Eigenvalues, eigenvectors, diagonalization

Def. $\mathbf{v} \in V_K$ is an **eigenvector** of the linear transformation $f : V \to V$ if $\mathbf{v} \neq \mathbf{0}$, and there is a scalar $\lambda \in K$ such that $f(\mathbf{v}) = \lambda\mathbf{v}$, that is, $f(\mathbf{v})$ is parallel to $\mathbf{v}$ (including the case when $f(\mathbf{v}) = \mathbf{0}$). Here λ is the **eigenvalue** corresponding to $\mathbf{v}$. The **spectrum** of f is the set of eigenvalues of f. The **eigenspace** corresponding to the eigenvalue λ is $V_\lambda = \{\, \mathbf{v} \in V \mid f(\mathbf{v}) = \lambda\mathbf{v} \,\} \leq V$, which consists of $\mathbf{0}$ and the eigenvectors for λ.

Example: The eigenvectors of an orthogonal projection onto a plane containing the origin are the nonzero vectors of the plane (with eigenvalue 1), and the nonzero vectors orthogonal to the plane (with eigenvalue 0). In other words, the plane itself is the eigenspace for 1, and the line through the origin which is perpendicular to the plane is the eigenspace for 0.

Def. The **eigenvectors, eigenvalues** and the **spectrum** of a matrix A are those of the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$.

Diagonalization (spectral decomposition)

$A \in K^{n\times n}$, $f : K^n \to K^n$, $f : \mathbf{x} \mapsto A\mathbf{x}$. If $\exists$ a basis $\mathcal{B} = \{\, \mathbf{b}_1, \dots, \mathbf{b}_n \,\}$ consisting of eigenvectors of f with eigenvalues $\lambda_1, \dots, \lambda_n$ then

$$[f]_{\mathcal{B}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \dots & 0 \\ 0 & 0 & \dots & \lambda_{n-1} & 0 \\ 0 & 0 & 0 & \dots & \lambda_n \end{bmatrix} = D$$

is a diagonal matrix, and with the transition matrix $P = [\mathbf{b}_1 \dots \mathbf{b}_n]$ we have $D = P^{-1}AP$, that is, $A = PDP^{-1}$. The latter is the spectral decomposition of A.

Def. $A \in K^{n\times n}$ is **diagonalizable** if there exists an invertible matrix P such that $P^{-1}AP$ is diagonal, i.e. $\exists$ a basis in K^n consisting of eigenvectors of A.

Powers of diagonalizable matrices

If $A = PDP^{-1}$, then $A^k = (PDP^{-1})(PDP^{-1})\cdots(PDP^{-1}) = PD^kP^{-1}$, and we obtain the kth power of a diagonal matrix simply by taking the kth powers of the diagonal elements.

Calculating eigenvalues and eigenvectors

$$\exists \mathbf{v} \neq \mathbf{0} : A\mathbf{v} = \lambda\mathbf{v} \;\Leftrightarrow$$
$$\exists \mathbf{v} \neq \mathbf{0} : (A - \lambda I)\mathbf{v} = \mathbf{0} \;\Leftrightarrow$$
$$|A - \lambda I| = 0$$

Characteristic polynomial

Def. The **characteristic polynomial** of the matrix A is

$$k_A(x) = \begin{vmatrix} a_{11} - x & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - x & \dots & a_{2n} \\ & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} - x \end{vmatrix}$$

Properties of the characteristic polynomial:

- the roots of $k_A(x)$ are the eigenvalues of A;
- $k_A(x) = (-1)^n x^n + (-1)^{n-1}(\operatorname{tr} A)x^{n-1} + \ldots + |A|$, where $\operatorname{tr} A = a_{11} + \ldots + a_{nn}$ is the **trace** of A;
- If $k_A(x)$ can be written as the product of linear polynomials: $k_A(x) = (-1)^n(x-\lambda_1)\ldots(x-\lambda_n)$ (the λ_i are the eigenvalues of A with multiplicities), then $\operatorname{tr} A = \lambda_1 + \ldots + \lambda_n$ and $|A| = \lambda_1 \cdots \lambda_n$.
- $A \sim B \Rightarrow k_A(x) = k_B(x)$, since $|P^{-1}AP - xI| = |P^{-1}AP - xP^{-1}IP| = |P^{-1}(A-xI)P| = |P|^{-1} \cdot |A-xI| \cdot |P| = |A - xI|$.

Exercises: Which of the following matrices are diagonalizable over $\mathbb{R}$ or $\mathbb{C}$?

$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$

$|A - xI| = \begin{vmatrix} 1-x & 2 \\ 0 & 2-x \end{vmatrix}$

$= (x-1)(x-2)$

eigenvalues: $\lambda = 1, 2$

$\exists$ eigenvector for each,,

they are indep. $\Rightarrow$

they form a basis $\Rightarrow$

A is diag.-able

$B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

$|B - xI| = x^2 + 1$

no real root $\Rightarrow$

B is not diag.-able over $\mathbb{R}$

(but diag.-able over $\mathbb{C}$)

$C = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

$|C - xI| = (x-1)^2$

eigenvalue: $\lambda = 1$

$(C - 1 \cdot I)\mathbf{v} = \mathbf{0}$, $\mathbf{v} = ?$

$\left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \Rightarrow \mathbf{v} = \begin{bmatrix} t \\ 0 \end{bmatrix}$

$\nexists$ two indep. eigenvectors $\Rightarrow$

C is not diagonalizable

(neither over $\mathbb{R}$ nor over $\mathbb{C}$)

Minimal polynomial

Def. For $A \in K^{n\times n}$ and $p(x) = c_0 + c_1 x + \cdots + c_m x^m \in K[x]$, we define $p(A) := c_0 I + c_1 A + \cdots + c_m A^m$.

Proposition. *For $A \in K^{n\times n}$ $\exists\ 0 \neq p(x) \in K[x]$, with $p(A) = 0$ (where 0 denotes the matrix with only 0 elements).*

Proof. $I, A, A^2, \ldots, A^{n^2} \in K^{n\times n}$, but $\dim K^{n\times n} = n^2 \Rightarrow$ these are linearly independent $\Rightarrow \exists\ c_0, \ldots, c_{n^2}$ not all 0: $c_0 I + c_1 A + \ldots + c_{n^2} A^{n^2} = 0$. □

Cayley–Hamilton theorem. $k_A(A) = 0$.

No proof.

Def. The minimal polynomial $m_A(x) \in K[x]$ of a matrix $A \in K^{n\times n}$ is the polynomial of minimal degree with main coefficient 1 such that $m_A(A) = 0$. (It follows from the Cayley–Hamilton theorem that $\deg m_A(x) \leq n$.)

Remark. The minimal polynomial remains the same over a larger field, for instance, the minimal polynomial of $A \in \mathbb{R}^{n\times n}$ is the same over $\mathbb{R}$ and $\mathbb{C}$, since $c_0 + c_1 A + \cdots + c_m A^m = 0$ is a linear system of equations for the unknown c_is, so if the coefficients are in K, and there is a nontrivial solution over a field $L \geq K$, then the Gaussian elimination method gives us a nontrivial solution over K as well.

Proposition. *For $p(x) \in K[x]$ we have $p(A) = 0 \Leftrightarrow m_A(x) | p(x)$, that is, $\exists\ q(x) \in K[x]$ such that $p(x) = m_A(x) q(x)$.*

Proof. $\Leftarrow$: $p(A) = m_A(A)q(A) = 0q(A) = 0$

$\Rightarrow$: The polynomial $p(x)$ can be written as $p(x) = m(x)q(x) + r(x)$, such that $\deg r(x) < \deg m(x)$. But $0 = p(A) = m(A)q(A) + r(A) = 0 \cdot q(A) + r(A) = r(A)$ and then $r(x) = 0$ follows from the minimality of $\deg m(x)$. □

Proposition. Ⓟ *Every eigenvalue of A is a root of $m_A(x)$.*

Proof. Let $\mathbf{v}$ be an eigenvector with eigenvalue λ.

$$\begin{aligned} A\mathbf{v} =& \lambda\mathbf{v} \\ A^2\mathbf{v} =& A(\lambda\mathbf{v}) = \lambda A\mathbf{v} = \lambda^2\mathbf{v} \\ &\vdots \\ A^k\mathbf{v} =& \lambda^k\mathbf{v} \\ p(A)\mathbf{v} =& p(\lambda)\mathbf{v} \ \forall p(x) \in K[x] \\ \mathbf{0} = m_A(A)\mathbf{v} =& m_A(\lambda)\mathbf{v} \\ m_A(\lambda) =& 0 \text{ because } \mathbf{v} \neq \mathbf{0}. \end{aligned}$$

□

Corollary. *If $A \in \mathbb{C}^{n\times n}$ and $k_A(x) = (-1)^n(x-\lambda_1)^{a_1}\cdots(x-\lambda_k)^{a_k}$, where $\lambda_1,\ldots,\lambda_k$ are different, then $m_A(x) = (x-\lambda_1)^{b_1}\cdots(x-\lambda_k)^{b_k}$ for some $1 \leq b_i \leq a_i \ \forall i$.*

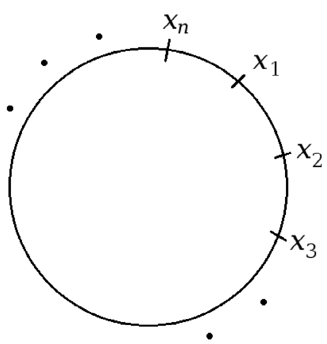
Exercise: Determine the characteristic and the minimal polynomial of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

Solution: $k_A(x) = |A - xI| = -(x-1)^2(x-2)$, so $m_A(x)$ can only be $(x-1)(x-2)$ or $(x-1)^2(x-2)$. We check if A is a 'root' of the first:

$$(A-I)(A-2I) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq 0,$$

so the minimal polynomial is $m_A(x) = (x-1)^2(x-2)$.

Example: The solution of the problem about the coins



$x_i =$ 0 or $1 \in \mathbb{F}_2$
(tail or head)

One round of the game:

$$x_1 \mapsto x_n + x_1$$
$$x_i \mapsto x_{i-1} + x_i$$

This is a linear transformation, its matrix: $\begin{bmatrix} 1 & 0 & 0 & \dots & 1 \\ 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \\ 0 & 0 & \dots & 1 & 1 \end{bmatrix}$.

k rounds: $\mathbf{x} \mapsto A^k\mathbf{x}$, the game ends: $\exists k$: $A^k\mathbf{x} = \mathbf{0}$.
the game ends for $\forall \mathbf{x} \Leftrightarrow$ the game ends for $\forall$ basis vector
$\Leftrightarrow \forall i\ \exists k_i:\ A^{k_i}\mathbf{e}_i = \mathbf{0}$
$\Leftrightarrow \exists k:\ \forall i\ A^k\mathbf{e}_i = \mathbf{0}$ (k is the maximal k_i)
$\Leftrightarrow \exists k:\ 0 = A^k[\mathbf{e}_1 \cdots \mathbf{e}_n] = A^kI = A^k$
$\Leftrightarrow \exists k:\ m_A(x)|x^k$
$\Leftrightarrow 0$ is the only eigenvalue
$\Leftrightarrow k_A(x) = (-1)^n x^n = x^n$ (over $\mathbb{F}_2$)

On the other hand, if we expand $|A - xI|$ by the first column, we get

$$|A - xI| = \begin{vmatrix} 1-x & 0 & 0 & \dots & 1 \\ 1 & 1-x & 0 & \dots & 0 \\ 0 & 1 & 1-x & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \\ 0 & 0 & \dots & 1 & 1-x \end{vmatrix} = (1-x)^n + (-1)^{n+1} \cdot 1 = (x+1)^n + 1$$

over $\mathbb{F}_2$, so the good numbers are those n for which $(x+1)^n + 1 = x^n$, i.e. $(x+1)^n = x^n + 1$ over $\mathbb{F}_2$. If n is a power of 2, then it is true. If $n = m \cdot 2^t$, where $m > 1$ is an odd number, then $(x+1)^n = (x^{2^t} + 1)^m = 1 + mx^{2^t} + \dots$, so the coeffiecient of x^{2^t} is not 0, thus $(x+1)^n \neq x^n + 1$.
Consequently, the game ends when started with any position of the coins if and only if n is a power of 2.

Invariants of matrices

If $A, B \in K^{n\times n}$, and $A \sim B$, then the following are the same for A and B:
a) characteristic polynomial (it has been proved)
b) determinant (it follows from a))
c) trace (it follows from a))
d) minimal polynomial (since $g(P^{-1}AP) = P^{-1}g(A)P$ for every polynomial $g(x)$)
e) rang ($\dim \operatorname{Im} f$ does not depend on the basis)
f) spectrum with multiplicities (the roots of the characteristic polynomial)
g) the dimension of the eigenspace ($A \sim B \Rightarrow (A - \lambda I) \sim (B - \lambda I)$, and $\dim V_\lambda = n - \operatorname{rank}(A - \lambda I)$)