Robert Andrew Martin

Linear Algebra

Vector Spaces

- A matrix can be thought of as a linear relationship between two vectors.
- Scalars are elements of a number field, e.g $\mathbb{R}$ or $\mathbb{C}$.
- A field is a set of elements on which addition and multiplication are defined, and are commutative, associative and distributive.
  - ↳ closed under add/mult
  - ↳ includes identity elements (0 for add, 1 for mult)
  - ↳ includes inverses for every element (except zero).
- Vectors are elements of a vector space, defined over some number field
  - ↳ vector addition and inner product are defined
  - ↳ closed under these ops
  - ↳ includes identity element for addition
- Let $S = \{\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots \underset{\sim}{e}_m\}$ be a subset of some vector space V. The span of S is the set of all vectors that are linear combinations of S.
- The vectors of S are linearly independent if no nontrivial linear combination of the vectors is zero
  i.e $\underset{\sim}{e}_i x_i = 0 \Rightarrow x_1 = \dots = x_n = 0$
- A basis is a set of linearly independent vectors that spans the space
  - ↳ all bases of V have the same number of elements – the dimension of the space.
  - ↳ any vector $x \in V$ can be written uniquely as $\underset{\sim}{e}_i x_i$
- We can convert between bases with a transformation matrix

  $\underset{\sim}{e}_j = \underset{\sim}{e}'_i R_{ij} \Rightarrow \underset{\sim}{e}'_i x'_i = \underset{\sim}{e}_j x_j = \underset{\sim}{e}'_i R_{ij} x_j$

Linear operators

- A linear operator $\underset{\sim}{A}$ acts on a vector space to produce other elements of V.

  $$A(\alpha \underset{\sim}{x} + \beta \underset{\sim}{y}) = \alpha A \underset{\sim}{x} + \beta A \underset{\sim}{y}$$

- A matrix is an array of numbers that can represent a linear operator. It contains the components of the operator with respect to a certain basis
  - ↳ because $\underset{\sim}{A}$ is linear, knowledge of its action on a basis is sufficient to know its action on any vector in the space

    $$\underset{\sim}{A}\,\underset{\sim}{x} = \underset{\sim}{e}_i A_{ij} x_j$$

- We can rewrite A in a different basis as follows:
  - ↳ we require $\underset{\sim}{e}_i A_{ij} x_j = \underset{\sim}{e}_i' A_{ij}' x_j'$
  - ↳ by using $\underset{\sim}{e}_j = \underset{\sim}{e}_i' R_{ij}$ and relabelling indices

$$\underset{\sim}{e}_k' R_{ki} A_{ij} x_j = e_k' A_{kj}' x_j'$$
$$\Rightarrow RA(R^{-1}\underset{\sim}{x}') = A'\underset{\sim}{x}'$$
$$\therefore \boxed{A' = RAR^{-1}}$$

Inner products

- The inner product $\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle$ is a scalar function of two vectors. It must:
  - be bilinear, i.e linear in the second argument and antilinear in the first:

$$\langle \underset{\sim}{x} | \alpha \underset{\sim}{y} \rangle = \alpha \langle \underset{\sim}{x} | \underset{\sim}{y} \rangle \quad \text{and} \quad \langle \alpha \underset{\sim}{x} | \underset{\sim}{y} \rangle = a^* \langle \underset{\sim}{x} | \underset{\sim}{y} \rangle$$

  - have Hermitian symmetry:

$$\langle \underset{\sim}{y} | \underset{\sim}{x} \rangle = \langle \underset{\sim}{x} | \underset{\sim}{y} \rangle^*$$

  - be positive definite:

$$\langle \underset{\sim}{x} | \underset{\sim}{x} \rangle \geq 0 \quad \leftarrow \text{equality iff } \underset{\sim}{x} = 0$$

  - ↳ distributive in the first argument:

$$\langle \underset{\sim}{x} + \underset{\sim}{y} | \underset{\sim}{z} \rangle = \langle \underset{\sim}{x} | \underset{\sim}{z} \rangle + \langle \underset{\sim}{y} | \underset{\sim}{z} \rangle$$

- In $\mathbb{R}^n$, $\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle = x_i y_i$.
- In $\mathbb{C}^n$, $\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle = x_i^* y_i$.

- The Cauchy-Schwarz inequality states:

$$|\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle|^2 \leq \langle \underset{\sim}{x} | \underset{\sim}{x} \rangle \langle \underset{\sim}{y} | \underset{\sim}{y} \rangle$$

or

$$\boxed{|\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle| \leq |\underset{\sim}{x}||\underset{\sim}{y}|}$$

  - ↳ equality when $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are linearly dependent
  - ↳ can be proven by considering $\langle \underset{\sim}{x} - a\underset{\sim}{y} | \underset{\sim}{x} - a\underset{\sim}{y} \rangle$ then later setting $|\alpha| = |\underset{\sim}{x}| / |\underset{\sim}{y}|$
  - ↳ we can use Cauchy-Schwarz to define $\cos\theta$ in $\mathbb{R}^n$

$$\cos\theta \equiv \frac{\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle}{|\underset{\sim}{x}||\underset{\sim}{y}|}$$

Hermitian matrices

- The Hermitian conjugate of a matrix is the complex conjugate of its transpose

$$A^\dagger \equiv (A^T)^* \quad \therefore (A^\dagger)_{ij} = A_{ji}^*$$

  - ↳ it obeys similar rules to the transpose:

$$(A^\dagger)^\dagger = A \qquad (AB)^\dagger = B^\dagger A^\dagger$$

  - ↳ the Hermitian conjugate of a scalar is just the conjugate, e.g $\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle^\dagger = \langle \underset{\sim}{x} | \underset{\sim}{y} \rangle^*$
  - ↳ a matrix is Hermitian if $A = A^\dagger$

Robert Andrew Martin

- Knowing $\langle \underset{\sim}{e}_i | \underset{\sim}{e}_j \rangle$ for basis vectors is sufficient to know $\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle$ because of bilinearity. If $\langle \underset{\sim}{e}_i | \underset{\sim}{e}_j \rangle = G_{ij}$:
$$\langle \underset{\sim}{x} | \underset{\sim}{y} \rangle = \langle \underset{\sim}{e}_i x_i | \underset{\sim}{e}_j y_j \rangle = x_i^* y_j G_{ij}$$
  ↳ G_{ij} are the metric coefficients
  ↳ G is Hermitian since $G_{ij} = G_{ji}^*$
  ↳ a basis is orthonormal if $\langle \underset{\sim}{e}_i | \underset{\sim}{e}_j \rangle = \delta_{ij}$

- The adjoint of a linear operator $\underset{\sim}{A}$ with respect to some inner product is another linear operator $\underset{\sim}{A}^{\dagger}$ such that:
$$\langle A^{\dagger} \underset{\sim}{x} | \underset{\sim}{y} \rangle = \langle \underset{\sim}{x} | A \underset{\sim}{y} \rangle$$
  ↳ for a given basis the components of $\underset{\sim}{A}^{\dagger}$ are the entries in the matrix $A^{\dagger}$.

Matrix Symmetry	Equation
symmetric	$A^T = A$
antisymmetric	$A^T = -A$
orthogonal	$AA^T = A^T A = I$
Hermitian	$A^{\dagger} = A$
anti-Hermitian	$A^{\dagger} = -A$
unitary	$A^{\dagger} = A^{-1}$
normal	$AA^{\dagger} = A^{\dagger} A$

complex analogues

these are all normal

Eigenvalues and Eigenvectors

- An eigenvector of a linear operator is a nonzero vector $\underset{\sim}{x}$ such that $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$. We can find the eigenvalues and eigenvectors by solving the characteristic equation $\det(A - \lambda I) = 0$
  ↳ if the n roots are distinct, there are n linearly independent eigenvectors (unique to a constant factor)
  ↳ if an eigenvalue is degenerate and occurs m times, there may be between 1 and m linearly independent eigenvectors for that eigenvalue, spanning the eigenspace
- In general, we can prove eigenvalue/eigenvector properties as follows, using the example of Hermitian matrices
  ↳ Consider two eigenvalue/vector pairs
  ① $Ax = \lambda x$ ② $Ay = \mu y$
  ↳ Take Hermitian conjugate of ② $\therefore y^{\dagger} A^{\dagger} = \mu^* y^{\dagger}$
  then use Hermitian property $\Rightarrow y^{\dagger} A = \mu^* y^{\dagger}$
  ↳ apply $y^{\dagger}$ to ① to get two expressions for $y^{\dagger} A x$
  $\therefore (\lambda - \mu^*) y^{\dagger} x = 0$
  ↳ suppose x and y are the same eigenvector (and $\lambda = \mu$)
  $\Rightarrow (\lambda - \lambda^*) x^T x = 0$.
  ↳ $x^T x \neq 0 \therefore \lambda = \lambda^*$
  ↳ so eigenvalues of Hermitian matrix are real

↳ if x and y are different eigenvectors:

$$(\lambda-\mu)y^\dagger x = 0$$

$$\Rightarrow y^\dagger x = 0 \quad \text{for } \lambda \neq \mu$$

↳ hence eigenvectors orthogonal for different eigenvalues.

$\boxed{\text{• The eigenvectors of normal matrices corresponding to distinct eigenvalues are orthogonal.}}$

↳ if A is Hermitian, iA is anti-Hermitian (& vice versa)

↳ if A is Hermitian, $\exp(iA)$ is unitary

↳ an eigenbasis can ALWAYS be constructed for a normal matrix (even if there are degenerate eigenvalues)

Symmetry	Eigenvalues	Interpretation
Hermitian	$\lambda^* = \lambda$	Real
anti-Hermitian	$\lambda^* = -\lambda$	Imaginary
Unitary	$\lambda^* = 1/\lambda$	Unit modulus

i.e. purely rotational

Diagonalisation of a matrix

- If $\underset{\sim}{x}' = R\underset{\sim}{x}$ for a transformation between basis vectors a linear operator can be transformed via $A' = RAR^{-1}$
- Two square matrices are *similar* if $B = S^{-1}AS$ where S is some invertible *similarity* matrix.
- A matrix is *diagonalisable* if it is similar to a diagonal matrix, i.e.: $A = S\Lambda S^{-1}$
- To diagonalise, we form S from the eigenvectors of A. The entries of Λ are then the corresponding eigenvalues:

$$S = \begin{pmatrix} \underset{\sim}{x}^{(1)} & \underset{\sim}{x}^{(2)} & \underset{\sim}{x}^{(3)} \\ \downarrow & \downarrow & \downarrow \end{pmatrix}$$

$$S^{-1}AS = S^{-1}A\begin{pmatrix} x^{(1)} & \cdots & x^{(n)} \\ \downarrow & \downarrow & \downarrow \end{pmatrix} = S^{-1}\begin{pmatrix} Ax^{(1)} & \cdots & Ax^{(n)} \end{pmatrix}$$

$$= S^{-1}\begin{pmatrix} x^{(1)} & \cdots & x^{(n)} \\ \downarrow & \downarrow & \downarrow \end{pmatrix}\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} = \Lambda$$

↳ note that S can only be inverted if its columns are linearly independent $\Rightarrow$ S is diagonalisable if and only if A has n linearly independent eigenvectors

- Thus normal matrices are diagonalisable, and the eigenvectors can be chosen to be orthonormal
- Intuitively, diagonalisation is the process of expressing a matrix in its eigenbasis – the simplest form. Hence the similarity matrix is unitary and $\boxed{A = U\Lambda U^\dagger}$
- Diagonalisation is useful because some operations are much easier to carry out on the diagonalised repr. $A = S\Lambda S^{-1}$:

$$A^m = S\Lambda^m S^{-1}$$

$$\left.\begin{aligned} \det A &= \det\Lambda \\ \operatorname{tr} A &= \operatorname{tr}\Lambda \end{aligned}\right\}$$

For any matrix: $\det A = \prod_i \lambda_i$, $\operatorname{tr} A = \sum_i \lambda_i$

- The transformation between orthonormal bases is described by a unitary matrix
  - ↳ a real symmetric matrix can be diagonalised by a real orthogonal transformation

Quadratic forms

- The quadratic form associated with a real symmetric matrix A is $Q(x) = x^T A x = A_{ij} x_i x_j$ ← hence the name
- Q is a homogeneous quadratic function of x, i.e $Q(\alpha x) = \alpha^2 Q(x)$. Any homogeneous quadratic is the quadratic form of some symmetric matrix.
- Because real symmetric matrices can be diagonalised by orthogonal transformations:

$$Q(x) = x^T A x = x'^T \Lambda x \, , \qquad x = S x'$$

  - ↳ the eigenvectors of A are the principal axes — in the eigenbasis, the quadratic form is just a sum of squares $Q = \lambda_i x_i'^2$
- Quadratic forms can represent quadratic surfaces

$$Q(x) = k \ (= \text{constant})$$

  - ↳ hence representing Q in its eigenbasis allows us to easily identify the shape.

- Given $\lambda_1 x^2 + \lambda_2 y^2 + \lambda_3 z^2 = k$:
  - λs have same sign ⇒ ellipsoid
  - λs have mixed sign ⇒ hyperboloid
  - $\lambda_1 = \lambda_2 = \lambda_3$ ⇒ sphere
  - $\lambda_1 = \lambda_2$ ⇒ surface of revolution about z axis
  - $\lambda_3 = 0$ ⇒ translation of conic section along z axis

Hermitian forms

- The Hermitian form is a complex extension of the quadratic form: $H(x) = x^\dagger A x$ (real scalar quantity)
- Hermitian matrices can be diagonalised with unitary transformations ⇒ $H(x) = x'^\dagger \Lambda x' = \lambda_i |x_i|^2$
- The Rayleigh quotient associated with a Hermitian matrix is the normalised Hermitian form: $\boxed{\lambda(x) = \dfrac{x^\dagger A x}{x^\dagger x}}$
  - ↳ if x is an eigenvector of A, λ is an eigenvalue (easily verified by substitution)
  - ↳ the Rayleigh-Ritz variational principle considers $\delta\lambda = \lambda(x + \delta x) - \lambda(x)$ and shows that the eigenvectors of A are the stationary points of $\lambda(x)$

Robert Andrew Martin

Cartesian Tensors

- In Cartesians, basis vectors are independent of position.
- To transform from basis vectors $\hat{\underline{e}}_j$ to $\hat{\underline{e}}_i$:

$$\underline{v} = v_j\hat{\underline{e}}_j = v_i'\hat{\underline{e}}_i' \Rightarrow v_i' = \hat{\underline{e}}_i'\cdot\underline{v} = \hat{\underline{e}}_i'\cdot\hat{\underline{e}}_j\, v_j$$

$$\boxed{v_i' = L_{ij}v_j \quad \text{with} \quad L_{ij} = \hat{\underline{e}}_i'\cdot\hat{\underline{e}}_j}$$

↳ L is the transformation matrix ← rotates the frame

↳ the same argument applies when interchanging v' and v. So $L^TL = LL^T = I \Rightarrow L$ is orthogonal

- A Cartesian vector v is defined as a set of coefficients v_i with respect to an orthonormal basis $\hat{\underline{e}}_i$ such that an orthogonal transformation transforms to another orthonormal basis $\hat{\underline{e}}_i'$, with coefficients v_i'
- Orthogonal matrices have determinant ± 1:

↳ $\det L = +1$ is a proper rotation

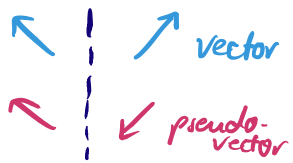
↳ $\det L = -1$ is an improper rotation (i.e rotation + reflection)

↳ if $L^{(1)}$ and $L^{(2)}$ are proper rotations, their composition is also a proper rotation: $v_i'' = L^{(1)}_{ij}L^{(2)}_{jk}v_k$

- A Cartesian pseudovector transforms via $\boxed{a_i' = \det L\, L_{ij}a_j}$

↳ i.e gains a sign change under any reflection (change of handedness)

↳ cross products are always pseudovectors.



Tensors

- A tensor of order (rank) n transforms between two orthonormal basis sets as described by the transformation law

$$\boxed{T'_{i_1\dots i_n} = L_{i_1j_1}\cdots L_{i_nj_n}T_{j_1\dots j_n}}$$

- The order of a tensor is equal to the number of indices needed to lablel it. Scalars are order zero; vectors are order 1; matrices are order 2.
- Pseudotensors are defined with an additional $\det L$ factor, changing the sign during reflection.
- The Kronecker delta is a second-order tensor:

$$\delta_{ij}' = L_{ip}L_{jq}\delta_{pq} = L_{ip}L_{jp} = \delta_{ij}$$ L is orthogonal.

- The Levi-Civita symbol is a third-order pseudotensor. This can be shown by verifying that one of the nonzero terms stays constant under a transformation:

$$\varepsilon'_{123} = \det L\, L_{1p}L_{2q}L_{3r}\varepsilon_{pqr} = (\det L)^2 = 1$$

- The inertia tensor relates the angular momentum $\underline{J}$ to the angular velocity $\underline{\omega}$. $d\underline{J} = dm\, \underline{x}\times(\underline{\omega}\times\underline{x}) = dm(|\underline{x}|^2\underline{\omega} - (\underline{\omega}\cdot\underline{x})\underline{x})$

$$\Rightarrow J_i = I_{ij}\omega_j \quad \text{with} \quad \boxed{I_{ij} = \int_V \rho(\underline{x})(x_kx_k\delta_{ij} - x_ix_j)\,dV}$$

- Susceptibility tensors (2nd order) relate the polarisation to the applied E-field. $P_i = \varepsilon_0\chi_{ij}E_j$

- Elastic deformation is described by the strain tensor ε_{ij}

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial x_i}{\partial x_j} + \frac{\partial x_j}{\partial x_i}\right)$$

↳ the associated stress tensor σ_{ij} defines the jth component of force on a plane perpendicular to i.

↳ they are related by a fourth-order stiffness tensor.

Properties of tensors

- If A and B are order-n tensors, then so is any linear combination of them. Proof: transform $C = \alpha A + \beta B$

$$C'_{i_1 \ldots i_n} = \alpha A'_{i_1 \ldots i_n} + \beta B'_{i_1 \ldots i_n}$$
$$= \alpha L_{i_1 j_1} \ldots L_{i_n j_n} A_{j_1 \ldots j_n} + \beta L_{i_1 j_1} \ldots L_{i_n j_n} B_{j_1 \ldots j_n}$$
$$= L_{i_1 j_1} \ldots L_{i_n j_n} (\alpha A_{j_1 \ldots j_n} + \beta B_{j_1 \ldots j_n})$$

- The tensor product of tensors of order n and m is a tensor of order $n+m$. ← also called outer product.

$$C = A \otimes B \Rightarrow C_{i_1 \ldots i_n i_{n+1} \ldots i_{n+m}} = A_{i_1 \ldots i_n} B_{i_{n+1} \ldots i_{n+m}}$$

↳ a general tensor can be written as $T = T_{i_1 \ldots i_n} \underline{e}_{i_1} \otimes \ldots \otimes \underline{e}_{i_n}$

↳ tensor $\otimes$ pseudotensor = pseudotensor

- A tensor contraction sets two indices equal and sums over, returning a tensor of order $n-2$.
- A tensor is symmetric in a pair of indices if $T_{\ldots i \ldots j \ldots} = T_{\ldots j \ldots i \ldots}$ and antisymmetric if $T_{\ldots i \ldots j \ldots} = -T_{\ldots j \ldots i \ldots}$. The (anti)symmetry of a tensor is invariant under a change of coordinates.
- If S_{ijk} is symmetric in i,j and A_{pqr} is antisymmetric in p,q, then the contraction $S_{ijk} A_{ijr} = 0$.

Second-order tensors

- 2nd order tensors can be represented as matrices and thus have matrix properties
- An antisymmetric second-order tensor is equivalent to a certain pseudovector – the dual vector.

$$A_{ij} = \varepsilon_{ijk} \omega_k = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix}$$

- Any symmetric second-order tensor can be uniquely written as the sum of a symmetric traceless tensor and a scalar multiple of the identity tensor:

$$S = \underbrace{S - \tfrac{1}{3}\operatorname{tr} S\, I}_{\text{traceless.}} + \tfrac{1}{3}\operatorname{tr} S\, I$$

- Symmetric second-order tensors can be diagonalised.

Isotropic tensors

- Isotropic tensors are invariant with respect to the frame and thus have no preferred direction.
- 0th order: all scalars are isotropic (transformation law)
- 1st order: only the zero vector is isotropic
- 2nd order: $\lambda \delta_{ij}$ for scalar λ

- 3^{rd} order: $\lambda \varepsilon_{ijk}$ for scalar λ
- 4^{th} order: $\lambda \delta_{ij}\delta_{kl} + \mu \delta_{ik}\delta_{jl} + \nu \delta_{il}\delta_{jk}$ for scalar λ, μ, ν.

- Isotropy may be used to evaluate integrals when the integration region is symmetric

$r' = r$, $dV' = dV$

$$X_i = \int_{r' \le a} x_i' \rho(r') dV' = \int_{r \le a} R_{ij} x_j \rho(r) dV = R_{ij} X_j = X_i'$$

↳ $X_i = R_{ij} X_j$ for general R_{ij} means X_i is isotropic

↳ the only isotropic vector is the zero vector, so $\underset{\sim}{X} = \underset{\sim}{0}$.

- E.g for a second-order tensor integral:

$$K_{ij} = \int_{r' \le a} x_i' x_j' \rho(r) dV = R_{ik} R_{jl} K_{kl} = K_{ij}'$$

$$\Rightarrow K_{ij} = \lambda \delta_{ij} \quad \text{with} \quad \lambda = \tfrac{1}{3} \operatorname{Tr} K$$

$$\Rightarrow K_{ij} = \left(\int_{r \le a} \tfrac{1}{3} r^2 \rho(r) dV \right) \delta_{ij}$$

Tensor fields

- A tensor field assigns a tensor to every position $\underset{\sim}{x}$ (in some domain)
  e.g a conductivity field (2^{nd} order tensor field)
- The divergence of a vector field is scalar - the contraction of the tensor product of two vector fields ∂_i and F_j
- $\nabla \times \underset{\sim}{F}$ is a pseudovector field – the contraction of the tensor product of pseudotensor ε_{ijk} and vectors ∂_l, F_m.
- The derivative of a second-order tensor field is a third-order tensor field $\partial_i \sigma_{jk}$.