

Determinants

The determinant $D : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ may be defined from a few initial properties:

- Multilinearity in columns:

$$\begin{aligned} D[\cdots, \mathbf{b}_i + \mathbf{c}_i, \cdots] &= D[\cdots, \mathbf{b}_i, \cdots] + D[\cdots, \mathbf{c}_i, \cdots] \\ D[\cdots, \lambda \mathbf{a}_i, \cdots] &= \lambda D[\cdots, \mathbf{a}_i, \cdots] \end{aligned}$$

- Alternating:

$$D[\cdots, \mathbf{a}_i, \mathbf{a}_{i+1}, \cdots] = 0 \qquad \text{when } \mathbf{a}_i = \mathbf{a}_{i+1}$$

- $D(I_n) = 1$

This may be used to give us the following definition of the determinant:

$$D[\mathbf{a}_1, \cdots, \mathbf{a}_n] = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{k=1}^n a_{k\sigma, k}$$

(this follows immediately from multilinearity and generalising the alternating property to non-adjacent columns)

There are a few ways of getting the more practical inductive definition of the determinant, including just using the above formula. Generally while the practical version is computationally easier to deal with, interpreting the determinant in terms of permutations makes several properties far more obvious than otherwise.

It’s probably useful to note that the determinant is the factor by which areas are squared. For evidence of this, note that any function which determined how a matrix scales area must satisfy the properties of the determinant, and that there is a unique determinantal function.

One of the key properties of the determinant is that it is zero for non-invertible (singular) matrices. To show this, we may first demonstrate that for any matrix A and an elementary matrix E , $\det(EA) = \det(E)\det(A)$. Thus we observe that $\det(S_{ij}) = -1$, $\det(M_i(\lambda)) = \lambda$, $\det(A_{ij}(\lambda)) = 1$, so the determinant is zero if and only if the original matrix does not have I_n as its row reduced echelon form.

Eigenvectors and Eigenvalues

As the simplest linear maps are simply scalings in linearly independent directions, we ideally want to be able to find for any linear map a basis which makes this obvious. Thus the question becomes for which linear maps do such transformations become simple?

Definition 1 *A vector $v \in V$ is called an eigenvector of T if $v \neq 0$ and $Tv = \lambda v$ for some $\lambda \in \mathbb{R}$. We call $\lambda \in \mathbb{R}$ an eigenvalue of T if $Tv = \lambda v$ for some non-zero $v \in V$.*

Claim 1 *λ is an eigenvalue of T if $\ker(T - \lambda I) \neq \{0\}$*

Corollary 1 *The following statements are equivalent:*

- λ is an eigenvalue of T .
- $\ker(T - \lambda I) \neq \{0\}$.
- $T - \lambda I$ is not invertible.
- $\det(T - \lambda I) = 0$.

From this we develop the notion of a characteristic polynomial – to find the eigenvalues of $A \in M_n(\mathbb{R})$ is to find the roots of $\det(A - xI_n)$. This is a degree n polynomial. Note that this is well-defined for a transformation given that for an arbitrary change of basis $P(A - xI_n)P^{-1} = PAP^{-1} - xI_n$, so $\det(PAP^{-1} - xI_n) = \det(P)\det(A - xI_n)\det(P^{-1}) = \det(I_n)\det(A - xI_n)$.

Claim 2 $\text{tr}(A) = \text{tr}(T)$ *is well defined for any transformation T where A represents T with respect to some basis.*

The proof is clear from $\text{tr}(AB) = \text{tr}(BA)$.

Claim 3 (Coefficients of characteristic polynomial) *The first coefficient is $(-1)^n$, the second $(-1)^{n-1}\text{tr}(A)$, and the last $\det A$.*

Proof: $\chi_A(0) = \det A$, and the other coefficients are found via Leibniz’ formula, noticing that only the identity permutation multiplies more than $n - 2$ elements along the diagonal.

From this we observe that $\text{tr } A = \sum_{i=1}^n \lambda_i$, and $\det A = \prod_{i=1}^n \lambda_i$.

Diagonalisation

The usefulness of eigenvalues and eigenvectors comes in the ability to uniquely characterise any linear map by them. We note that for any set of distinct eigenvalues, any set of corresponding eigenvectors must be linearly independent (otherwise for a minimal non-LI subset of eigenvectors, applying $T - \lambda I$ gives a smaller subset of non-LI eigenvectors).

If we have a basis of eigenvectors then we may straightforwardly change coordinates to diagonalise the linear map. Thus diagonalisability is equivalent to the existence of a basis of eigenvectors.

Geometric and algebraic multiplicity
For a linear transformation $T : V \rightarrow V$, and an eigenvalue λ for T , <p style="text-align: center;">$E_\lambda = \ker(T - \lambda I) = \{v \in V \mid Tv = \lambda v\}$.</p> To characterise the existence of eigenvectors and eigenvalues for T , we refer to the notions of geometric and algebraic multiplicity. The geometric multiplicity is the dimension of E_λ , while the algebraic multiplicity is the multiplicity of λ as a root of $\chi_T(x)$.
We have that the algebraic multiplicity must be at least the geometric multiplicity, as if we take a basis for E_λ and extend it to a basis for V , the corresponding matrix for T has top left $g_\lambda \times g_\lambda$ matrix as λI_{g_λ} (all zeroes below), so $\chi_T(x) = (\lambda - x)^{g_\lambda} h(x)$.

Spectral Theorem

Recall from Linear Algebra I that a basis is orthonormal with respect to some inner product if for any pair of distinct bases their inner product is 0, and the norm of any base is 1. By the Gram-Schmidt procedure we demonstrate for real symmetric matrices that there exists an orthonormal basis of eigenvectors.

Given $\{u_1, \dots, u_n\}$ a basis for \mathbb{R}^n , we construct $\{v_1, \dots, v_n\}$ with the property that for all $k \in \{1, \dots, n\}$, $\langle u_1, \dots, u_k \rangle = \langle v_1, \dots, v_k \rangle$.

$$\begin{aligned} w_1 &:= u_1 & v_1 &:= \frac{w_1}{||w_1||} \\ w_2 &:= u_2 - (u_2 \cdot v_1)v_1 & v_2 &:= \frac{w_2}{||w_2||} \\ \dots & & \dots & \\ w_k &:= u_k - \sum_{j=1}^{k-1} (u_k \cdot v_j)v_j & v_k &:= \frac{w_k}{||w_k||} \\ \dots & & \dots & \\ w_n &:= u_n - \sum_{j=1}^{n-1} (u_n \cdot v_j)v_j & v_n &:= \frac{w_n}{||w_n||} \end{aligned}$$

We may then show that all eigenvalues of a real symmetric matrix must also be real, and by induction using the Gram-Schmidt process we can generate a coordinate change producing a diagonal matrix.

To see the importance of this result, note that symmetric matrices are characterised by the property that $(Au) \cdot v = u \cdot (Av)$. Further, for any inner product $\langle -, - \rangle : V^2 \rightarrow \mathbb{R}$, a map $T : V \rightarrow V$ is self-adjoint if $\langle Tu, v \rangle = \langle u, Tv \rangle$ for all $u, v \in V$. Thus, by the spectral theorem a self-adjoint map T on a finite-dimensional real inner product space V has real eigenvalues, and there exists an orthonormal basis for V consisting of eigenvectors of T .

Inner products on finite-dimensional real vector spaces are a basis free way of discussing the dot product.

Quadrics
One useful application of the spectral theorem is for the purpose of classifying quadratic forms. With an equation of the form $Q(x_1, \dots, x_n) = \sum_{i=1}^n \sum_{j=1}^n a_{ij}x_ix_j = (x_1 \cdots x_n) A \begin{pmatrix} x_1 \\ \cdots \\ x_n \end{pmatrix}, \quad A = (a_{ij})$ we can write A as symmetric without loss of generality ($\frac{1}{2}(A + A^\top)$). Consequently we may diagonalise the equation via a change of variables.
Within \mathbb{R}^3 , we thus get equations of the form $\mu_1 Y_1^2 + \mu_2 Y_2^2 + \mu_3 Y_3^2 = c \in \{-1, 0, 1\}$. For μ_1, μ_2, μ_3 all positive, we get an ellipsoid for $c = 1$, an empty set for c negative, and $\{0\}$ for $c = 0$. For one μ_i negative (without loss of generality, as with two negative multiply by -1), we get a 1-sheet hyperboloid with $c = 1$, a 2-sheet hyperboloid with $c = -1$, and a cone for $c = 0$ (the crossover point).
The general definition of a quadric is as the set of roots in \mathbb{R}^3 of a degree 2 equation $\begin{aligned} f(x_1, x_2, x_3) &= \sum_{i=1}^3 \sum_{j=1}^3 a_{ij}x_ix_j + \sum_{i=1}^3 b_ix_i + c = 0 \\ &= \mathbf{x}^\top A \mathbf{x} + \mathbf{b}^\top \mathbf{x} + c && \text{for some symmetric } A \in M_3(\mathbb{R}) \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2 + \mathbf{B}^\top \mathbf{y} + c && \mathbf{y} = P^\top \mathbf{x} \\ &= \lambda_1 Y_1^2 + \lambda_2 Y_2^2 + \lambda_3 Y_3^2 + C && Y_i = y_i + \frac{B_i}{2\lambda_i} \end{aligned}$