Handbook on matrices

Andriy Zhugayevych (http://zhugayevych.me)

July 28, 2022

§1. Conventions

Capitals A are matrices, normal symbols a are vectors, and Greek symbols α are scalars. Upper index enumerates vectors and matrices while lower indexes enumerates their elements. For a matrix component A_{ij} i means row and j means column. Throughout this paper we denote n to be the dimension of matrices and vectors. We denote identity matrix simply by 1 so that $\lambda - A$ means $(\lambda - A)_{ij} = \lambda\delta_{ij} - A_{ij}$. Transposed matrix is denoted by $A^\top$: $(A^\top)_{ij} = A_{ji}$. Hermitian conjugated matrix is denoted by A^+: $(A^+)_{ij} = \overline{A_{ji}}$.

There are two kinds of linear transformations. A linear map given by a nondegenerate matrix T transforms $x \to Tx$ so that $(Ax) \to T(Ax)$ and thus $A \to TAT^{-1}$. Two matrices are similar (or equivalent) if there exists a transformation mapping one matrix into another. The second kind of linear transformations is a *basis change* (basis is an outer object for the matrix theory). Let a new basis is given by $e'_i = \sum_j e_j B_{ji}$, where B is a nondegenerate matrix. Then a vector defined as $x = \sum_i x_i e_i$ and a matrix defined by the scalar product $(x, Ay) = \sum_{ij} \overline{x}_i A_{ij} y_j$ transform as follows:

$$x' = B^{-1}x, \qquad A' = B^+AB, \tag{1.1}$$

so that Hermitian matrices remain Hermitian. In particular the basis overlap matrix $S_{ij} = (e_i, e_j)$ transforms as $S' = B^+SB$. Both kinds of transformations are equivalent iff the transformation matrix is unitary.

We call matrix A reducible if there is such a partition of its index set $\{1,\dots,n\} = \Xi \cup \Gamma$ ($\Xi \cap \Gamma = \emptyset$) that $\forall \xi \in \Xi\ \forall \gamma \in \Gamma\ A_{\xi\gamma} = 0$. If in addition $A_{\gamma\xi} = 0$ than the matrix A is decomposable. Thus the natural chain is {decomposable}⊃{indecomposable reducible}⊃{irreducible}.

As default real symmetric matrices are symbolized by S, Hermitian by H, real orthogonal by O or R, unitary by U or V, projective by P.

§2. Vector spaces

The Gram matrix of a set of vectors $\{x_1,\dots,x_n\}$ is defined by its elements to be (x_i,x_j). The determinant of this matrix, the Gram determinant Γ, has the following properties: 1) $\Gamma \geqslant 0$ and $\Gamma \neq 0$ iff $\{x_1,\dots,x_n\}$ is linear independent; 2) $\sqrt{\Gamma}$ gives the volume of m-dimensional parallelepiped constructed on $\{x_1,\dots,x_n\}$; 3) Hadamard's inequality asserts $\Gamma[x_1,\dots,x_n] \leqslant \Gamma[x_1]\dots\Gamma[x_n]$ and the equality is only for orthogonal system of vectors.

A finite set of linear independent vectors is called basis. Any basis can be orthogonalized.

For any basis $\{e_1,\dots,e_n\}$ there exists a dual basis $\{e^1,\dots,e^n\}$ such that $(e^i,e_j) = \delta_{ij}$. The metric tensor is defined as $g_{ij} = (e_i,e_j)$, so that $g^{ij} = (e^i,e^j)$, $g_i^j = \delta_{ij}$, $e^i = \sum_j g^{ij} e_j$ and $e_i = \sum_j g_{ij} e^j$. Let G be a matrix with the elements $G_{ij} = g_{ij}$, then $\left(G^{-1}\right)_{ij} = g^{ij}$. The basis $\{e_1,\dots,e_n\}$ can be represented as the columns of a matrix T: $T_{\alpha j} = (e_j)_\alpha$, so that $T^\top T = G$. If $\tilde{T}_{i\beta} = (e^i)_\beta$ then $\tilde{T}T = 1$. If the basis is complete, then for any vector x, $x = \sum_i (e^i,x)e_i = \sum_i (e_i,x)e^i$. If the basis is incomplete, then the best expansion $x = \sum_i c_i e_i + u$ minimizes the basis deficiency $||u||^2$. In L_2 norm the solution is given by $c_i = (e^i,x)$ with $||u||^2 = 1 - \sum_i (e^i,x)(e_i,x)$.

§3. Basic matrix theory

Some notions. Let there is n-dimensional matrix A.

Minor M of an order p is the determinant of the matrix obtained from A by deleting $n-p$ rows and columns. The *principal* minors are the "diagonal" minors. The (i,j) *cofactor* α_{ij} of the element A_{ij} is the coefficient at A_{ij} in the determinant of A . Note that $\alpha_{ij} = (-1)^{i+j} M_{ij}$, where M_{ij} is the minor of the order $n-1$ obtained by deleting i-th row and j-th column. Thus the well known formula for the determinant arises. The *determinant* is

$$\det A = \sum_{i_1,\dots,i_n=1}^{n} e^{1\dots n}_{i_1\dots i_n} A_{i_1 1}\dots A_{i_n n} = \sum_{j=1}^{n} A_{1j}\alpha_{1j},$$

where e is absolute antisymmetric tensor, and the *permanent* is the absolutely symmetric sum. The *trace* is $\operatorname{tr} A = \sum_{i=1}^n A_{ii}$.

3.1. Types of matrices

The matrix A is *symmetric* (*Hermitian*) if $A^\top = A$ ($A^+ = A$), *antisymmetric* (*skew Hermitian*) if $A^\top = -A$ ($A^+ = -A$), *orthogonal* (*unitary*) if $A^\top = A^{-1}$ ($A^+ = A^{-1}$), *normal* if $A^+A = AA^+$, *nilpotent* if $A^m = 0$, *projective* if $A^2 = A$, *involutive* if $A^2 = 1$. Other types include diagonal, band, block, upper and lower triangular matrices.

Note that unitary matrices correspond to isometric operators and projective matrices to projectors. Hermitian projective matrices corresponds to orthogonal projectors.

Any matrix A can be uniquely presented via the Hermitian matrices by $A = H + \mathrm{i}H'$.

Any matrix can be decomposed into the product of two symmetric matrices anyone of which can be chosen nondegenerate.

Any matrix can be decomposed into the sum of two mutually commuting diagonalizable and nilpotent matrices.

3.2. Characteristic matrix $\lambda - A$

The *characteristic polynomial* is

$$\Delta(\lambda) = \det(\lambda - A) = \lambda^n - p_1\lambda^{n-1} - \ldots - p_n.$$

Note that $p_1 = \operatorname{tr} A$ and $p_n = (-1)^{n-1}\det A$.

The *adjoint* matrix $B(\lambda)$ is defined by its elements $B_{ij}(\lambda)$ equal to cofactor of the element $\lambda\delta_{ji} - A_{ji}$ of the matrix $\lambda - A$. The adjoint matrix can be evaluated by the formula $B(\lambda) = \lambda^{n-1} + B_1\lambda^{n-2} + \ldots + B_{n-1}$, where $B_k = AB_{k-1} - p_k$ and $B_0 = 1$. Note that $B_n = 0$ and $B_{n-1} = p_n A^{-1}$.

Let $D(\lambda)$ be the least common divisor of all $B_{ij}(\lambda)$. The *reduced adjoint* matrix is defined as $C(\lambda) = B(\lambda)/D(\lambda)$.

Note that $\Delta(A) = 0$. The *minimal polynomial* $\psi(\lambda)$ is the annihilating polynomial (i.e. $\psi(A) = 0$) of the lowest degree with the leading coefficient equal unity. It equals to $\Delta(\lambda)/D(\lambda)$.

There are some useful relations:

$$(\lambda - A)B(\lambda) = B(\lambda)(\lambda - A) = \Delta(\lambda),$$
$$(\lambda - A)C(\lambda) = C(\lambda)(\lambda - A) = \psi(\lambda),$$
$$C(\lambda) = \Psi(\lambda, A), \text{ where } \Psi(\lambda,\mu) = \frac{\psi(\lambda) - \psi(\mu)}{\lambda - \mu} \text{ is the polynomial(!).}$$

3.3. Spectrum

The *eigenvalues* λ and corresponding *eigenvectors* h are the nontrivial solutions of the equation $Ah = \lambda h$. Let there are s different eigenvalues λ_σ of algebraic multiplicity m_σ and r eigenvectors h_ρ with $m_\rho - 1$ adjoint eigenvectors. Obviously $1 \leqslant s \leqslant r \leqslant n$ and for each λ_σ the geometric multiplicity m_σ^{g} that is the number of corresponding eigenvectors satisfies the inequality $1 \leqslant m_\sigma^{\mathrm{g}} \leqslant m_\sigma$. A matrix is *nondefective* if all the geometric multiplicities are equal to the algebraic ($m_\sigma^{\mathrm{g}} = m_\sigma \iff r = n$), otherwise it is *defective*. A matrix is *simple* if all the geometric multiplicities are equal to the unity ($m_\sigma^{\mathrm{g}} = 1 \iff s = r$). *Simple nondefective* matrix has all the eigenvalues distinct and thus $s = r = n$.

The eigenvalues are the continuous functions of matrix elements in the complex domain, though the eigenvectors are generally not. Moreover for any matrix A and $\forall\varepsilon > 0\ \exists A(\varepsilon)\ \|A - A(\varepsilon)\| < \varepsilon$ and all the eigenvalues of $A(\varepsilon)$ are distinct.

Gershgorin circle theorem: For a matrix A denote $\rho_i = \sum_{j\neq i}|A_{ij}|$, then in every circle $|\lambda - A_{ii}| \leqslant \rho_i$, $i = \overline{1,n}$ there is minimum one eigenvalue.

Matrix $R_A(\lambda) = (\lambda - A)^{-1}$ is the *resolvent* of the matrix A. The *spectral radius* $\rho(A)$ is maximum $|\lambda(A)|$. For any matrix A and $\forall\varepsilon > 0\ \exists c\ \forall i,j,k\ |(A^k)_{ij}| \leqslant c(\rho(A) + \varepsilon)^k$.

Note that $\sum_{i=1}^n \lambda_i^k = \operatorname{tr}(A^k)$.

Hermitian matrix has real eigenvalues (moreover the product of two Hermitian matrices, one of which is positively defined, has real eigenvalues). Unitary matrix has $|\lambda| = 1$ and orthogonal has $\lambda = \mathrm{e}^{\pm\mathrm{i}\phi}$.

For Hermitian matrices let us assume their eigenvalues be arranged in ascending order $\lambda_1 \leqslant \ldots \leqslant \lambda_n$. If we arrange the diagonal elements of the Hermitian matrix H and denote them d_k then $\sum_{i=1}^k \lambda_i(H) \leqslant \sum_{i=1}^k d_i$ for any k. For any two Hermitian matrices $\sum_{i=1}^k(\lambda_i(H) + \lambda_i(H')) \leqslant \sum_{i=1}^k \lambda_i(H + H')$ and $\lambda_k(H) + \lambda_1(H') \leqslant \lambda_k(H + H') \leqslant \lambda_k(H) + \lambda_n(H')$ for any k.

If H is Hermitian and depends on a set of parameters indexed by i, j, then the derivatives of the eigenvalues of H are given by

$$\partial_i\lambda = \langle h|\partial_i H|h\rangle, \qquad \partial_i\partial_j\lambda = \langle h|\partial_i\partial_j H|h\rangle + 2\sum_{\lambda'\neq\lambda}\frac{\langle h|\partial_i H|h'\rangle\langle h'|\partial_j H|h\rangle}{\lambda - \lambda'}, \tag{3.1}$$

where h (h') is the eigenvector corresponding to the eigenvalue λ (λ').

§4. Canonical forms of matrices

4.1. Diagonal form

A matrix is similar to some diagonal matrix iff it is nondefective. A matrix is unitary (orthogonally) similar to some diagonal matrix iff it is normal (real normal). A matrix has complete orthonormal set of eigenvectors iff it is normal. Two Hermitian matrices can be diagonalized simultaneously iff they commute. Any set of mutually commuting normal matrices has a common orthonormal basis in which all the matrices are diagonal.

The canonical form of a Hermitian matrix H is the diagonal matrix $D_H = U^+HU$ of its eigenvalues. All the eigenvectors can be chosen mutually orthogonal, and any two eigenvectors corresponding to different eigenvalues are always orthogonal. Therefore the transformation matrix U can be chosen unitary, containing the orthonormal eigenvectors of H as its columns.

Generalized eigenvalue problem is formulated as $Hh = \lambda Sh$, where S is a Hermitian positive-defined matrix (basis overlap). It can be reduced to ordinary eigenvalue problem $\tilde{H}\tilde{h} = \lambda\tilde{h}$ by the transformation $\tilde{h} = L^+h$ and $\tilde{H} = L^{-1}H(L^{-1})^+$ where L is the lower triangular matrix in Cholesky LU-decomposition of the overlap matrix, $S = LL^+$. If H is Hermitian, the eigenvectors are orthogonal wrt scalar product $(h_1, h_2) = h_1^+Sh_2$. The left eigenvectors are simply conjugated right vectors. If T is the matrix of eigenvectors of H as its columns then the following identities hold:

$$HT = ST\Lambda, \quad T^+ST = 1, \quad T^+HT = \Lambda, \quad TT^+S = 1, \quad ST\Lambda T^+S = H. \tag{4.1}$$

The generalized eigenvalue problem is invariant under basis change Eq. (1.1), i.e. $H'T' = S'T'\Lambda$ where $T' = B^{-1}T$.

Canonical form of the diagonal block for the complex conjugated eigenvalues is

$$\begin{pmatrix} \Re\lambda & -\Im\lambda \\ \Im\lambda & \Re\lambda \end{pmatrix}.$$

Canonical form of an orthogonal matrix is block diagonal matrix with elements

$$(\pm 1) \text{ and } \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

The key parameter of an orthogonal matrix is its determinant equal to ± 1. Note that hyperbolic rotations can be described by the blocks

$$\begin{pmatrix} \cosh\psi & \sinh\psi \\ \sinh\psi & \cosh\psi \end{pmatrix}.$$

The determinant is still ± 1 but the inverse matrix does not coincide with the transposed.

Canonical form of a skew Hermitian matrix is block diagonal matrix with elements

$$(0) \text{ and } \begin{pmatrix} 0 & -v \\ v & 0 \end{pmatrix}.$$

4.2. Jordan normal form

Jordan block matrix is a block-diagonal matrix, each Jordan block of an order m has the form

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ 0 & 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda \end{pmatrix}.$$

Any matrix A is uniquely similar to its Jordan form J_A so that $A = TJ_AT^{-1}$,

$$J_A = \begin{pmatrix} J_{m_1}(\lambda_1) & 0 & \dots & 0 \\ 0 & J_{m_2}(\lambda_2) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & J_{m_s}(\lambda_s) \end{pmatrix}, \quad \sum_{\sigma=1}^{s} m_\sigma = n.$$

However, the Jordan form is not a continuous function of the matrix elements. Each Jordan block $J_m(\lambda)$ corresponds to the minimal irreducible invariant subspace of the matrix A. The basis of this subspace are the eigenvector $h(\lambda)$ and $m-1$ *adjoined vectors* h^k. This basis is unique and can be determined from the set of equations

$$(A-\lambda)h = 0, \quad (A-\lambda)h^k = h^{k-1}, \quad k = \overline{1, m-1}, \quad h^0 \equiv h.$$

The transformation matrix is constructed of all eigenvectors and adjoined vectors as its columns:

$$T = \begin{pmatrix} h_1(\lambda_1) & h_1^1(\lambda_1) & \dots & h_1^{m_1}(\lambda_1) & h(\lambda_2) & h^1(\lambda_2) & \dots \\ h_2(\lambda_1) & h_2^1(\lambda_1) & \dots & h_2^{m_1}(\lambda_1) & h_2(\lambda_2) & h_2^1(\lambda_2) & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ h_n(\lambda_1) & h_n^1(\lambda_1) & \dots & h_2^{m_1}(\lambda_1) & h_n(\lambda_2) & h_n^1(\lambda_2) & \dots \end{pmatrix}. \tag{4.2}$$

Note that T^{-1} has the left eigenvectors $g(\lambda)$ and adjoint vectors g^k as its rows but in inverse order:

$$T^{-1} = \begin{pmatrix} g_1^{m_1}(\lambda_1) & g_2^{m_1}(\lambda_1) & \dots & g_n^{m_1}(\lambda_1) \\ g_1^{m_1-1}(\lambda_1) & g_2^{m_1-1}(\lambda_1) & \dots & g_n^{m_1-1}(\lambda_1) \\ \dots & \dots & \dots & \dots \\ g_1^2(\lambda_1) & g_2^2(\lambda_1) & \dots & g_n^2(\lambda_1) \\ g_1(\lambda_1) & g_2(\lambda_1) & \dots & g_n(\lambda_1) \\ g_1^{m_2}(\lambda_2) & g_2^{m_2}(\lambda_2) & \dots & g_n^{m_2}(\lambda_2) \\ \dots & \dots & \dots & \dots \end{pmatrix}. \tag{4.3}$$

This also implies that left and right vectors are bi-orthogonal.

For real matrix A the real Jordan form can be defined with the following real Jordan blocks corresponding to two complex Jordan blocks $J_m(\lambda)$ and $J_m(\overline{\lambda})$:

$$J_{2m}(\lambda, \overline{\lambda}) = \begin{pmatrix} \begin{pmatrix} \Re\lambda & -\Im\lambda \\ \Im\lambda & \Re\lambda \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \dots \\ \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \begin{pmatrix} \Re\lambda & -\Im\lambda \\ \Im\lambda & \Re\lambda \end{pmatrix} & \dots \\ \dots & \dots & \dots \end{pmatrix}.$$

The transformation matrix T gets the form

$$T = \begin{pmatrix} \Re h(\lambda_1) & -\Im h(\lambda_1) & \Re h^1(\lambda_1) & -\Im h^1(\lambda_1) & \dots \end{pmatrix}.$$

If one want to reduce a matrix to its Jordan form partially that is only on the basis of some selected eigenspaces then one proceeds as follows. Let the selected eigenspaces are defined by the rectangular matrix $H_{n\times m}$ which is some closed part of the matrix (4.2) (closed means that all the eigenvectors are selected with their adjoint vectors). Let the matrix $G_{m\times n}$ be the corresponding selection from the matrix (4.3). It can be constructed by the right eigenvectors and adjoint vectors providing that they are properly normalized that is they are surely orthogonal (GH is diagonal) so one has only to normalize them to get $GH = 1$. Let $E_{n\times(n-m)}$ be any rectangular matrix such that its columns complement the columns of H to the complete basis, and let $F_{n\times(n-m)}$ complements G to the complete basis in the dual space (which is actually the same space). Then the transformation matrix that reduces the first m coordinates to the Jordan form is given by

$$T = \begin{pmatrix} H & PE \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} G \\ (FPE)^{-1}FP \end{pmatrix}, \tag{4.4}$$

where $P \equiv 1 - HG$ is the projector onto orthogonal complement to the selected eigenspaces and FPE is the square matrix of the dimension $(n-m)$. The reduced matrix is

$$T^{-1}AT = \begin{pmatrix} J_{m\times m} & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & (FPE)^{-1}FPAPE \end{pmatrix}.$$

In particular, if A is Hermitian then

$$T = \begin{pmatrix} H & K \end{pmatrix}, \quad T^{+} = \begin{pmatrix} H^{+} \\ K^{+} \end{pmatrix}, \quad T^{+}AT = \begin{pmatrix} H^{+}AH & 0 \\ 0 & K^{+}AK \end{pmatrix}, \tag{4.5}$$

provided that T is unitary.

4.3. Singular value decomposition and pseudoinverse

In this subsection M is rectangular $n \times m$ matrix. The matrix M^+M is Hermitian non-negatively definite and the square roots of its eigenvalues are the *singular values* of the matrix M. For square matrices the ratio of the greatest and least singular values gives the Euclidean (square norm) *condition number* of the matrix.

Singular value decomposition: any matrix M can be decomposed into the product $M = U\Sigma V^+$, where Σ is the $n \times m$ matrix whose diagonal contains the singular values of M and all other entries are zero, U is the $n \times n$ and V is the $m \times m$ unitary matrices. If $n = m$ then $M = \tilde{U}H$, where $H = V\Sigma V^+$ is Hermitian and $\tilde{U} = UV^+$ is unitary matrices.

The *Frobenius norm* of a matrix M is defined as follows:

$$||M||_{\mathrm{F}}^2 = \sum_{i,j} |M_{ij}|^2 \equiv \mathrm{tr}(M^+M) \equiv \sum_k \sigma_k^2, \tag{4.6}$$

where σ are singular values of M.

Pseudoinverse matrix M^{-1} is the least squares solution of the equation $MX = 1$ in the sense that it minimizes the Euclidean (square) norm of the matrix $1 - M^{-1}M$ (and simultaneously the norm of $1 - MM^{-1}$). For the case of degenerate matrices this definition is ambiguous, that is there are p-parametric class of pseudoinverse matrices, where p is the order of degeneracy. Pseudoinverse matrix can be evaluated by the formula $M^{-1} = V\Sigma^{-1}U^+$, where the Σ^{-1} is the $m \times n$ matrix whose diagonal contains the inverse singular values of M (if the singular value is zero then one can choose any value, e.g. zero) and all other entries are zero. The matrix $1 - M^{-1}M$ $(1 - MM^{-1})$ is the orthogonal projector onto the right (left) $\ker M$.

If M is a rectangular matrix whose columns are linearly independent, then $P = M(M^+SM)^{-1}M^+S$ is the orthogonal (wrt to S) projector onto M, i.e. $P^2 = P$, $P.M = M$, and $M'SM = 0$ where $M' = \mathrm{Span}(1 - P)$.

4.4. Other forms

Any matrix A is unitary similar to its *Schur form* T_A so that $A = UT_AU^+$, where T_A is the upper triangular matrix. Though the superdiagonal elements are determined not uniquely, the sum $\sum_{i<j} |T_{ij}|^2$ is determined uniquely. Any real matrix A is orthogonally similar to its real Schur form T_A so that $A = OT_AO^\top$, where T_A is the block upper triangular matrix with standard 2×2 blocks for each complex conjugated eigenvalues.

§5. Matrix functions

5.1. Using Jordan normal form

For any differentiable function f the matrix $f(At) = Tf(J_At)T^{-1}$ can be evaluated by using the formula

$$f\left[J_m(\lambda)t\right] = \begin{pmatrix} f(\lambda t) & f'(\lambda t)t & f''(\lambda t)\frac{t^2}{2} & \dots & f^{(m-1)}(\lambda t)\frac{t^{m-1}}{(m-1)!} \\ 0 & f(\lambda t) & f'(\lambda t)t & \dots & f^{(m-2)}(\lambda t)\frac{t^{m-2}}{(m-2)!} \\ 0 & 0 & f(\lambda t) & \dots & f^{(m-3)}(\lambda t)\frac{t^{m-3}}{(m-3)!} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & f(\lambda t) \end{pmatrix}. \tag{5.1}$$

Any function f applied to the real Jordan block yields (5.1) but with

$$f(\lambda t) \to \begin{pmatrix} \Re f(\lambda t) & -\Im f(\lambda t) \\ \Im f(\lambda t) & \Re f(\lambda t) \end{pmatrix}, \qquad f'(\lambda t) \to \dots$$

For an example of exponential function one get

$$\exp\left[J_{2m}(\lambda, \overline{\lambda})t\right] = \mathrm{e}^{\Re\lambda t} \begin{pmatrix} \cos\Im\lambda t & -\sin\Im\lambda t \\ \sin\Im\lambda t & \cos\Im\lambda t \end{pmatrix} \times \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{m-1}}{(m-1)!} \\ 0 & 1 & t & \dots & \frac{t^{m-2}}{(m-2)!} \\ 0 & 0 & 1 & \dots & \frac{t^{m-3}}{(m-3)!} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix},$$

where the multiplication sign “×” means the direct product.

Another example is the resolvent $(s-A)^{-1}$:

$$(s-J_m(\lambda))^{-1}=\frac{1}{s-\lambda}+\frac{N_1}{(s-\lambda)^2}+\ldots+\frac{N_{m-1}}{(s-\lambda)^m}, \tag{5.2}$$

where N_μ is m-dimensional matrix which has 1 on its μ-th superdiagonal (so that N_μ is nilpotent of $m-\mu$ order).

5.2. Interpolating polynomial

From the Jordan form of function of matrix one can conclude that any differentiable function of matrix can be evaluated via the so called Lagrange–Sylvester interpolating polynomial:

$$f(A)=\sum_{\sigma=1}^{s}\frac{1}{(m_\sigma-1)!}\left[\frac{C(\lambda)}{\psi^\sigma(\lambda)}f(\lambda)\right]^{(m_\sigma-1)}\Bigg|_{\lambda=\lambda_\sigma}\equiv\sum_{\sigma=1}^{s}\left(f(\lambda_\sigma)Z_{\sigma1}+f'(\lambda_\sigma)Z_{\sigma2}+\ldots+\frac{1}{(m_\sigma-1)!}f^{(m_\sigma-1)}(\lambda_\sigma)Z_{\sigma m_\sigma}\right) \tag{5.3}$$

where

$$Z_{\sigma j}=\frac{1}{(m_\sigma-j)!}\left[\frac{C(\lambda)}{\psi^\sigma(\lambda)}\right]^{(m_\sigma-j)}\Bigg|_{\lambda=\lambda_\sigma},$$

σ enumerates eigenvalues, and m_σ is the algebraic multiplicity of λ_σ. In particular, the resolvent

$$(\lambda-A)^{-1}=\sum_{\sigma=1}^{s}\left(\frac{Z_{\sigma1}}{\lambda-\lambda_\sigma}+\frac{Z_{\sigma2}}{(\lambda-\lambda_\sigma)^2}+\ldots+\frac{Z_{\sigma m_\sigma}}{(\lambda-\lambda_\sigma)^{m_\sigma}}\right). \tag{5.4}$$

By comparing this expression with (5.2) one can obtain that

$$Z_{\sigma\mu}=\sum_{\lambda=\lambda_\sigma}\left(\frac{h^1\times g^{m-\mu+1}}{g^{m-\mu+1}h^1}+\frac{h^2\times g^{m-\mu}}{g^{m-\mu}h^2}+\ldots+\frac{h^{m-\mu+1}\times g^1}{g^1h^{m-\mu+1}}\right), \tag{5.5}$$

where h^k (g^k) is the right (left) eigenvector for $k=1$ and adjoint vector for $k>1$ and the summation is over all the geometrically distinct eigenvalues equal to the given λ_σ. In particular, if λ_σ is simple eigenvalue then $m_\sigma=1$ and $Z_{\sigma1}=h\times g/(gh)$.

To obtain series expansion of the resolvent at some eigenvalue λ_1 we reduce the matrix partially to the Jordan form yielding

$$(\lambda-A)^{-1}=H(\lambda_1)\left[\lambda-J(\lambda_1)\right]^{-1}G(\lambda_1)+PE\left[\lambda-(FPE)^{-1}FPAPE\right]^{-1}(FPE)^{-1}FP,$$

where all notations are from (4.4). The first term is the singular part which can be evaluated via (5.4) and the second term can be expanded in regular Taylor series.

For generalized eigenvalue problem $Hh=\lambda Sh$ all the formulas are the same but the scalar product is generalized and instead of $f(A)$ we must write $S^{-1}f(HS^{-1})\equiv f(S^{-1}H)S^{-1}$. Because all the eigenvalues are nondefective the formula (5.3) in conjunction with (5.5) reduces to

$$f(S^{-1}Ht)=\sum_{\lambda}f(\lambda t)\frac{h\times h^+S}{h^+Sh}$$

(note that Maple 10 returns normalized eigenvectors if they are real).

5.3. Special relations

Campbell–Hausdorff formula:

$$\ln\left(e^Ae^B\right)=A+B+\frac{1}{2}[A,B]+\frac{1}{12}\left([A,[A,B]]+[[A,B],B]\right)+\ldots=\sum_{m,n=1}^{\infty}\frac{(-1)^{m+1}}{m}\sum\frac{A^{k_1}B^{l_1}\ldots A^{k_m}B^{l_m}}{k_1!l_1!\ldots k_m!l_m!},$$

where the summation is over $\{k_i \geqslant 0,\ l_i \geqslant 0,\ k_i + l_i \geqslant 1,\ k_1 + l_1 + \ldots + k_m + l_m = n\}$. Related formula:

$$e^A X e^{-A} = X + [A, X] + \frac{1}{2!}[A, [A, X]] + \frac{1}{3!}[A, [A, [A, X]]] + \ldots$$

Derivative:

$$\frac{\partial\, \mathrm{tr}(AB)}{\partial A} = B^\top.$$

§6. Other topics

6.1. Linear equations

Consider equation $Ax = b$. Its solution can be found by the Kramer's rule: $x_i = \det A_i / \det A$, where A_i is the matrix obtained from A by replacing i-th column by b.

The equation $AX = XB$ has nontrivial solution iff A and B have common eigenvalues (can be solved by using Jordan form).

6.2. Bilinear forms

For any matrix A the bilinear form is defined: $(x, Ay) = \sum_{ij} A_{ij} x_i y_j$. For the symmetric matrix A the quadratic form is defined: $(x, Ax) = \sum_{ij} A_{ij} x_i x_j$. The symmetric matrix A is positively defined if and only if

$$A_{11} > 0, \quad \det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} > 0, \quad \ldots, \quad \det A > 0. \tag{6.1}$$

The symmetric matrix A is non-negatively defined if and only if all the principal minors (not only Eq. (6.1)) are non-negative. The conditions of negatively (non-positively) defined matrices are the same but with $(-1)^p$ multiplier where p is the order of a minor.

6.3. Matrices with positive elements

We call the matrix A positive (non-negative) and denote $A > 0$ ($A \geqslant 0$) if all its elements are positive (non-negative). The same notations are for vectors.

Perron–Frobenius theorem: Any positive matrix $A > 0$ has the simple eigenvalue $\lambda_0 > 0$ which is greater than the absolute value of any other eigenvalue, $\lambda_0 > |\lambda|$; the eigenvector corresponding to this eigenvalue is positive, $h_0 > 0$. Any irreducible non-negative matrix $A \geqslant 0$ has the simple eigenvalue $\lambda_0 > 0$ which is greater than the absolute value of any other eigenvalue, $\lambda_0 > |\lambda|$, except for possible $m - 1$ eigenvalues $\lambda = \lambda_0 \sqrt[m]{1}$ (in this case the whole spectrum is invariant under the rotation on the angle $2\pi/m$); the eigenvector corresponding to λ_0 is positive, $h_0 > 0$. And finally, any non-negative matrix $A \geqslant 0$ has the eigenvalue $\lambda_0 > 0$ which is not less than the absolute value of any other eigenvalue, $\lambda_0 \geqslant |\lambda|$; the eigenvector corresponding to this eigenvalue is non-negative, $h_0 \geqslant 0$.

The resolvent $R_A(\lambda)$ and some other matrices of the non-negative matrix $A \geqslant 0$ for $\lambda > \lambda_0(A) \equiv \rho(A)$ has the following properties: $R(\lambda) \geqslant 0$, $\mathrm{d}R(\lambda)/\,\mathrm{d}\lambda \leqslant 0$, $B(\lambda) \geqslant 0$ and $C(\lambda) \geqslant 0$. Also for the non-negative matrix $A \geqslant 0$ one has $\min_i \sum_j A_{ij} \leqslant \lambda_0 \leqslant \max_i \sum_j A_{ij}$.

6.4. Perturbation theory for generalized eigenvalue problem

Let H be a Hermitian matrix and S be a positive-definite Hermitian matrix. Let $H = H_0 + \epsilon V$ and $S = S_0 + \epsilon Q$, where ϵ is a small parameter. Then the solution of the generalized eigenvalue problem $HT = STE$ can be expanded in series of ϵ, provided that the spectral problem $H_0 T_0 = S_0 T_0 E^{(0)}$ is nondegenerate:

$$E = \sum_{n=0}^{\infty} \epsilon^n E^{(n)}, \quad T = T_0 \sum_{n=0}^{\infty} \epsilon^n X^{(n)},\ X^{(0)} = 1.$$

The expansion coefficients can be determined iteratively using the equation:

$$E^{(n)} + X^{(n)} E^{(0)} - E^{(0)} X^{(n)} = v X^{(n-1)} - \sum_{k+m=1}^{n} X^{(k)} E^{(m)} - q \sum_{k+m=0}^{n-1} X^{(k)} E^{(m)},$$

where $v = T_0^+ V T_0$, $q = T_0^+ Q T_0$, and the two summations run through all possible values of k, m with $k+m$ varying in the given limits. In particular,

$$E_\alpha^{(1)} = v_{\alpha\alpha} - q_{\alpha\alpha} E_\alpha^{(0)}, \qquad X_{\beta\alpha}^{(1)} = \frac{v_{\beta\alpha} - q_{\beta\alpha} E_\alpha^{(0)}}{E_{\alpha\beta}},\ \alpha \neq \beta,\ X_{\alpha\alpha}^{(1)} = -\frac{q_{\alpha\alpha}}{2},$$

$$E_\alpha^{(2)} = \sum_{\beta \neq \alpha} \frac{\left| v_{\beta\alpha} - q_{\beta\alpha} E_\alpha^{(0)} \right|^2}{E_{\alpha\beta}} - q_{\alpha\alpha} \left(v_{\alpha\alpha} - q_{\alpha\alpha} E_\alpha^{(0)} \right), \qquad \text{where } E_{\alpha\beta} = E_\alpha^{(0)} - E_\beta^{(0)}.$$

§7. Special relations for special matrices

7.1. Tridiagonal matrices

Let A be a tridiagonal matrix. Denote A_k to be the $A_{1..k,1..k}$ matrix, then

$$\det A_{k+1} = A_{k+1,k+1} \det A_k - A_{k+1,k} A_{k,k+1} \det A_{k-1}.$$

In particular, if

$$A = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \dots & 0 \\ \mu_1 & -\mu_1 - \lambda_1 & \lambda_1 & \dots & 0 \\ 0 & \mu_2 & -\mu_2 - \lambda_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & -\lambda_n \end{pmatrix}$$

then $\Delta_k(\lambda) = \det(\lambda - A_k)$ can be calculated by the recurrence relations

$$\Delta_k - (\lambda + \lambda_k + \mu_k)\Delta_{k-1} + \lambda_{k-1}\mu_k \Delta_{k-2} = 0 \quad \text{with } \Delta_{-1} = 1,\ \Delta_0 = \lambda + \lambda_0.$$

Now if all $\lambda_i, \mu_i > 0$ then the eigenvalues of the matrix A_k are distinct, non-positive, and interleave the eigenvalues of the matrix A_{k-1}.

Let S and H are “cyclic” tridiagonal matrices:

$$S_{ii} = 1,\ S_{i,i+1} = S_{i,i-1} = s/2, \quad H_{ii} = 0,\ H_{i,i+1} = H_{i,i-1} = 1, \quad i = \overline{1, n},$$

where “cyclic” means $n + i \sim i$. Then

$$(S^{-1})_{ij} = \frac{1+q^2}{1-q^2} \frac{q^{|i-j|} + q^{n-|i-j|}}{1-q^n}, \tag{7.1}$$

$$S^{-1/2} H S^{-1/2} = \frac{2}{s} \left(1 - S^{-1} \right), \tag{7.2}$$

where

$$q = -\frac{s}{1 + \sqrt{1-s^2}}, \quad q + \frac{1}{q} = -\frac{2}{s}.$$

For pure tridiagonal matrices S and H

$$(S^{-1})_{ij} = \text{r.h.s.}(7.1) - \frac{1+q^2}{1-q^2} \frac{1-q^{n+2}}{1-q^{2n+2}} \left(\frac{q^{i+j} + q^{2n+2-i-j}}{1-q^{n+2}} + \frac{q^{n+i-j} + q^{n-i+j}}{1-q^n} \right),$$

while (7.2) holds unchanged.

7.2. Block matrices

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det\left(D - CA^{-1}B\right), \qquad \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix}$$

where

$$A' = \left(A - BD^{-1}C\right)^{-1}, \quad D' = \left(D - CA^{-1}B\right)^{-1},$$
$$B' = \left(C - DB^{-1}A\right)^{-1} \equiv -A^{-1}BD', \quad C' = \left(B - AC^{-1}D\right)^{-1} \equiv -D^{-1}CA'.$$

In particular, if A is a scalar denoted by α then the above formulas reduce to

$$\begin{pmatrix} \alpha & b \\ c & D \end{pmatrix}^{-1} = \begin{pmatrix} \alpha' & b' \\ c' & D' \end{pmatrix},$$

where

$$\alpha' = \frac{1}{\alpha - bD^{-1}c}, \quad b' = -\alpha' bD^{-1}, \quad c' = -\alpha' D^{-1}c, \quad D' = D^{-1} + \frac{1}{\alpha'}c' \times b'.$$

Another particular case is when the only nonzero entry of the matrix B is $B_{ij'} = x$ and the only nonzero entry of the matrix C is $C_{i'j} = y$ then

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det D - xy\alpha_{ij}\delta_{i'j'},$$

where α and δ are cofactors of the matrices A and D, respectively.

7.3. Outer product matrices

$$(A + u \times v)^{-1} = A^{-1} - \frac{1}{1 + vA^{-1}u}A^{-1}u \times vA^{-1},$$
$$\det\left(A + u \times v\right) = \det A + \sum_{i=1}^{n} \det A_i,$$

where A_i is constructed from A by replacing i-th row with corresponding row of $u \times v$ matrix. The spectrum of A is determined from the equation $1 + v(A - \lambda)^{-1}u = 0$.

In the symmetric case

$$(A + u \times v + v \times u)^{-1} = A^{-1} + \frac{vA^{-1}v}{\Delta}A^{-1}u \times uA^{-1} + \frac{uA^{-1}u}{\Delta}A^{-1}v \times vA^{-1}$$
$$- \frac{1 + uA^{-1}v}{\Delta}A^{-1}u \times vA^{-1} - \frac{1 + vA^{-1}u}{\Delta}A^{-1}v \times uA^{-1},$$
$$\Delta = \left(1 + uA^{-1}v\right)\left(1 + vA^{-1}u\right) - \left(uA^{-1}u\right)\left(vA^{-1}v\right).$$

In particular, if

$$M = \begin{pmatrix} \beta + \alpha_1 & \beta & \dots & \beta \\ \beta & \beta + \alpha_2 & \dots & \beta \\ \dots & \dots & \dots & \dots \\ \beta & \beta & \dots & \beta + \alpha_n \end{pmatrix}$$

then

$$\det M = \left(1 + \sum_{k=1}^{n} \beta\alpha_k^{-1}\right)\prod_{i=1}^{n}\alpha_i, \qquad \left(M^{-1}\right)_{ij} = \delta_{ij}\alpha_i^{-1} - \frac{\beta\alpha_i^{-1}\alpha_j^{-1}}{1 + \sum_{k=1}^{n}\beta\alpha_k^{-1}}, \qquad \sum_{i=1}^{n}\frac{1}{\lambda - \alpha_i} = \frac{1}{\beta},$$

so that for $\beta > 0$ the i-th eigenvalue lies just above α_i.

7.4. 2 by 2 matrices

$$\det(A+B) = \det A + \det B + \operatorname{tr} A \operatorname{tr} B - \operatorname{tr} AB.$$

For a matrix

$$A = a + b\begin{pmatrix} \cos\beta & e^{-i\alpha}\sin\beta \\ e^{i\alpha}\sin\beta & -\cos\beta \end{pmatrix}$$

with real α and β, the eigenvalues are $a \pm b$ and the eigenvectors are

$$T = \begin{pmatrix} \cos\frac{\beta}{2} & -e^{-i\alpha}\sin\frac{\beta}{2} \\ e^{i\alpha}\sin\frac{\beta}{2} & \cos\frac{\beta}{2}, \end{pmatrix}$$

so that

$$e^{At} = e^{at}\left[\cosh bt + \sinh bt \begin{pmatrix} \cos\beta & e^{-i\alpha}\sin\beta \\ e^{i\alpha}\sin\beta & -\cos\beta \end{pmatrix}\right].$$

7.5. Other matrices

1. Binomial matrices: Because of orthogonality of binomial coefficients given by the identity

$$\sum_{k=m}^{n} \binom{s+n}{s+k}\binom{s+k}{s+m}(-1)^{k-m} = \delta_{nm},$$

the solution of the equation

$$\sum_{k=0}^{n} \binom{s+n}{s+k} x_k = y_k$$

is given by

$$x_n = \sum_{k=0}^{n} (-1)^{n-k}\binom{s+n}{s+k} y_k.$$

2. Matrices with main diagonal, one column, and one row:

$$A = \begin{pmatrix} c_1 & b_2 & b_3 & b_4 & \dots & b_n \\ a_2 & c_2 & 0 & 0 & \dots & 0 \\ a_3 & 0 & c_3 & 0 & \dots & 0 \\ a_4 & 0 & 0 & c_4 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_n & 0 & 0 & 0 & \dots & c_n \end{pmatrix}.$$

Denote $C = c_2 c_3 \dots c_n$, then

$$\det A = C\left(c_1 - \frac{a_2 b_2}{c_2} - \frac{a_3 b_3}{c_3} - \dots - \frac{a_n b_n}{c_n}\right), \qquad A^{-1} = \frac{C}{\det A} B,$$

where

$$B_{11} = 1, \quad B_{i1} = -\frac{a_i}{c_i}, \quad B_{1j} = -\frac{b_j}{c_j}, \quad B_{ii} = \frac{1}{c_i}\left(\frac{\det A}{C} + \frac{a_i b_i}{c_i}\right), \quad B_{ij} = \frac{a_i b_j}{c_i c_j}.$$