Jordan canonical form

- Jordan block
- Jordan canonical form
- Extra material. Normal matrices

Definition. A Jordan block $J_k(\lambda)$ is a $k \times k$ matrix with λ on the main diagonal and 1 above the main diagonal:

$$J_k^{\lambda_o} = \begin{pmatrix} \lambda_o & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_o & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda_o & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \lambda_o & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \lambda_o & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_o \end{pmatrix}$$

Properties of Jordan block:
a) It has only one eigenvalue $\lambda = \lambda_o$ with algebraic multiplicity k:

$$\mathrm{Det}(J_k^{\lambda_o}) = \mathrm{Det}\begin{pmatrix} \lambda_o-\lambda & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & \lambda_o-\lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \lambda_o-\lambda & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & \lambda_o-\lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & \lambda_o-\lambda & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_o-\lambda \end{pmatrix}$$

$$= (\lambda_o - \lambda)^k.$$

b) Geometric multiplicity of $\lambda = \lambda_o$ is 1:

$$\mathrm{Ker}\begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} = \mathrm{span}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}.$$

c) Denote

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \dots \\ 0 \\ 0 \end{pmatrix}, \dots, e_k = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \dots \\ 0 \\ 1 \end{pmatrix},$$

then

$$J_k^{\lambda_o} e_1 = \lambda_o e_1,\ J_k^{\lambda_o} e_i = \lambda_o e_i + e_{i-1},\ i = 2, 3, \ldots, k.$$

d) For $B = J_k^{\lambda_o} - \lambda_o I_k$ we have

$$0 = Be_1,\ e_1 = Be_2,\ e_2 = Be_3, \ldots, e_{k-1} = Be_k.$$

Definition. A *Jordan Canonical Form* is a block-diagonal $n \times n$ matrix:

$$J = \begin{pmatrix} J_{\mathrm{k}_1}^{\lambda_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & J_{\mathrm{k}_2}^{\lambda_2} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & J_{\mathrm{k}_m}^{\lambda_m} \end{pmatrix},$$

with m Jordan blocks $J_{\mathrm{k}_1}^{\lambda_1}$, $J_{\mathrm{k}_2}^{\lambda_2}$, ..., $J_{\mathrm{k}_m}^{\lambda_m}$, such that $k_1 + k_2 + \ldots k_m = n$, and $\mathbf{0}$ denotes a zero matrix.

Properties of the Jordan Canonical form

$$\text{a)}\det(J - \lambda I) = (\lambda_1 - \lambda)^{k_1} (\lambda_2 - \lambda)^{k_2} \ldots (\lambda_m - \lambda)^{k_m}.$$

Proof Determinant of an upper triangular matrix.

b) The Jordan Canonical form has m eigenvectors $x_1^1, x_2^1, \ldots, x_m^1$. Each eigenvector corresponds to one Jordan block.

Proof By induction on the number of Jordan blocks. Hint: show that all eigenvectors of J can be chosen to be orthogonal.

Theorem. Every $n \times n$ matrix A is similar to a **Jordan Canonical Form**:

$$S^{-1} A S = J = \begin{pmatrix} J_{\mathrm{k}_1}^{\lambda_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & J_{\mathrm{k}_2}^{\lambda_2} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & J_{\mathrm{k}_m}^{\lambda_m} \end{pmatrix}.$$

Proof: Suppose $x_1^1, x_2^1, \ldots, x_m^1$ are <u>all</u> linearly independent eigenvectors with (not necessarily distinct) eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$:

$$A x_k^j = \lambda x_k^j,\ j = 1. \tag{1}$$

If we construct exactly n linearly independent vectors $x_1^2, x_1^3, \ldots,\ x_2^2, x_2^3, \ldots, \ldots$ such that

$$A x_k^j = \lambda x_k^j + x_k^{j-1},\ j > 1. \tag{2}$$

then

$$AS = SJ.$$

We construct the vectors that satisfy (2) *by induction* on the dimension of the matrix.

If the matrix is one-dimensional, the statement is obvious.

Assume that all $(n-r) \times (n-r)$ matrices have Jordan form.

Since matrices A and $A - \mu I$ have the same eigenvectors (Why?), without loss of generality, we can assume that one of the eigenvalues of A is zero (Why?), in other words A is singular.
Therefore the image $\mathrm{Im}(A)$ has dimension $r < n$. Hence on the sub-space $\mathrm{Im}(A)$ the theorem holds by induction: there are r linearly independent vectors $y_k^j \in \mathrm{Im}(A)$ so that (2) holds when the matrix A is viewed as a linear map A from the image to the image of A:

$$A : \mathrm{Im}(A) \to \mathrm{Im}(A). \tag{3}$$

Denote by $V = \mathrm{Im}(A)$, and by A_v the restriction of A on V. Then (3) is

$$A_v : V \to V.$$

Suppose dim(Ker(A_v) =dim(Im(A) $\cap$ Ker(A))= p. Then, there are p vectors $x_k \notin V$ such that

$$Ax_k = y_k^j.$$

In other words we can add x_k to the "ends" of chains in V.
Dim(Ker(A)= $n-r$), hence there are $n-r-p$ basis vectors z_i, $z_i \perp \mathrm{Ker}(A) \cap \mathrm{Im}(A)$, $z_i \in \mathrm{Ker}(A)$.
These z's correspond to blocks $J = (\mathbf{0})$.
We constructed vectors y_k^j, x_k and z_k, there are exactly n of them, they satisfy (1) and (2).
Main issue in the proof: why all y_k^j, x_k and z_k are linearly independent?
Suppose

$$\sum_{j,k} c_k^j y_k^j + \sum_k d_k x_k + \sum_i g_i z_i = 0.$$

Apply A, since $Az_i = 0$

$$\sum_{j,k,\lambda \neq 0} c_k^j \begin{pmatrix} \lambda_k y_k^1 \\ \text{or} \\ \lambda_k y_k^j + y_k^{j-1} \end{pmatrix} + \sum_{j,k,\lambda=0} c_k^j \begin{pmatrix} 0 \\ \text{or} \\ y_k^{j-1} \end{pmatrix} + \sum_{k,j=j_{\max},\lambda=0} d_k y_k^j = 0.$$

Since, by induction, all y_k^j are linearly independent, all $d_k = 0$. Hence

$$\sum_{j,k} c_k^j y_k^j = -\sum_i g_i z_i$$

which is an equality between an element in the column space and an element in the space, orthogonal to column space. Hence the last equality is an equality of zeroes.

Extra material: Normal matrices *Recall* that a (real-valued) matrix is symmetric if $A = A^t$.
Definition. The conjugate transpose of a (complex-valued) matrix is

$$A^* = \bar{A}^t.$$

In other words we take a transpose of a matrix and then take a complex conjugate of each of its entries.
Question. What is the conjugate transpose of a (real-valued) symmetric matrix?
Question. Recall that the (real-valued) symmetric matrices "commute" with respect to the (real-valued) inner product:

$$\langle x, Ay \rangle = \langle Ax, y \rangle.$$

Define the complex-valued inner product as

$$\langle x, y \rangle = \sum \bar{x}_i y_i.$$

Note that if

$$\langle x, y \rangle = \langle y, x \rangle,$$

then $\langle x, y \rangle$ is real.

How to define complex-valued matrices, that commute with the complex-valued inner product?
Definition. A (complex-valued) matrix is Hermitian if

$$A^* = A.$$

Hermitian matrices is a complex analog of symmetric matrices.
Observation. Every eigenvalue of a hermitian matrix is real.
Indeed, $\langle x, Ax \rangle$ is real, hence for any eigenvector x

$$\langle x, Ax \rangle = \lambda ||x||^2,$$

and λ is real.
Definition. A (complex-valued) matrix A is normal if

$$A^* A = AA^*,$$

that is, it commutes with its conjugate transpose.
Definition. A (complex-valued) matrix A is unitary if

$$A^{-1} = A^*,$$

Lemma If a matrix A is normal, then it is diagonalizable.
Note: Normal matrices are exactly those, that possess a complete set of *orthonormal* eigenvectors.
Observe, that here we get even more, not only diagonalizability, but also orthogonality of eigenvectors.
Proof: Step 1. For any square matrix A there is a unitary matrix U such that

$$U^{-1} A U = T,$$

where T is upper triangular with non-unity elements on the diagonal.
Indeed, for any matrix A there is at least one eigenvector/eigenvalue:

$$A x_1 = \lambda_1 x_1.$$

Let $x_1, x_2, \ldots, x_n$ be an orthonormal basis, where x_1 is the eigenvector and $x_2, x_3, \ldots, x_n$ are arbitrary. Then

$$AU_1 = U_1 M_1, \quad M_1 = \begin{pmatrix} \lambda_1 & * & * & * & \ldots & * & * \\ 0 & * & * & * & \ldots & * & * \\ 0 & * & * & * & \ldots & * & * \\ 0 & * & * & * & \ldots & * & * \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & * & * & * & \ldots & * & * \end{pmatrix} = \begin{pmatrix} \lambda_1 & * \\ 0 & A_1 \end{pmatrix}$$

Continue inductively for $A_1, A_2, \ldots, A_{n-1}$, and

$$U_2 = \begin{pmatrix} 1 & 0 \\ 0 & M_2 \end{pmatrix}$$

Step 2. If A is normal then $T = U^{-1}AU$ is also normal (why?), but a triangular normal matrix must be diagonal. Why? because the length of every row must equal the length of every column:

$$||Tx|| = ||T^*x||.$$

Corollary. Symmetric and hermitian, orthogonal and unitary matrices are diagonalizable.
Proof: They are normal.
Question. For hermitian, orthogonal and unitary matrices

$$A = UDU^{-1}$$

where U is complex-valued in general. Can we guarantee that U is always real-valued for symmetric matrices?
Question. Yes, because all its eigenvectors can be chosen to be real. Why?