
  $=\mathbf{DIAG}((\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}, (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1})\, [\mathbf{I}, -\mathbf{B}\mathbf{D}^{-1}; -\mathbf{C}\mathbf{A}^{-1}, \mathbf{I}]$
  $=\mathbf{DIAG}(\mathbf{A}^{-1}, \mathbf{0}) + [-\mathbf{A}^{-1}\mathbf{B}; \mathbf{I}]\, (\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}[-\mathbf{C}\mathbf{A}^{-1}, \mathbf{I}]$
  $=\mathbf{DIAG}(\mathbf{0}, \mathbf{D}^{-1}) + [\mathbf{I}; -\mathbf{D}^{-1}\mathbf{C}]\, (\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}[\mathbf{I}, -\mathbf{B}\mathbf{D}^{-1}]$
  - $[\mathbf{A}, \mathbf{0}; \mathbf{C}, \mathbf{D}]^{-1} = [\mathbf{A}^{-1}, \mathbf{0}; -\mathbf{D}^{-1}\mathbf{C}\mathbf{A}^{-1}, \mathbf{D}^{-1}]$
    $=[\mathbf{I}, \mathbf{0}; -\mathbf{D}^{-1}\mathbf{C}, \mathbf{I}]\, \mathbf{DIAG}(\mathbf{A}^{-1}, \mathbf{D}^{-1})$
    $=\mathbf{DIAG}(\mathbf{A}^{-1}, \mathbf{D}^{-1})\, [\mathbf{I}, \mathbf{0}; -\mathbf{C}\mathbf{A}^{-1}, \mathbf{I}]$
  - $[\mathbf{A}, \mathbf{B}; \mathbf{C}, \mathbf{0}]^{-1} = \mathbf{DIAG}(\mathbf{A}^{-1}, \mathbf{0}) - [-\mathbf{A}^{-1}\mathbf{B}; \mathbf{I}]\, (\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}[-\mathbf{C}\mathbf{A}^{-1}, \mathbf{I}]$

- $[\mathbf{A}, \mathbf{b}; \mathbf{c}^T, d]^{-1} = [\mathbf{Q}^{-1}, -d^{-1}\mathbf{Q}^{-1}\mathbf{b}; -d^{-1}\mathbf{c}^T\mathbf{Q}^{-1}, d^{-1}(1+d^{-1}\mathbf{c}^T\mathbf{Q}^{-1}\mathbf{b})]$ where $\mathbf{Q} = (\mathbf{A}-d^{-1}\mathbf{b}\mathbf{c}^T)$,
  $= [\mathbf{A}^{-1}(\mathbf{I}+p^{-1}\mathbf{b}\mathbf{c}^T\mathbf{A}^{-1}), -p^{-1}\mathbf{A}^{-1}\mathbf{b}; -p^{-1}\mathbf{c}^T\mathbf{A}^{-1}, p^{-1}]$ where $p = (d-\mathbf{c}^T\mathbf{A}^{-1}\mathbf{b})$
  $=[\mathbf{I}, -\mathbf{A}^{-1}\mathbf{b}; -d^{-1}\mathbf{c}^T, 1]\, \mathbf{DIAG}((\mathbf{A}-d^{-1}\mathbf{b}\mathbf{c}^T)^{-1}, (d-\mathbf{c}^T\mathbf{A}^{-1}\mathbf{b})^{-1})$
  $=\mathbf{DIAG}((\mathbf{A}-d^{-1}\mathbf{b}\mathbf{c}^T)^{-1}, (d-\mathbf{c}^T\mathbf{A}^{-1}\mathbf{b})^{-1})\, [\mathbf{I}, -\mathbf{b}d^{-1}; -\mathbf{c}^T\mathbf{A}^{-1}, 1]$
  $=\mathbf{DIAG}(\mathbf{A}^{-1}, 0) + (d-\mathbf{c}^T\mathbf{A}^{-1}\mathbf{b})^{-1}[\mathbf{A}^{-1}\mathbf{b}; -1]\, [\mathbf{c}^T\mathbf{A}^{-1}, -1]$
  $=\mathbf{DIAG}(\mathbf{0}, d^{-1}) + [\mathbf{I}; -d^{-1}\mathbf{c}^T]\, (\mathbf{A}-d^{-1}\mathbf{b}\mathbf{c}^T)^{-1}[\mathbf{I}, -d^{-1}\mathbf{b}]$
  - $[\mathbf{A}, \mathbf{0}; \mathbf{c}^T, d]^{-1} = [\mathbf{A}^{-1}, \mathbf{0}; -d^{-1}\mathbf{c}^T\mathbf{A}^{-1}, d^{-1}]$
    $=[\mathbf{I}, \mathbf{0}; -d^{-1}\mathbf{c}^T, 1]\, \mathbf{DIAG}(\mathbf{A}^{-1}, d^{-1})$
    $=\mathbf{DIAG}(\mathbf{A}^{-1}, d^{-1})\, [\mathbf{I}, \mathbf{0}; -\mathbf{c}^T\mathbf{A}^{-1}, 1]$
  - $[\mathbf{A}, \mathbf{b}; \mathbf{c}^T, 0]^{-1} = \mathbf{DIAG}(\mathbf{A}^{-1}, 0) - (\mathbf{c}^T\mathbf{A}^{-1}\mathbf{b})^{-1}[\mathbf{A}^{-1}\mathbf{b}; -1]\, [\mathbf{c}^T\mathbf{A}^{-1}, -1]$

See also: Generalized Inverse, Pseudoinverse, Inversion Lemma

---

Kernel

The kernel (or null space) of $\mathbf{A}$ is the subspace of vectors $\mathbf{x}$ for which $\mathbf{Ax} = \mathbf{0}$. The dimension of this subspace is the nullity of $\mathbf{A}$.

- The kernel of $\mathbf{A}$ is the orthogonal complement of the range of $\mathbf{A}^H$

Linear Independence

The columns of $\mathbf{A}$ are *linearly independent* iff the only solution to $\mathbf{Ax}=\mathbf{0}$ is $\mathbf{x}=\mathbf{0}$.

- rank($\mathbf{A}_{[m\#n]}$) = n iff its columns are linearly independent. [1.5]
- If the columns of $\mathbf{A}_{[m\#n]}$ are linearly independent then $m >= n$ [1.3, 1.5]
- If $\mathbf{A}$ has linearly independent columns and $\mathbf{A}=\mathbf{F}_{[m\#r]}\mathbf{G}_{[r\#n]}$ then r>=n. [1.1]

Matrix Norms

A *matrix norm* is a real-valued function of a square matrix satisfying the four axioms listed below. A *generalized matrix norm* satisfies only the first three.

1. Positive: $\|\mathbf{X}\|=0$ iff $\mathbf{X}=0$ else $\|\mathbf{X}\|>0$
2. Homogeneous: $\|c\mathbf{X}\|=|c|\ \|\mathbf{X}\|$ for any real or complex scalar c
3. Triangle Inequality: $\|\mathbf{X}+\mathbf{Y}\|<=\|\mathbf{X}\|+\|\mathbf{Y}\|$
4. Submultiplicative: $\|\mathbf{XY}\|<=\|\mathbf{X}\|\ \|\mathbf{Y}\|$

Induced Matrix Norm

If $\|\mathbf{y}\|$ is a vector norm, then we define the *induced matrix norm* to be $\|\mathbf{X}\|=\max(\|\mathbf{Xy}\|$ for $\|\mathbf{y}\|=1)$

Euclidean or Frobenius Norm

The *Euclidean* or *Frobenius* norm of a matrix $\mathbf{A}$ is given by $\|\mathbf{A}\|_F = \text{sqrt}(\text{sum}(\mathbf{ABS}(\mathbf{A}).^2))$. It is always a real number. The closely related *Hilbert-Schmidt* norm of a square matrix $\mathbf{A}_{n\#n}$ is given by $\|\mathbf{A}\|_{HS} = n^{-½}\ \|\mathbf{A}\|_F$.

- $\|\mathbf{A}\|_F = \|\mathbf{A}^T\|_F = \|\mathbf{A}^H\|_F$
- $\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}^H\mathbf{A}) = \text{sum}(\text{CONJ}(\mathbf{A})\mathbf{.*}\mathbf{A})$
- [$\mathbf{Q}$: orthogonal]: $\|\mathbf{A}\|_F = \|\mathbf{QA}\|_F = \|\mathbf{AQ}\|_F$

p-Norms

$\|\mathbf{A}\|_p = \max(\|\mathbf{Ax}\|_p)$ where the max() is taken over all $\mathbf{x}$ with $\|\mathbf{x}\|_p = 1$ where $\|\mathbf{x}\|_p = \text{sum}(\mathbf{abs}(\mathbf{x})^{\bullet p})^{(1/p)}$ denotes the vector *p*-norm for $p>=1$.

- $\|\mathbf{AB}\|_p <= \|\mathbf{A}\|_p\ \|\mathbf{B}\|_p$
- $\|\mathbf{Ax}\|_p <= \|\mathbf{A}\|_p\ \|\mathbf{x}\|_p$
- [$\mathbf{A}_{:m\#n}$]: $\|\mathbf{A}\|_2 <= \|\mathbf{A}\|_F <= n^{½}\ \|\mathbf{A}\|_2$
- [$\mathbf{A}:_{m\#n}$]: $\max(\mathbf{ABS}(\mathbf{A})) <= \|\mathbf{A}\|_2 <= \text{sqrt}(mn)\max(\mathbf{ABS}(\mathbf{A}))$
- $\|\mathbf{A}\|_2 <= \text{sqrt}(\|\mathbf{A}\|_1\ \|\mathbf{A}\|_{\text{inf}})$
- $\|\mathbf{A}\|_1 = \max(\mathbf{sum}(\mathbf{ABS}(\mathbf{A}^T)))$
- $\|\mathbf{A}\|_{\text{inf}} = \max(\mathbf{sum}(\mathbf{ABS}(\mathbf{A})))$
- [$\mathbf{A}:_{m\#n}$]: $\|\mathbf{A}\|_{\text{inf}} <= \text{sqrt}(n)\ \|\mathbf{A}\|_2 <= \text{sqrt}(mn)\ \|\mathbf{A}\|_{\text{inf}}$
- [$\mathbf{A}:_{m\#n}$]: $\|\mathbf{A}\|_1 <= \text{sqrt}(m)\ \|\mathbf{A}\|_2 <= \text{sqrt}(mn)\ \|\mathbf{A}\|_1$
- [$\mathbf{Q}$: orthogonal]: $\|\mathbf{A}\|_2 = \|\mathbf{QA}\|_2 = \|\mathbf{AQ}\|_2$

Minor

A *k*th-order *minor* of $\mathbf{A}$ is the determinant of a $k\#k$ submatrix of $\mathbf{A}$.

A *principal minor* is the determinant of a submatrix whose diagonal elements lie on the principal diagonal of $\mathbf{A}$.

Null Space

The null space (or kernel) of $\mathbf{A}$ is the subspace of vectors $\mathbf{x}$ for which $\mathbf{Ax} = \mathbf{0}$.

- The null space of **A** is the orthogonal complement of the range of $\mathbf{A}^H$
- The dimension of the null space of **A** is the nullity of **A**.
- Given a vector **x**, we can choose a Householder matrix **P=I-2vv**^{H} with **v** = (**x** + k**e**$_1$)/||**x** + k**e**$_1$|| where k=sgn(x(1))*||**x**|| and **e**$_1$ is the first column of the identity matrix. The first row of **P** equals -k^{-1}**x**^{T} and the remaining rows form an orthonormal basis for the null space of **x**^{T}.

Nullity

The nullity of a matrix **A** is the dimension of the null space of **A**.

- The nullity of A is the geometric multiplicity of the eigenvalue 0.

Observability

The pair of matrices {$\mathbf{A}_{[n\#n]}$, $\mathbf{C}_{[m\#n]}$} are *observable* iff {$\mathbf{A}^H$, $\mathbf{C}^H$} are reachable.

- If {**A**, **C**} are observable then they are constructible and detectable.
- If det(**A**)!=0 and {**A**, **C**} are constructible then they are observable.

Permanent

For an $n\#n$ matrix **A**, pet(**A**) is a scalar number defined by pet(**A**)=sum(**prod**(**A**(1:n,**PERM**(n))))

This is the same as the determinant except that the individual terms within the sum are not multiplied by the signatures of the column permutations.

Properties of Permanents

- pet(**A**.') = pet(**A**)
- pet(**A**') = conj(pet(**A**))
- pet(c**A**) = c^n pet(**A**)
- [**P**: permutation matrix]: pet(**PA**) = pet(**AP**) = pet(**A**)
- [**D**: diagonal matrix]: pet(**DA**) = pet(**AD**) = pet(**A**) pet(**D**) = pet(**A**) prod(**diag**(**D**))

Permanents of simple matrices

- pet([a b; c d]) = $ad + bc$
- The permanent of a diagonal or triangular matrix is the product of its diagonal elements.
- The permanent of a permutation matrix equals 1.

Potency

The potency of a non-negative matrix **A** is the smallest n>0 such that diag($\mathbf{A}^n$) > 0 i.e. all diagonal elements of $\mathbf{A}^n$ are strictly positive. If no such n exists then **A** is impotent.

Pseudoinverse

The *pseudoinverse* (also called the *Natural Inverse* or *Moore-Penrose Pseudoinverse*) of $\mathbf{X}_{m\#n}$ is the unique [1.20] $n\#m$ matrix $\mathbf{X}^+$ that satisfies:

1. $\mathbf{XX^+X=X}$ (i.e. $\mathbf{X}^+$ is a generalized inverse of $\mathbf{X}$).
2. $\mathbf{X^+XX^+=X^+}$ (i.e. $\mathbf{X}$ is a generalized inverse of $\mathbf{X}^+$).
3. $(\mathbf{XX}^+)^H=\mathbf{XX}^+$
4. $(\mathbf{X}^+\mathbf{X})^H=\mathbf{X}^+\mathbf{X}$

- If $\mathbf{X}$ is square and non-singular then $\mathbf{X}^+=\mathbf{X}^{-1}$.
- If $\mathbf{X=UDV}^H$ is the singular value decomposition of $\mathbf{X}$, then $\mathbf{X^+=VD^+U}^H$ where $\mathbf{D}^+$ is formed by inverting all the non-zero elements of $\mathbf{D}^T$.
  - If $\mathbf{D}$ is a (not necessarily square) diagonal matrix, then $\mathbf{D}^+$ is formed by inverting all the non-zero elements of $\mathbf{D}^T$.
- The pseudoinverse of $\mathbf{X}$ is the generalized inverse having the lowest Frobenius norm.
- If $\mathbf{X}$ is real then so is $\mathbf{X}^+$.
- $(\mathbf{X}^+)^+=\mathbf{X}$
- $(\mathbf{X}^T)^+=(\mathbf{X}^+)^T$
- $(\mathbf{X}^H)^+=(\mathbf{X}^+)^H$
- $(c\mathbf{X})^+=c^{-1}\mathbf{X}^+$ for any real or complex scalar c.
- $\mathbf{X}^+=\mathbf{X}^H(\mathbf{XX}^H)^+=(\mathbf{X}^H\mathbf{X})^+\mathbf{X}^H$.
- If $\mathbf{X}_{m\#n} = \mathbf{F}_{m\#r}\,\mathbf{G}_{r\#n}$ has rank r then $\mathbf{X}^+= \mathbf{G}^+\mathbf{F}^+= \mathbf{G}^H(\mathbf{F}^H\mathbf{X}\mathbf{G}^H)^{-1}\mathbf{F}^H$.
  - If $\mathbf{X}_{m\#n}$ has rank n (i.e. the columns are linearly independent) then $\mathbf{X}^+=(\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H$ and $\mathbf{X^+X=I}$.
  - If $\mathbf{X}_{m\#n}$ has rank m (i.e. the rows are linearly independent) then $\mathbf{X}^+=\mathbf{X}^H(\mathbf{XX}^H)^{-1}$ and $\mathbf{XX^+=I}$.
  - If $\mathbf{X}$ has orthonormal rows or orthonormal columns then $\mathbf{X}^+= \mathbf{X}^H$.
- $\mathbf{XX}^+$ is a projection onto the column space of $\mathbf{X}$.
- [rank(X)=1]: $\mathbf{X}^+ = \mathbf{X}^H/\mathrm{tr}(\mathbf{X}^H\mathbf{X}) = \mathbf{X}^H/ ||\mathbf{X}||_F^2$ where $||\mathbf{X}||_F$ is the Frobenius Norm (see rank-1 matrices)
  - $(\mathbf{xy}^H)^+ = \mathbf{yx}^H/(\mathbf{x}^H\mathbf{xy}^H\mathbf{y})$
  - $\mathbf{x}^+ = \mathbf{x}^H/(\mathbf{x}^H\mathbf{x})$

See also: Inverse, Generalized Inverse

Rank

The rank of an $m\#n$ matrix $\mathbf{A}$ is the smallest r for which there exist $\mathbf{F}_{[m\#r]}$ and $\mathbf{G}_{[r\#n]}$ such that $\mathbf{A=FG}$. Such a decomposition is a *full-rank* decomposition. As a special case, the rank of $\mathbf{0}$ is 0.

- $\mathbf{A=F}_{[m\#r]}\mathbf{G}_{[r\#n]}$ implies that rank($\mathbf{A}$) <= r .
- rank($\mathbf{A}$)=1 iff $\mathbf{A} = \mathbf{xy}^T$ for some $\mathbf{x}$ and $\mathbf{y}$.
- rank($\mathbf{A}_{[m\#n]}$) <= min(m,n). [1.3]
- rank($\mathbf{A}_{[m\#n]}$) = n iff its columns are linearly independent. [1.5]
- rank($\mathbf{A}$) = rank($\mathbf{A}^T$) = rank($\mathbf{A}^H$)
- rank($\mathbf{A}$) = maximum number of linearly independent columns (or rows) of $\mathbf{A}$.
- rank($\mathbf{A}$) is the dimension of the range of $\mathbf{A}$.
- rank($\mathbf{A}_{[n\#n]}$) + nullity($\mathbf{A}_{[n\#n]}$) = n
  - rank($\mathbf{A}_{[n\#n]}$) = n - 1 if 0 is an eigenvalue of $\mathbf{A}$ with algebraic multiplicity 1.
- det($\mathbf{A}_{[n\#n]}$)=0 iff rank($\mathbf{A}_{[n\#n]}$)<n.
- rank($\mathbf{A} + \mathbf{B}$) <= rank($\mathbf{A}$) + rank($\mathbf{B}$)
- rank([$\mathbf{A}$ $\mathbf{B}$]) = rank($\mathbf{A}$) + rank($\mathbf{B} - \mathbf{AA}^\#\mathbf{B}$) where $\mathbf{A}^\#$ is a generalized inverse of $\mathbf{A}$.
  - rank([$\mathbf{A}$; $\mathbf{C}$]) = rank($\mathbf{A}$) + rank($\mathbf{C} - \mathbf{CA}^\#\mathbf{A}$)

- rank([**A B**; **C 0**]) = rank(**B**) + rank(**C**) + rank((**I** - **BB**$^{\#}$)**A**(**I** - **CC**$^{\#}$))

- rank(**AA**^{H}) = rank(**A**^{H}**A**) = rank(**A**) [see grammian]
- rank(**AB**) + rank(**BC**) <= rank(**B**) + rank(**ABC**)
  - rank($\mathbf{A}_{[m\#n]}$) + rank(**B**) - n <= rank(**AB**) <= min(rank(**A**), rank(**B**))
- [**X**: non-singular]: rank(**XA**) = rank(**AX**) = rank(**A**)
- rank(**KRON**(**A**,**B**)) = rank(**A**)rank(**B**)
- rank(**DIAG**(**A**,**B**,...,**Z**)) = sum(rank(**A**), rank(**B**), ..., rank(**Z**))

Range

The range (or image) of **A** is the subspace of vectors that equal **Ax** for some **x**. The dimension of this subspace is the rank of **A**.

- [**A**:$m\#n$] The range of **A** is the orthogonal complement of the null space of **A**^{H}.

Reachability

The pair of matrices {$\mathbf{A}_{[n\#n]}$, $\mathbf{B}_{[n\#m]}$} are *reachable* iff any of the following equivalent conditions are true

1. rank(**C**)=n where **C** = [**B AB A**^{2}**B** ... **A**$^{n-1}$**B**]$_{[n\#mn]}$ is the *controllability matrix*.
2. If **x**^{H}**A**^{r}**B** = **0** for 0<=r<n then **x** = **0**.
3. If **x**^{H}**B** = **0** and **x**^{H}**A** = k**x**^{H} then **x** = **0**.
4. For any **v**, it is possible to choose $\mathbf{L}_{[n\#m]}$ such that **eig**(**A**+**BL**^{H})=**v**.

- If {**A**, **B**} are reachable then they are controllable and stabilizable.
- If det(**A**)!=0 and {**A**, **B**} are controllable then they are reachable.
- {**DIAG**(**a**), **b**} are reachable iff all elements of **a** are distinct and all elements of **b** are non-zero.

Schur Complement

Given a block matrix **M** = [$\mathbf{A}_{[m\#m]}$, **B**; **C**, $\mathbf{D}_{[n\#n]}$], then $\mathbf{P}_{[n\#n]}$=**D**-**CA**$^{-1}$**B** and $\mathbf{Q}_{[m\#m]}$=**A**-**BD**$^{-1}$**C** are respectively the Schur Complements of **A** and **D** in **M**.

- det([**A**, **B**; **C**, **D**]) = det([**D**, **C**; **B**, **A**]) = det(**A**)*det(**P**) = det(**Q**)*det(**D**) [3.1]
- [**A**, **B**; **C**, **D**]$^{-1}$ = [**Q**$^{-1}$, -**Q**$^{-1}$**BD**$^{-1}$; -**D**$^{-1}$**CQ**$^{-1}$, **D**$^{-1}$(**I**+**CQ**$^{-1}$**BD**$^{-1}$)]= [**A**$^{-1}$(**I**+**BP**$^{-1}$**CA**$^{-1}$), -**A**$^{-1}$**BP**$^{-1}$; -**P**$^{-1}$**CA**$^{-1}$, **P**$^{-1}$] [3.5]

Spectral Radius

The *spectral radius*, rho(**A**), of $\mathbf{A}_{[n\#n]}$ is the maximum modulus of any of its eigenvalues.

- rho(**A**) <= ||**A**|| where ||**A**|| is any matrix norm.
- For any a>0, there exists a matrix norm such that ||**A**|| - a <= rho(**A**) <= ||**A**||.
- If **ABS**(**A**)<=**B** then rho(**A**)<=rho(**ABS**(**A**))<=rho(**B**)
  - [**A**,**B**: real] If **B**>=**A**>=**0** then rho(**B**)>=rho(**A**)
- [**A**: real] If **A**>=**0** then rho(**A**)>=a_{ij} for all i,j
- [**A**,**B**: Hermitian] **abs**(**eig**(**A**+**B**)-**eig**(**A**))<=rho(**B**) where **eig**(**A**) contains the eigenvalues of **A** sorted into ascending order. This shows that perturbing a hermitian matrix slightly doesn't have too big an effect on its eigenvalues.

Spectrum

The spectrum of $\mathbf{A}_{[n\#n]}$ is the set of all its eigenvalues.

Stabilizability

The pair of matrices $\{\mathbf{A}_{[n\#n]}, \mathbf{B}_{[n\#m]}\}$ are *stabilizable* iff either of the following equivalent conditions are true

1. If $\mathbf{x}^T\mathbf{B} = \mathbf{0}$ and $\mathbf{x}^T\mathbf{A} = k\mathbf{x}^T$ then either $|k|< 1$ or else $\mathbf{x} = \mathbf{0}$.
2. It is possible to choose $\mathbf{L}_{[n\#m]}$ such that all elements of $\mathbf{eig}(\mathbf{A}+\mathbf{B}\mathbf{L}^H)$ have absolute value < 1.

- If $\{\mathbf{A}, \mathbf{B}\}$ are reachable or controllable then they are stabilizable.
- $\{\mathbf{DIAG}(\mathbf{a}), \mathbf{b}\}$ are stabilizable iff all elements of **a** with modulus >=1 are distinct and all the corresponding elements of **b** are non-zero.

Submatrix

A *submatrix* of **A** is a matrix formed by the elements $a(i,j)$ where i ranges over a subset of the rows and j ranges over a subset of the columns.

Trace

The trace of a square matrix is the sum of its diagonal elements: tr(**A**)=sum(**diag**(**A**))

In the formulae below, we assume that matrix dimensions ensure that the argument of tr() is square.

- $\mathrm{tr}(a\mathbf{A}) = a \times \mathrm{tr}(\mathbf{A})$
- $\mathrm{tr}(\mathbf{A}^T) = \mathrm{tr}(\mathbf{A})$
- $\mathrm{tr}(\mathbf{A}^H) = \mathrm{tr}(\mathbf{A})^C$
- $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$
- $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ [1.17]
    - $\mathrm{tr}((\mathbf{AB})^k) = \mathrm{tr}((\mathbf{BA})^k)$
    - $\mathbf{tr}(\mathbf{ab}^T) = \mathbf{a}^T\mathbf{b}$
    - $\mathbf{tr}(\mathbf{Xba}^T) = \mathbf{a}^T\mathbf{Xb}$
    - $\mathrm{tr}(\mathbf{ab}^H) = (\mathbf{a}^H\mathbf{b})^C$
    - $\mathrm{tr}(\mathbf{ABCD}) = \mathrm{tr}(\mathbf{BCDA}) = \mathrm{tr}(\mathbf{CDAB}) = \mathrm{tr}(\mathbf{DABC})$
    - Similar matrices have the same trace: $\mathrm{tr}(\mathbf{X}^{-1}\mathbf{AX}) = \mathrm{tr}(\mathbf{A})$
- $\mathrm{tr}(\mathbf{AB}) = \mathbf{A{:}}^T\mathbf{B}^T\mathbf{{:}} = \mathbf{A}^T\mathbf{{:}}^T\mathbf{B{:}} = \mathbf{A}^H\mathbf{{:}}^H\mathbf{B{:}} = (\mathbf{A{:}}^H\mathbf{B}^H\mathbf{{:}})^C$ [1.18]
    - $\mathrm{tr}(\mathbf{A}^T\mathbf{B}) = \mathrm{tr}(\mathbf{AB}^T) = \mathrm{sum}(\mathbf{A{:}} \bullet \mathbf{B{:}}) = \mathbf{A{:}}^T\, \mathbf{B{:}}$
    - $\mathrm{tr}(\mathbf{A}^H\mathbf{B}) = \mathrm{tr}(\mathbf{BA}^H) = \mathrm{sum}(\mathbf{A}^C\mathbf{{:}} \bullet \mathbf{B{:}}) = \mathbf{A{:}}^H\, \mathbf{B{:}}$
        - $\mathrm{tr}(\mathbf{A}^H\mathbf{A}) = \mathrm{tr}(\mathbf{AA}^H) = \mathbf{A{:}}^H\, \mathbf{A{:}} = (\,\|\mathbf{A}\|_F\,)^2$ where $\|\mathbf{A}\|_F$ is the Frobenius matrix norm.
- $\mathrm{tr}([\mathbf{A}\ \mathbf{B}]^T\, [\mathbf{C}\ \mathbf{D}]) = \mathrm{tr}(\mathbf{A}^T\mathbf{C}) + \mathrm{tr}(\mathbf{B}^T\mathbf{D})$ [1.19]
    - $\mathrm{tr}([\mathbf{A}\ \mathbf{b}]^T\, [\mathbf{C}\ \mathbf{d}]) = \mathrm{tr}(\mathbf{A}^T\mathbf{C}) + \mathbf{b}^T\mathbf{d}$
    - $\mathrm{tr}([\mathbf{A}\ \mathbf{B}]^T\, \mathbf{X}[\mathbf{C}\ \mathbf{D}]) = \mathrm{tr}(\mathbf{A}^T\mathbf{XC}) + \mathrm{tr}(\mathbf{B}^T\mathbf{XD})$
        - $\mathrm{tr}([\mathbf{A}\ \mathbf{b}]^T\, \mathbf{X}[\mathbf{C}\ \mathbf{d}]) = \mathrm{tr}(\mathbf{A}^T\mathbf{XC}) + \mathbf{b}^T\mathbf{Xd}$
- $\mathrm{tr}(\mathbf{A} \otimes \mathbf{B}) =$ [**A**,**B**: $n\#n$] $\mathrm{tr}(\mathbf{A})\, \mathrm{tr}(\mathbf{B})$ where $\otimes$ denotes the Kroneker product.
- [D is diagonal] $\mathrm{tr}(\mathbf{XDX}^T) = \mathrm{sum}_i(d_i\ \ \mathbf{x}_i^T\mathbf{x}_i)$ and $\mathrm{tr}(\mathbf{XDX}^H) = \mathrm{sum}_i(d_i\ \ \mathbf{x}_i^H\mathbf{x}_i) = \mathrm{sum}_i(d_i\ \ |\mathbf{x}_i|^2)$ [1.16]

Transpose

$\mathbf{X}=\mathbf{Y}^T$ is the transpose of $\mathbf{Y}$ iff $x(i,j)=y(j,i)$.

Vectorization

The vector formed by concatenating all the columns of $\mathbf{X}$ is written **vec**($\mathbf{X}$) or, in this website, $\mathbf{X}$**:**. If $\mathbf{y} = \mathbf{X}_{[m\#n]}$**:** then $y_{i+m(j-1)} = x_{i,j}$.

- $\mathbf{a} \otimes \mathbf{b}=(\mathbf{ba}^T)$**:** where $\otimes$ denotes the Kroneker product.
- sum(($\mathbf{A} \bullet \mathbf{B}$)**:**) = tr($\mathbf{A}^T\mathbf{B}$) = sum($\mathbf{A}$**:** $\bullet$ $\mathbf{B}$**:**) = $\mathbf{A}$**:**^{T} $\mathbf{B}$**:** = ($\mathbf{A}^T$**:**)^T $\mathbf{B}^T$**:** where $\mathbf{A} \bullet \mathbf{B}$ denotes the Hadamard or elementwise product.
- tr($\mathbf{A}^H\mathbf{B}$) = sum($\mathbf{A}^C$**:** $\bullet$ $\mathbf{B}$**:**) = $\mathbf{A}$**:**^{H} $\mathbf{B}$**:**
  - [**A**, **B** Hermitian] tr($\mathbf{A}^H\mathbf{B}$) = tr($\mathbf{B}^H\mathbf{A}$) = $\mathbf{A}$**:**^{H} $\mathbf{B}$**:** = $\mathbf{B}$**:**^{H} $\mathbf{A}$**:** is real-valued.
- ($\mathbf{ABC}$)**:** = ($\mathbf{C}^T \otimes \mathbf{A}$) $\mathbf{B}$**:**
  - ($\mathbf{AB}$)**:** = ($\mathbf{I} \otimes \mathbf{A}$) $\mathbf{B}$**:** = ($\mathbf{B}^T \otimes \mathbf{I}$) $\mathbf{A}$**:**= ($\mathbf{B}^T \otimes \mathbf{A}$) $\mathbf{I}$**:**
  - ($\mathbf{Abc}^T$)**:** = ($\mathbf{c} \otimes \mathbf{A}$) $\mathbf{b}$ = $\mathbf{c} \otimes \mathbf{Ab}$
  - $\mathbf{ABc}$ = ($\mathbf{c}^T \otimes \mathbf{A}$) $\mathbf{B}$**:**
  - $\mathbf{a}^T\mathbf{Bc}$ = ($\mathbf{c} \otimes \mathbf{a}$)^T $\mathbf{B}$**:** = ($\mathbf{c}^T \otimes \mathbf{a}^T$) $\mathbf{B}$**:** = ($\mathbf{ac}^T$)**:**^{T} $\mathbf{B}$**:** = $\mathbf{B}$**:**^{T} ($\mathbf{a} \otimes \mathbf{c}$) = $\mathbf{B}$**:**^{T} ($\mathbf{ca}^T$)**:**
  - $\mathbf{ab}^H \otimes \mathbf{cd}^H$ = ($\mathbf{a} \otimes \mathbf{c}$)($\mathbf{b} \otimes \mathbf{d}$)^H = ($\mathbf{ca}^T$)**:**($\mathbf{db}^T$)**:**^{H}
  - $\mathbf{a}^H\mathbf{bc}^H\mathbf{d}$ = $\mathbf{a}^H\mathbf{b} \otimes \mathbf{c}^H\mathbf{d}$ = ($\mathbf{a} \otimes \mathbf{c}$)^H($\mathbf{b} \otimes \mathbf{d}$) = ($\mathbf{ca}^T$)**:**^{H}($\mathbf{db}^T$)**:**
- ($\mathbf{ABC}$)**:**^{T} = $\mathbf{B}$**:**^{T} ($\mathbf{C} \otimes \mathbf{A}^T$)
  - ($\mathbf{AB}$)**:**^{T} = $\mathbf{B}$**:**^{T} ($\mathbf{I} \otimes \mathbf{A}^T$) = $\mathbf{A}$**:**^{T} ($\mathbf{B} \otimes \mathbf{I}$) = $\mathbf{I}$**:**^{T} ($\mathbf{B} \otimes \mathbf{A}^T$)
  - ($\mathbf{Abc}^T$)**:**^{T} = $\mathbf{b}^T$($\mathbf{c}^T \otimes \mathbf{A}^T$) = $\mathbf{c}^T \otimes \mathbf{b}^T\mathbf{A}^T$
  - $\mathbf{a}^T\mathbf{B}^T\mathbf{C}$ = $\mathbf{B}$**:**^{T} ($\mathbf{a} \otimes \mathbf{C}$)
- If $\mathbf{Y}$=$\mathbf{AXB}$+$\mathbf{CXD}$+... then $\mathbf{X}$**:** = ($\mathbf{B}^T \otimes \mathbf{A}$ + $\mathbf{D}^T \otimes \mathbf{C}$+...)$^{-1}$ $\mathbf{Y}$**:** however this is a slow and often ill-conditioned way of solving such equations.
- ($\mathbf{A}_{[m\#n]}{}^T$)**:** = **TVEC**(m,n) ($\mathbf{A}$**:**) [see vectorized transpose]

Vector Norms

A *vector norm* is a real-valued function of a vector satisfying the three axioms listed below.

1. Positive: $||\mathbf{x}||=0$ iff $\mathbf{x}=0$ else $||\mathbf{x}||>0$
2. Homogeneous: $||c\mathbf{x}||=|c|\ ||\mathbf{x}||$ for any real or complex scalar c
3. Triangle Inequality: $||\mathbf{x}+\mathbf{x}||<=||\mathbf{x}||+||\mathbf{x}||$

Inner Product Norm

If $<\mathbf{x}, \mathbf{y}>$ is an inner product then $||\mathbf{x}|| = \mathrm{sqrt}(<\mathbf{x}, \mathbf{x}>)$ is a vector norm.

- A vector norm may be derived from an inner product iff it satisfies the *parallelogram identity*: $||\mathbf{x}+\mathbf{y}||^2+||\mathbf{x}-\mathbf{y}||^2=2||\mathbf{x}||^2+2||\mathbf{y}||^2$
- If $||\mathbf{x}||$ is derived from $<\mathbf{x}, \mathbf{y}>$ then $4\mathrm{Re}(<\mathbf{x}, \mathbf{y}>) = ||\mathbf{x}+\mathbf{y}||^2-||\mathbf{x}-\mathbf{y}||^2 = 2||\mathbf{x}+\mathbf{y}||^2-||\mathbf{x}||^2-||\mathbf{y}||^2$

Euclidean Norm

The Euclidean norm of a vector $\mathbf{x}$ equals the square root of the sum of the squares of the absolute values of all its elements and is written $||\mathbf{x}||$. It is always a real number and corresponds to the normal notion of the vector's length.

- $||\mathbf{x}||^2 = \mathbf{x}^H\mathbf{x} = \mathrm{tr}(\mathbf{x}\mathbf{x}^H)$
- Cauchy-Schwartz inequality: $|\mathbf{x}^H\mathbf{y}| <= ||\mathbf{x}||\ ||\mathbf{y}||$
- [**Q**: orthogonal]: $||\mathbf{Q}\mathbf{x}|| = ||\mathbf{x}||$

Hölder Norms or p-Norms

The p-norm of a vector **x** is defined by $||\mathbf{x}||_p = \mathrm{sum}(\mathbf{abs}(\mathbf{x})^{\bullet p})^{(1/p)}$ for p>=1. The most common values of p are 1, 2 and infinity.

- City-Block Norm: $||\mathbf{x}||_1 = \mathrm{sum}(\mathbf{abs}(\mathbf{x}))$
- Euclidean Norm: $||\mathbf{x}|| = ||\mathbf{x}||_2 = \mathrm{sqrt}(\mathbf{x}'\mathbf{x})$
- Infinity Norm: $||\mathbf{x}||_{inf} = \max(\mathbf{abs}(\mathbf{x}))$
- Hölder inequality: $\mathbf{abs}(\mathbf{x})^T\mathbf{abs}(\mathbf{y}) <= ||\mathbf{x}||_p\ ||\mathbf{y}||_q$ where $1/p + 1/q = 1$
- $||\mathbf{x}||_{inf} <= ||\mathbf{x}||_2 <= ||\mathbf{x}||_1 <= \mathrm{sqrt}(n)\ ||\mathbf{x}||_2 <= n\ ||\mathbf{x}||_{inf}$
