1.4.1. Definition.
$B \in \mathbb{K}^{n\times n}$ is similar to $A \in \mathbb{K}^{n\times n}$ if $B = P^{-1}AP$ for some invertible $n \times n$-matrix P.

1.4.2. Definition.
A is diagonalizable if it is similar to a diagonal matrix

$$B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

1.4.4. Definition.
λ is an eigenvalue of A if there is a nonzero $v \in \mathbb{K}^n$ such that

$$Av = \lambda v.$$

1.4.5. Lemma.
λ is an eigenvalue of A if and only if

$$\det(\lambda I - A) = 0$$

The polynomial $\operatorname{char}_A(x) = \det(xI - A)$ is called the characteristic polynomial of A.

REMARK.

(a) Two similar matrices have the same characteristic polynomial.

(b) The *multiplicity* of an eigenvalue λ_0 is the multiplicity of λ_0 as a zero of char_A; i.e., the largest number n such that $(x - \lambda)^n$ divides char_A.

Reminder: Computing determinants

The determinant of A behaves as follows under row transformations:

(a) Adding a multiple of a row to another row doesn't change the determinant.

(b) Exchanging two rows multiplies the determinant by -1.

(c) Multiplying a row by λ multiplies the determinant by λ.

The same is true for columns. So we can compute the determinant by transforming A into echelon form using row and column operations.

Another common way of computing determinants is by *developing* a certain row or column:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \dots & a_{jn} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

$$= \sum_{k=1}^{n} (-1)^{j+k} a_{jk} \begin{vmatrix} a_{11} & \dots & a_{1(k-1)} & a_{1(k+1)} & \dots & a_{1n} \\ \vdots & & & & \vdots & \\ a_{(j-1)1} & \dots & a_{(j-1)(k-1)} & a_{(j-1)(k+1)} & \dots & a_{(j-1)n} \\ a_{(j+1)1} & \dots & a_{(j+1)(k-1)} & a_{(j+1)(k+1)} & \dots & a_{(j+1)n} \\ \vdots & & & & \vdots & \\ a_n1 & \dots & a_{n(k-1)} & a_{n(k+1)} & \dots & a_{nn} \end{vmatrix}.$$

Example.

$$\begin{vmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & -1 & 3 \\ 1 & 1 & 1 & 1 \end{vmatrix} = -1 \cdot \begin{vmatrix} 2 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 2 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{vmatrix}$$

$$= -\left(2 \cdot \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix}\right) + 2$$

$$= -(2 \cdot 2 - 1) + 1 = -3 + 1$$

$$= -2.$$

1.4.1. Definition.
$B \in \mathbb{K}^{n\times n}$ is similar to $A \in \mathbb{K}^{n\times n}$ if $B = P^{-1}AP$ for some invertible $n \times n$-matrix P.

1.4.2. Definition.
A is diagonalizable *if it is similar to a* diagonal matrix

$$B = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

1.4.4. Definition.
λ is an eigenvalue of A if there is a nonzero $v \in \mathbb{K}^n$ such that

$$Av = \lambda v.$$

1.4.5. Lemma.
λ is an eigenvalue of A if and only if

$$\det(\lambda I - A) = 0$$

The polynomial $\operatorname{char}_A(x) = \det(xI - A)$ is called the characteristic polynomial *of A.*

Remark.

(a) Two similar matrices have the same characteristic polynomial.

(b) The *multiplicity* of an eigenvalue λ_0 is the multiplicity of λ_0 as a zero of char_A; i.e., the largest number n such that $(x - \lambda)^n$ divides char_A.

1.4.6. Proposition.
The following are equivalent:

(a) A is diagonalizable.

(b) The characteristic polynomial of A can be decomposed into linear factors, i.e.

$$\operatorname{char}_A(x) = (x - \lambda_1)^{m_1} \cdot \dots \cdot (x - \lambda_k)^{m_k},$$

and for every eigenvalue λ_j, the multiplicity m_k of λ_j equals the dimension of the eigenspace $\operatorname{Eig}(\lambda)$;

(c) $\mathbb{K}^n$ has a basis consisting of eigenvectors of A;

(d) $\mathbb{K}^n$ is the direct sum of the eigenspaces of A.

1.4.6. Proposition.
The following are equivalent:

(a) A is diagonalizable.

(b) The characteristic polynomial of A can be decomposed into linear factors, i.e.

$$\mathrm{char}_A(x) = (x - \lambda_1)^{m_1} \cdot \dots \cdot (x - \lambda_k)^{m_k},$$

and for every eigenvalue λ_j, the multiplicity m_k of λ_j equals the dimension of the eigenspace $\mathrm{Eig}(\lambda)$*;*

(c) $\mathbb{K}^n$ has a basis consisting of eigenvectors of A;

(d) $\mathbb{K}^n$ is the direct sum of the eigenspaces of A.

Recipe for diagonalizing A:

(a) Compute the characteristic polynomial $\mathrm{char}_A(x) = \det(xI - A)$, and find all its zeros (i.e., the eigenvalues of A) and their multiplicity.

(b) If the multiplicities do not add up to n; i.e., if $\mathrm{char}_A(x)$ does not decompose into linear factors, then A is not diagonalizable.

(c) For every eigenvalue λ, compute the eigenspace $\mathrm{Eig}(\lambda)$; i.e., all solutions $v \in \mathbb{K}^n$ of

$$(\lambda I - A)v = 0.$$

(d) If the dimensions of the eigenspaces add up to n, then we can pick a basis of $\mathbb{K}^n$ consisting of eigenvectors of A, and put these as the columns of a matrix P. This is a matrix which diagonalizes A.

(e) Otherwise, A is not diagonalizable.

Recipe for diagonalizing A:

(a) Compute the characteristic polynomial $\text{char}_A(x) = \det(xI - A)$, and find all its zeros (i.e., the eigenvalues of A) and their multiplicity.

(b) If the multiplicities do not add up to n; i.e., if $\text{char}_A(x)$ does not decompose into linear factors, then A is not diagonalizable.

(c) For every eigenvalue λ, compute the eigenspace $\text{Eig}(\lambda)$; i.e., all solutions $v \in \mathbb{K}^n$ of

$$(\lambda I - A)v = 0.$$

(d) If the dimensions of the eigenspaces add up to n, then we can pick a basis of $\mathbb{K}^n$ consisting of eigenvectors of A, and put these as the columns of a matrix P. This is a matrix which diagonalizes A.

(e) Otherwise, A is not diagonalizable.

1.4.7. Examples.(a) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$;

(b) $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$;

(c) $\varphi : \text{Pol}_2(\mathbb{R}) \to \text{Pol}_2(\mathbb{R}); \varphi(ax^2 + bx + c) = (c - b)x^2 + (2a + 3b - 2c)x + c$.

Recall our goal:

Given a matrix $A \in \mathbb{R}^{n\times n}$ (or $A \in \mathbb{C}^{n\times n}$, find a matrix B similar to A which is as "simple" as possible.

2×2-matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then $\mathrm{char}_A(x) = (x-a)(x-d) - bc = x^2 - (a+d)x + ad - bc$.

(a) $\mathrm{char}_A(x)$ has two distinct real roots: A is diagonalizable over $\mathbb{R}$.

(b) $\mathrm{char}_A(x)$ has no real roots, but two distinct complex roots $a \pm ib$: A is not diagonalizable over $\mathbb{R}$, but *is* diagonalizable over $\mathbb{C}$.

(c) $\mathrm{char}_A(x)$ has only one real root λ: then either $A = \lambda I$, or A is *not* diagonalizable over $\mathbb{C}$.

(However, A is similar to $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.)

Recall our goal:

Given a matrix $A \in \mathbb{R}^{n\times n}$ (or $A \in \mathbb{C}^{n\times n}$, find a matrix B similar to A which is as "simple" as possible.

2×2-matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then $\mathrm{char}_A(x) = (x-a)(x-d) - bc = x^2 - (a+d)x + ad - bc$.

(a) $\mathrm{char}_A(x)$ has two distinct real roots: A is diagonalizable over $\mathbb{R}$.

(b) $\mathrm{char}_A(x)$ has no real roots, but two distinct complex roots $a \pm ib$: A is not diagonalizable over $\mathbb{R}$, but *is* diagonalizable over $\mathbb{C}$.

(c) $\mathrm{char}_A(x)$ has only one real root λ: then either $A = \lambda I$, or A is *not* diagonalizable over $\mathbb{C}$.

(However, A is similar to $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$.)

The general case

(a) Any real symmetric matrix is diagonalizable;

(b) Any real matrix whose characteristic polynomial splits into linear factors is similar to an *upper triangular matrix*;

(c) Any complex matrix is similar to a (complex) *upper triangular matrix*;

(We can replace "upper triangular matrix" by "matrix in Jordan normal form".)

1.4.7. Proposition.
Let $A \in \mathbb{K}^{n\times n}$ *and suppose that*

$$\mathrm{char}_A(x) = (x-\lambda_1)\cdot(x-\lambda_2)\cdot\dots\cdot(x-\lambda_n).$$

Then A *is similar to an upper triangular matrix*

$$B = \begin{pmatrix} \lambda_1 & * & * & * & \dots & * & * \\ 0 & \lambda_2 & * & * & \dots & * & * \\ 0 & 0 & \lambda_3 & * & \dots & * & * \\ \vdots & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

1.4.7. Proposition.
Let $A \in \mathbb{K}^{n\times n}$ *and suppose that*

$$\mathrm{char}_A(x) = (x-\lambda_1)\cdot(x-\lambda_2)\cdot\dots\cdot(x-\lambda_n).$$

(with the λ_j *not necessarily different).*
Then A *is similar to an upper triangular matrix*

$$B = \begin{pmatrix} \lambda_1 & * & * & * & \dots & * & * \\ 0 & \lambda_2 & * & * & \dots & * & * \\ 0 & 0 & \lambda_3 & * & \dots & * & * \\ \vdots & & & & & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda_n \end{pmatrix}.$$

1.4.9. Definition (Jordan matrices)**.**
A Jordan block *is an* $m \times m$*-matrix of the form*

$$J(\lambda, m) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

A Jordan matrix *is a square matrix of the following form (where each block* $\boxed{0}$ *represents a matrix consisting only of zeros):*

$$B = \begin{pmatrix} \boxed{J(\lambda_1, m_1)} & \boxed{0} & \dots & \boxed{0} \\ \boxed{0} & \boxed{J(\lambda_2, m_2)} & \dots & \boxed{0} \\ \boxed{0} & \boxed{0} & \dots & \boxed{J(\lambda_k, m_k)} \end{pmatrix}.$$

1.4.10. Proposition (Jordan normal form)**.**
Let $A \in \mathbb{K}^{n\times n}$ *and suppose that*

$$\mathrm{char}_A(x) = (x-\lambda_1)\cdot(x-\lambda_2)\cdot\dots\cdot(x-\lambda_n).$$

(with the λ_j *not necessarily different).*
Then A *is similar to a Jordan matrix. This Jordan matrix is unique up to the order of its Jordan blocks along the diagonal.*

1.4.9. Definition (Jordan matrices).
A Jordan block is an $m \times m$-matrix of the form

$$J(\lambda, m) = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & 1 & \dots & 0 & 0 \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & \dots & \lambda & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda \end{pmatrix}.$$

A Jordan matrix is a square matrix of the following form (where each block $\boxed{0}$ represents a matrix consisting only of zeros):

$$B = \begin{pmatrix} \boxed{J(\lambda_1, m_1)} & \boxed{0} & \dots & \boxed{0} \\ \boxed{0} & \boxed{J(\lambda_2, m_2)} & \dots & \boxed{0} \\ \boxed{0} & \boxed{0} & \dots & \boxed{J(\lambda_k, m_k)} \end{pmatrix}.$$

1.4.10. Proposition (Jordan normal form).
Let $A \in \mathbb{K}^{n \times n}$ and suppose that

$$\mathrm{char}_A(x) = (x - \lambda_1) \cdot (x - \lambda_2) \cdot \dots \cdot (x - \lambda_n).$$

(with the λ_j not necessarily different).

Then A is similar to a Jordan matrix. This Jordan matrix is unique up to the order of its Jordan blocks along the diagonal.

3×3 matrices

- Case 1: A is diagonalizable. Jordan form is diagonal.
- Case 2: $\mathrm{char}_A(x) = (x - \lambda_1)^2 (x - \lambda_2)$, with only l.i. eigenvectors v_1, v_3. Jordan form must be

$$\begin{pmatrix} \lambda_1 & 1 & 0 \\ 0 & \lambda_1 & 0 \\ 0 & 0 & \lambda_2 \end{pmatrix}.$$

  Solution: Find a vector with $Av_2 = \lambda v_2 + v_1$.

- Case 3: $\mathrm{char}_A(x) = (x - \lambda)^3$; only two l.i. eigenvectors.

  Solution: v_2 arbitary non-eigenvector. $v_1 := Av_2 - \lambda v_2$. v_3 is an eigenvector which is l.i. to v_2.

- Case 4: $\mathrm{char}_A(x) = (x - \lambda)^3$; only one l.i. eigenvectors. Solution: similar to case 2 or 3.