Applications of determinants. Eigenvalues.

Zdeněk Dvořák

April 7, 2016

Lemma 1 (Recursive formula for determinant)**.** *Let A be an $n \times n$-matrix, and let $i \in \{1, \ldots, n\}$. Then*

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(A^{ij}).$$

Theorem 2 (Fundamental theorem of algebra)**.** *Every complex polynomial $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ is equal to $a_n(x - c_1)(x - c_2)\cdots(x - c_n)$, where $c_1, \ldots, c_n$ are the roots of p (i.e., $p(c_1) = \ldots = p(c_n) = 0$).*

The roots may be complex numbers even if p is real, and they do not have to be pairwise distinct. The multiplicity of the root is the number of times it appears among c_1, …, c_n.

1 Geometric meaning of determinants

Theorem 3. *Let v_1, …, v_k be vectors in $\mathbf{R}^n$. Let $A = (v_1|v_2|\ldots|v_k)$. The k-dimensional volume of the parallelepiped with edges v_1, …, v_k is $\sqrt{\det(A^T A)}$ (and thus if $k = n$, then the volume is $|\det(A)|$).*

Proof. We prove the claim by induction on k. If $k = 1$, then the 1-dimensional volume of v_1 is its length $\|v_1\| = \sqrt{v_1^T v_1}$. Hence, assume that $k > 1$.

Let p be the projection of v_k to $\operatorname{span}(v_1, \ldots, v_{k-1})$. Note that $\operatorname{Vol}(v_1, \ldots, v_k) = \operatorname{Vol}(v_1, \ldots, v_{k-1}, v_k - p)$, since both parallelepipeds have the same height. Let $B = (v_1|v_2|\ldots|v_{k-1}|v_k - p)$. Since $p \in \operatorname{span}(v_1, \ldots, v_{k-1})$, we have $p = \alpha_1 v_1 + \ldots + \alpha_{k-1} v_{k-1}$ for some $\alpha_1, \ldots, \alpha_{k-1}$, and $B = AR$, where

$$R = \begin{pmatrix} 1 & 0 & 0 & \ldots & -\alpha_1 \\ 0 & 1 & 0 & \ldots & -\alpha_2 \\ & & \ldots & & \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}.$$

Note that $\det(R) = 1$. Hence, $\det(B^T B) = \det((AR)^T AR) = \det(R^T(A^T A)R) = \det(R^T)\det(A^T A)\det(R) = \det(A^T A)$.

Therefore, it suffices to prove that $\mathrm{Vol}(v_1, \ldots, v_{k-1}, v_k - p) = \sqrt{\det(B^T B)}$. Let $C = (v_1|v_2|\ldots|v_{k-1})$. By the induction hypothesis, we have $\mathrm{Vol}(v_1, \ldots, v_{k-1}) = \sqrt{\det(C^T C)}$. Since $v_k - p$ is perpendicular to $v_1, \ldots, v_{k-1}$, the height of the parallelepiped is $\|v_k - p\|$, and thus $\mathrm{Vol}(v_1, \ldots, v_{k-1}, v_k - p) = \|v_k - p\|\sqrt{\det(C^T C)}$. Furthermore,

$$B^T B = \begin{pmatrix} C^T C & o \\ o^T & \|v_k - p\|^2 \end{pmatrix},$$

and $\det(B^T B) = \|v_k - p\|^2 \det(C^T C)$ by Lemma 1.

Therefore, $\mathrm{Vol}(v_1, \ldots, v_k) = \mathrm{Vol}(v_1, \ldots, v_{k-1}, v_k - p) = \|v_k - p\|\sqrt{\det(C^T C)} = \sqrt{\det(B^T B)} = \sqrt{\det(A^T A)}$. □

Corollary 4. *Let $f : \mathbf{R}^n \to \mathbf{R}^n$ be defined by $f(x) = Ax$ for an $n \times n$ matrix A. For any measurable set $S \subseteq \mathbf{R}^n$, the volume of the set $f(S) = \{f(x) : x \in S\}$ is $|\det(A)|$ times the volume of S.*

Proof. If S is the cube with edges $\alpha e_1, \ldots, \alpha e_n$ for some $\alpha > 0$ (whose volume is α^n), then $f(S)$ is the parallelepiped with edges $\alpha A_{\star,1}, \ldots, \alpha A_{\star,n}$, whose volume is $\alpha^n|\det(A)|$ by Theorem 3. All other measurable sets can be approximated arbitrarily well by a disjoint union of cubes. □

Lemma 5 (Meaning of the sign of the determinant)**.** *Let $v_1, \ldots, v_{n-1}$ be independent vectors in $\mathbf{R}^n$, and let $\mathbf{U}$ be the hyperplane span$(v_1, \ldots, v_{n-1})$. Then $\mathbf{R}^n \setminus \mathbf{U}$ has two connected components X_1 and X_2, such that*

$$\det(v_1|\ldots|v_{n-1}|x) > 0 \textit{ for } x \in X_1, \textit{ and}$$
$$\det(v_1|\ldots|v_{n-1}|x) < 0 \textit{ for } x \in X_2.$$

Proof. Note that $\det(v_1|\ldots|v_{n-1}|x) = 0$ if and only if $x \in \mathbf{U}$. If $c \subseteq \mathbf{R}^n \setminus \mathbf{U}$ is a continuous curve, then $\det(v_1|\ldots|v_{n-1}|x) \neq 0$ for every $x \in c$, and by continuity, the sign of $\det(v_1|\ldots|v_{n-1}|x)$ is the same for all $x \in c$. Hence, $\det(v_1|\ldots|v_{n-1}|x)$ has the same sign for every vector x in each connected component of $\mathbf{R}^n \setminus \mathbf{U}$. Furthermore, if $\det(v_1|\ldots|v_{n-1}|x) > 0$, then $\det(v_1|\ldots|v_{n-1}|-x) < 0$ and vice versa, and thus the signs in distinct components are opposite. □

2 Resultants

Definition 1. *The* <u>*Sylvester matrix*</u> *of two polynomials* $p(x) = a_0 + a_1 x + \ldots + a_n x^n$ *and* $q(x) = b_0 + b_1 x + \ldots + b_m x^m$ *is*

$$S(p,q) = \begin{pmatrix} a_n & a_{n-1} & \ldots & a_0 & 0 & 0 & 0 & \ldots \\ 0 & a_n & a_{n-1} & \ldots & a_0 & 0 & 0 & \ldots \\ 0 & 0 & a_n & a_{n-1} & \ldots & a_0 & 0 & \ldots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \ldots & 0 & a_n & a_{n-1} & \ldots & a_1 & a_0 \\ b_m & b_{m-1} & \ldots & b_0 & 0 & 0 & 0 & \ldots \\ 0 & b_m & b_{m-1} & \ldots & b_0 & 0 & 0 & \ldots \\ 0 & 0 & b_m & b_{m-1} & \ldots & b_0 & 0 & \ldots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \ldots & 0 & b_m & b_{m-1} & \ldots & b_1 & b_0 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\} m \text{ rows} \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\\a\end{matrix}}\right\} n \text{ rows} \end{matrix},$$

and their <u>*resultant*</u> *is*

$$Res(p,q) = \det(S(p,q)).$$

Theorem 6. *If two non-zero polynomials* $p(x)$ *and* $q(x)$ *have a common root, then* $Res(p,q) = 0$.

Proof. Let $n = \deg(p)$ and $m = \deg(q)$. If $p(a) = 0 = q(a)$, then $p(x) = (x-a)p_0(x)$ and $q(x) = (x-a)q_0(x)$ for some non-zero polynomials $p_0(x) = c_0 + c_1 x + \ldots + c_{n-1}x^{n-1}$ and $q_0(x) = d_0 + d_1 x + \ldots + d_{m-1}x^{m-1}$.

Note that $p(x)q_0(x) - q(x)p_0(x) = 0$. On the other hand,

$$\begin{aligned} p(x)q_0(x) - q(x)p_0(x) = & d_0 p(x) + d_1(xp(x)) + d_2(x^2p(x)) + \ldots + d_{m-1}(x^{m-1}p(x)) \\ & - c_0 q(x) - c_1(xq(x)) - c_2(x^2 q(x)) - \ldots - c_{n-1}(x^{n-1}q(x)). \end{aligned}$$

Let $A = S(p,q)$. Observe that $p(x)q_0(x) - q(x)p_0(x) = 0$ is equivalent to

$$\begin{aligned} o = & d_0 A_{m,\star} + d_1 A_{m-1,\star} + d_2 A_{m-2,\star} + \ldots + d_{m-1}A_{1,\star} \\ & - c_0 A_{m+n,\star} - c_1 A_{m+n-1,\star} - c_2 A_{m+n-2,\star} - \ldots - c_{n-1}A_{m+1,\star}. \end{aligned}$$

Hence, A is not regular, and thus $\text{Res}(p,q) = \det(A) = 0$. □

Example 1. *Solve the system of equations*

$$\begin{aligned} x^2 + xy - y^2 &= -5 \\ x^2 - 2xy + y^2 &= 4. \end{aligned}$$

Let $p(x,y) = x^2 + xy + 5 - y^2$ *and* $q(x,y) = x^2 - 2xy - 4 + y^2$. *We need to find values of* x *and* y *such that* $p(x,y) = q(x,y) = 0$.

For a complex number a, let p_a and q_a be the polynomials obtained from p and q by substituting a for y: $p_a(x) = x^2 + ax + (5 - a^2)$ and $q_a(x) = x^2 - 2ax + (a^2 - 4)$. If (b, a) is a solution, then p_a and q_a have a common root b, and thus

$$0 = Res(p_a, q_a) = \det \begin{pmatrix} 1 & a & 5 - a^2 & 0 \\ 0 & 1 & a & 5 - a^2 \\ 1 & -2a & a^2 - 4 & 0 \\ 0 & 1 & -2a & a^2 - 4 \end{pmatrix} = (a^2 - 9)^2.$$

Hence, $a = 3$ or $a = -3$. If $a = 3$, then $p_a(x) = x^2 + 3x - 4 = (x + 4)(x - 1)$ and $p_a(x) = x^2 - 6x + 5 = (x - 5)(x - 1)$ have common root $b = 1$. If $a = -3$, then $p_a(x) = x^2 - 3x - 4 = (x - 4)(x + 1)$ and $p_a(x) = x^2 + 6x + 5 = (x + 5)(x + 1)$ have common root $b = -1$.

Therefore, the solutions are $(x, y) = (1, 3)$ and $(x, y) = (-1, -3)$.

3 Eigenvectors and eigenvalues

Example 2. *Let*

$$A = \begin{pmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{11}} & \frac{1}{\sqrt{11}} & \frac{3}{\sqrt{11}} \\ \frac{7}{\sqrt{66}} & -\frac{4}{\sqrt{66}} & -\frac{1}{\sqrt{66}} \end{pmatrix}.$$

The matrix A is orthonormal, and thus the function $f(x, y, z) = A(x, y, z)^T$ is an isometry of $\mathbf{R}^3$. Furthermore, $\det(A) = 1$, hence it preserves orientation. In fact, it is a rotation of $\mathbf{R}^3$ around an axis passing through $(0, 0, 0)$. How to find this axis?

Any point (x, y, z) of the axis is mapped by f to itself, and thus it satisfies $A(x, y, z)^T = (x, y, z)$. Equivalently,

$$\begin{array}{llll} \left(\frac{1}{\sqrt{6}} - 1\right) x & + \frac{2}{\sqrt{6}} y & - \frac{1}{\sqrt{6}} z & = 0 \\ \frac{1}{\sqrt{11}} x & + \left(\frac{1}{\sqrt{11}} - 1\right) y & + \frac{3}{\sqrt{11}} z & = 0 \\ \frac{7}{\sqrt{66}} x & - \frac{4}{\sqrt{66}} y & - \left(\frac{1}{\sqrt{66}} + 1\right) z & = 0. \end{array}$$

The set of solutions is span$((7 - \sqrt{11}, 3\sqrt{6} - 4, \sqrt{66} - \sqrt{6} - \sqrt{11} - 1))$, and this also is the axis of the rotation.

Example 3. *In quantum mechanics, the state of the system is described by its wave function, which assigns to each point in space and time a complex* *number. Let* $\mathbf{W}$ *denote the space of wave functions (which must satisfy further constraints). The evolution of the state* $f \in \mathbf{W}$ *of the system over time is described by Schrödinger's equation*

$$i\frac{\partial}{\partial t}f(x,t) = H(f(x,t)),$$

where the function $H : \mathbf{W} \to \mathbf{W}$*, called the* <u>*Hamiltonian*</u>*, is a linear function encoding the forces acting on the system.*

Suppose that the Hamiltonian does not depend on time. In this case, it turns out that we can find solutions of form $f(x,t) = g(x)h(t)$*; substituting to Schrödinger's equation, we get*

$$\begin{aligned} ig(x)\frac{d}{dt}h(t) &= h(t)H(g(x)) \\ i\frac{1}{h(t)}\frac{d}{dt}h(t) &= \frac{H(g(x))}{g(x)} \end{aligned}$$

Now, the left-hand side depends only on time, while the right hand side depends only on the space coordinates. Therefore, they may only be equal always and everywhere if they are both equal to the same constant λ*. The equation*

$$i\frac{1}{h(t)}\frac{d}{dt}h(t) = \lambda$$

has solution $h(t) = e^{-\lambda(t-t_0)i}$ *for some constant* t_0*. Thus, the full solution is* $f(x,t) = g(x)e^{-\lambda(t-t_0)i}$*, where* g *is the solution to the equation*

$$H(g) = \lambda g.$$

Note that this equation does not necessarily have a solution for all values of λ*. If such a solution exists, it turns out that this constant* λ *is equal to the energy of the system in the state described by* $f(x,t) = g(x)e^{-\lambda(t-t_0)i}$*.*

Definition 2. *Let* A *be a square matrix whose entries are complex numbers. If* $Av = \lambda v$ *for a complex number* λ *and a non-zero vector* v*, then* λ *is an* <u>*eigenvalue*</u> *of* A*, and* v *is the corresponding* <u>*eigenvector*</u>*.*

More generally, λ *is an eigenvalue of a linear function* $f : \mathbf{V} \to \mathbf{V}$ *if* $f(v) = \lambda v$ *for some* $v \neq o$*.*

Observation 7. *If* v *is an eigenvector of* A*, then any non-zero multiple of* v *is an eigenvector of* A *with the same eigenvalue.*

Lemma 8. *A complex number* λ *is an eigenvalue of* A *if and only if* $\det(A - \lambda I) = 0$*; and the set of eigenvectors corresponding to* λ *is Ker*$(A - \lambda I) \setminus \{o\}$*.*

Proof. A complex number λ is an eigenvalue if and only if $Av = \lambda v = \lambda I v$ for some $v \neq o$, which is equivalent to $(A - \lambda I)v = o$. Note that $(A - \lambda I)v = o$ for some non-zero v if and only if $A - \lambda I$ is non-regular, which is equivalent to $\det(A - \lambda I) = 0$. Furthermore, $(A - \lambda I)v = o$ is equivalent to $v \in \operatorname{Ker}(A - \lambda I)$. □

Example 4. *Determine the eigenvalues and eigenvectors of*

$$A = \begin{pmatrix} 3 & 1 & 2 \\ 3 & 3 & 4 \\ -2 & -1 & -1 \end{pmatrix}.$$

$$\det(A - \lambda I) = \det \begin{pmatrix} 3-\lambda & 1 & 2 \\ 3 & 3-\lambda & 4 \\ -2 & -1 & -1-\lambda \end{pmatrix} = -(\lambda - 1)(\lambda - 2)^2.$$

Hence, the eigenvalues are 1 *and* 2*. To obtain the eigenvectors, we need to solve the systems*

$$(A - I)v = \begin{pmatrix} 2 & 1 & 2 \\ 3 & 2 & 4 \\ -2 & -1 & -2 \end{pmatrix} v = o$$

(the set of solutions is span$((0, -2, 1))$*, and thus any non-zero multiple of* $(0, -2, 1)$ *is an eigenvector corresponding to* 1*), and*

$$(A - 2I)v = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 1 & 4 \\ -2 & -1 & -3 \end{pmatrix} v = o$$

(the set of solutions is span$((1, 1, -1))$*, and thus any non-zero multiple of* $(1, 1, -1)$ *is an eigenvector corresponding to* 2*).*

Definition 3. *Let A be a square matrix. Then*

$$p(x) = \det(A - Ix)$$

is the characteristic polynomial of A.

Example 5. *The characteristic polynomial of the matrix from Example 4 is*

$$-(x - 1)(x - 2)^2 = -x^3 + 5x^2 - 8x + 4.$$

Observation 9. *The characteristic polynomial p of an $n \times n$ matrix A has degree exactly n. Each eigenvalue of A is a root of p, and thus A has at most n distinct eigenvalues.*

Theorem 10. *Let A be a square matrix. Let λ be an eigenvalue of A, with a corresponding eigenvector v.*

1. *A is non-regular if and only if 0 is an eigenvalue of A.*
2. *If A is regular, then $1/\lambda$ is an eigenvalue of A^{-1} corresponding to eigenvector v.*
3. *λ^2 is an eigenvalue of A^2 corresponding to eigenvector v.*
4. *$\alpha\lambda$ is an eigenvalue of αA corresponding to eigenvector v.*
5. *$\lambda + \alpha$ is an eigenvalue of $A + \alpha I$ corresponding to eigenvector v.*
6. *λ is an eigenvalue of A^T, possibly with a different eigenvector.*
7. *If A is real, then $\overline{\lambda}$ is an eigenvalue of A corresponding to eigenvector $\overline{v}$.*

Proof.

1. 0 is an eigenvalue of A if and only if $\det(A - 0I) = \det(A)$ is 0, which holds if and only if A is non-regular.
2. We have $v = Iv = A^{-1}Av = A^{-1}(\lambda v) = \lambda(A^{-1}v)$, and thus $A^{-1}v = \frac{1}{\lambda}v$.
3. We have $A^2 v = A(\lambda v) = \lambda(Av) = \lambda^2 v$.
4. We have $(\alpha A)v = \alpha\lambda v$.
5. We have $(A + \alpha I)v = Av + \alpha v = (\lambda + \alpha)v$.
6. A and A^T have the same characteristic polynomial.
7. Since A is real, we have $\overline{A} = A$. Hence, $A\overline{v} = \overline{A}\overline{v} = \overline{Av} = \overline{\lambda v} = \overline{\lambda}\overline{v}$.

□

Observation 11. *If A is an upper or lower triangular matrix with elements $a_1, \ldots, a_n$ on the diagonal, then its characteristic polynomial is $(a_1 - x)(a_2 - x) \cdots (a_n - x)$ and its eigenvalues are $a_1, \ldots, a_n$.*

Definition 4. *Let A be a square matrix, let p be the characteristic polynomial of A, and let λ be an eigenvalue of A. The algebraic multiplicity of λ is its multiplicity as the root of p, that is, the largest integer m such that $(x - \lambda)^m$ divides p. The geometric multiplicity of λ is the dimension of the space $Ker(A - \lambda I)$ of its eigenvectors.*

Example 6. *Let A be the matrix from Example 4. Then 1 is an eigenvalue of A with both algebraic and geometric multiplicity 1, and 2 is an eigenvalue of A with algebraic multiplicity 2 and geometric multiplicity 1.*

Observation 12. *The sum of algebraic multiplicities of eigenvalues of an $n \times n$ matrix is equal to n.*

When we list eigenvalues of a matrix, we list them with their algebraic multiplicities. As we will see later, the geometric multiplicity of each eigenvalue is at most as large as its algebraic multiplicity.

Example 7. *Determine the eigenvalues of the $n \times n$ matrix J whose entries are all equal to 1.*

The matrix J is non-regular, and thus 0 is an eigenvalue of J. Furthermore, $rank(J) = 1$, and thus $\dim(Ker(J)) = n-1$. Hence, 0 is an eigenvalue of J of geometric multiplicity $n-1$.

If j is the vector of all ones, then $Jj = nj$, and thus n is another eigenvalue of 1. The algebraic multiplicity of 0 is at least $n-1$ and the algebraic multiplicity of n is at least 1, and their sum is n; hence, 0 has algebraic multiplicity exactly $n-1$, n has algebraic multiplicity exactly 1, and J has no other eigenvalues.

The eigenvalues of J are $\underbrace{0, \ldots, 0}_{(n-1)\times}, n$.

Definition 5. *The <u>trace</u> of an $n \times n$ matrix A is the sum of its diagonal entries,*

$$Tr(A) = A_{11} + A_{22} + \ldots + A_{nn}.$$

Lemma 13. *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of an $n \times n$ matrix A, and let p be the characteristic polynomial of A. Then*

$$\begin{aligned} p(x) &= (\lambda_1 - x)(\lambda_2 - x) \cdots (\lambda_n - x) \\ \det(A) &= \lambda_1 \lambda_2 \cdots \lambda_n \\ Tr(A) &= \lambda_1 + \ldots + \lambda_n \end{aligned}$$

Proof. Since $\lambda_1, \ldots, \lambda_n$ are the roots of p, we have $p(x) = \alpha(x-\lambda_1)\cdots(x-\lambda_n)$ for some α. Since $p(x) = \det(A - Ix)$, the coefficient of p at x^n is $(-1)^n$, and thus $\alpha = (-1)^n$, and

$$p(x) = (-1)^n (x - \lambda_1) \cdots (x - \lambda_n) = (\lambda_1 - x)(\lambda_2 - x) \cdots (\lambda_n - x).$$

Hence, $\det(A) = p(0) = \lambda_1 \lambda_2 \cdots \lambda_n$.

The coefficient of $p(x) = (\lambda_1 - x) \cdots (\lambda_n - x)$ at x^{n-1} is $(-1)^{n-1}(\lambda_1 + \ldots + \lambda_n)$. On the other hand, the coefficient of $\det(A - Ix)$ at x^{n-1} is $(-1)^{n-1}(A_{11} + A_{22} + \ldots + A_{nn}) = (-1)^{n-1}\mathrm{Tr}(A)$. Hence, $\mathrm{Tr}(A) = \lambda_1 + \ldots + \lambda_n$. □