Math Camp

Module #2: Vector spaces and linear maps

Part II: enacting violence

Remember: “The introduction of numbers as coordinates is an act of violence.” *still H. Weyl*

Coordinates

Let V be a real vector space of dimension n, and let $A = \{a_1, ..., a_n\}$ be a basis for V.

- $v \in V \implies \exists! \{\alpha_k^v\}_{i=k}^n$ s.t. $v = \sum_{k=1}^{n} \alpha_k^v a_k$
- The scalars $\{\alpha_1^v, \ldots, \alpha_n^v\}$ comprise the *coordinate representation* of v with respect to the basis A

Dirac function : given two sets X and Y, $\delta : X \times Y \to \{0, 1\}$ is defined by $\delta_{xy} = 1$ if and only if $x = y$.

- $X = Y = \mathbb{Z}$ is illustrative: $\delta_{ij} = 1 \Leftrightarrow i = j$

$\mathbb{R}^n$

- $x \in \mathbb{R}^n$ implies $x = (x_1, \ldots, x_n)$ with $x_i \in \mathbb{R}$
- Always think of x as a column
- The *canonical basis* for $\mathbb{R}^n$ is $\mathscr{E} = \{e_1, \ldots, e_n\}$, where $e_j = (e_{1j}, \ldots, e_{nj})$ and $e_{ij} = \delta_{ij}$.
- Let $\dim V = n$ with basis A. Define $\varphi : V \to \mathbb{R}^n$ by setting $\varphi(a_i) = e_i$.

$\mathbb{R}^{m\times n}$

- A real $m \times n$ matrix is a rectangular array of real numbers with m rows and n columns
- If $A \in \mathbb{R}^{m\times n}$ then $A = (a_{ij})$, where $i = 1,\ldots,m$ and $j = 1,\ldots,n$, I.e.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

- $\mathbb{R}^{m\times n}$ is a vector space with canonical basis $\{e^{ij}\}$ where

$$(e^{ij}_{kl}) = \delta_{(i,j)(k,l)}$$

Linear functionals

Let V be a real vector space

- A *linear functional* is a linear map from V to $\mathbb{R}$
- The *dual space* V^* of V is the vector space of linear functionals from V to $\mathbb{R}$
- Let $\dim V = n$ with basis $A = \{a_1, \ldots, a_n\}$. Define $a_i^* : A \to \mathbb{R}$ on A by $a_i^*(a_j) = \delta_{ij}$, and extend linearly. Then A^* is a basis for V^*.
- The coordinate representation of $v^* \in V^*$ with respect to A^* is $(v^*(a_1), \ldots, v^*(a_n)) \in \mathbb{R}^n$.

Inner products

An *inner product* on a real vector space V is a symmetric, positive definite, bilinear form, i.e. a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ such that

- $\langle v, v \rangle \geq 0$ with equality only when $v = 0$ (positive definiteness)
- $\langle v, w \rangle = \langle w, v \rangle$ (symmetry)
- For any $v \in V$, the maps $\langle v, \cdot \rangle : V \to \mathbb{R}$ and $\langle \cdot, v \rangle : V \to \mathbb{R}$ are linear (bilinearity). Thus

$$\langle v, \alpha w_1 + \beta w_2 \rangle = \alpha \langle v, w_1 \rangle + \beta \langle v, w_2 \rangle$$
$$\langle \alpha w_1 + \beta w_2, v \rangle = \alpha \langle w_1, v \rangle + \beta \langle w_2, v \rangle$$

Inner products and linear functionals on $\mathbb{R}^n$

For $v, w \in \mathbb{R}^n$, define $\langle v, w \rangle = \sum_{i=1}^n v_i w_i$.

- If $v \in \mathbb{R}^n$ then $v^* = \langle \cdot, v \rangle \in (\mathbb{R}^n)^*$.
- If v^* in $(\mathbb{R}^n)^*$ then there exists $v \in \mathbb{R}^n$ such that $v^* = \langle \cdot, v \rangle$
- $\mathbb{R}^n \cong (\mathbb{R}^n)^*$ with the isomorphism given by $v \to \langle \cdot, v \rangle$
- The kernel of v^* is the subspace of $\mathbb{R}^n$ *orthogonal* to v.
- Inner products impart geometry

Row vectors are linear functionals

Let $v \in \mathbb{R}^n$, viewed as a column vector. Then v^T is the corresponding row vector.

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \implies v^T = (v_1, \ldots, v_n)$$

- $v^T w = \sum_{i=1}^{n} v_i w_i = v_j w_j = \langle v, w \rangle$
- $v \in \mathbb{R}^n$ implies $v^T \in (\mathbb{R}^n)^*$

Linear maps are columns of linear functionals

Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be any function.

- $f(x) = \begin{pmatrix} f^1(x) \\ \vdots \\ f^m(x) \end{pmatrix}$ where $f^i : \mathbb{R}^n \to \mathbb{R}$
- If f is linear then f^i is a linear functional
- If f is linear then f is a column vector of row vectors

Matrices are linear maps

Let $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_m\}$ be bases of V and W respectively. Let $T : V \to W$ be linear.

Define the $m \times n$ matrix $\beta(T)$ as follows: for $1 \leq j \leq n$, the j^{th}-column of $\beta(T)$ is the coordinate representation of $T(a_j) \in \mathbb{R}^m$ against the basis B: $T(a_j) = \sum_{i=1}^{m} \beta(T)_{ij} b_i$. Then

$$
\begin{aligned}
T(v) = T\left(\sum_{j=1}^{n} \alpha_j^v a_j\right) &= \sum_{j=1}^{n} \alpha_j^v T(a_j) \\
&= \sum_{j=1}^{n} \alpha_j^v \left(\sum_{i=1}^{m} \beta(T)_{ij} b_i\right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \beta(T)_{ij} \alpha_j^v\right) b_i
\end{aligned}
$$

Thus the coordinate representation of $T(v)$ against B is $\beta(T)\alpha^v$.

Matrices as linear maps

Matrices are exactly linear maps represented against bases

$$\begin{tikzcd}
V \arrow[r, "T"] \arrow[d, "\varphi_A"'] & W \\
\mathbb{R}^n \arrow[r, "\beta(T)\alpha^\nu"'] & \mathbb{R}^m \arrow[u, "\varphi_B^{-1}"']
\end{tikzcd}$$

Here φ_A and φ_B are the canonical isomorphisms and $\beta(T)\alpha^\nu$ is obtained by the matrix multiplication.

In particular, matrix multiplication is composition of linear maps.

The Transpose

Definition. The *transpose* A^T of an $m \times n$ matrix A is an $n \times m$ is given by $a^T_{ij} = a_{ji}$.

- If $v \in \mathbb{R}$ is viewed as a column matrix then v^T can be viewed as a row matrix

- Under matrix multiplication, a row vector is a linear functional.

- $\langle v, w \rangle = v^T w$.

The Transpose

If $S : V \to W$ then $S^* : W^* \to V^*$ is given by $S^*(w^*)(v) = w^*(S(v))$.
The following diagram commutes:

$$\begin{tikzcd} V \arrow[r, "S"] & W \\ V^* \arrow[u, "\varphi_V"] & W^* \arrow[l, "S^*"] \arrow[u, "\varphi_W"'] \end{tikzcd}$$

Fixing bases, $\beta(S)^T = \beta(S^*)$.

The Determinant

The *determinant* is a map $\det : \mathbb{R}^{n\times n} \to \mathbb{R}$ that is magical:

- $\det(I_n) = 1$
- $\det(AB) = \det(A)\det(B)$
- $\det(\alpha A) = \alpha^n \det(A)$

There are two very important points worth emphasizing

- the determinant is a polynomial of degree n in its entries
- because it is a polynomial in its entries, the determinant is computable

The Determinant

There are two ways to compute the determinant: geometrically, and using eigenvalues

Geometry.

- Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear function (square matrix) and let $S \subset \mathbb{R}^n$ be the unit cube, that is, $S = [0,1]^n$
- Because T is linear, the image of S under T is an m-dimensional parallelpiped, where m is the rank of T.
- The determinant of T is the signed volume of this parallelepiped.
- $det(T) = 0 \Leftrightarrow dim(ker(T)) > 0$

The Determinant

Eigenvalues.

- Recall that $\lambda \in \mathbb{R}$ is an eigenvalue of T provided there is $v \neq 0$ such that $T(v) = \lambda v$, or $(T - \lambda I_n)(v) = 0$.
- Thus $\dim ker(T - \lambda I_n) > 0$, whence $\phi_T(\lambda) \equiv \det(T - \lambda I_n) = 0$
- $\phi_T(\lambda)$ is the *characteristic polynomial* of T
- The eigenvalues of T are the roots of this polynomial, and may be complex.

The Determinant

- $\phi_T(\lambda) \equiv \det(T - \lambda I_n)$
- Every $n \times n$ matrix has exactly n eigenvalues corresponding to the n roots of the characteristic polynomial
- The determinant of T is equal to the product of the eigenvalues
- $\det(T) = 0$ iff zero is an eigenvalue of T

Invertibility

Given sets X and Y, let $f : X \to Y$ be any map.

- f is *injective* if $f(x_1) = f(x_2)$ implies $x_1 = x_2$
- f is surjective provided that for any $y \in Y$ there is an $x \in X$ so that $y = f(x)$
- a function that is both surjective and injective is *bijective*
- Bijectivity is necessary and sufficient for invertibility of a linear map
- If $\dim V < \infty$ then linear map from V to V is invertible if and only if its nullity is zero.

Invertibility

A matrix is invertible provided that the associated linear map is invertible

Theorem 3.2 A square matrix is invertible if and only if its determinant is non-zero

Why is this theorem so important?

Column and Row Space

Let $M \in \mathbb{R}^{m \times n}$, and denote by $\{M^i\}_{i=1}^m \subset \mathbb{R}^n$ the rows of M (but "written" as column vectors), and by $\{M_j\}_{j=1}^n \subset \mathbb{R}^m$ the columns of M.

- the row space of M is $\text{span}\left(\{M^i\}_{i=1}^m\right)$
- the column space of M is $\text{span}\left(\{M_j\}_{j=1}^n\right)$
- the columns of M span the range of the associated linear map
- the dimension of the column and row space are equal

Column and Row Space

Let $T : V \to W$ linear.

$$\begin{aligned}
\text{dim of row space of } \beta(T) &= \text{dim of column space of } \beta(T)^T \\
&= \dim(T^*(W^*)) \\
&= \dim(T(V)) \\
&= \text{dim of column space of } \beta(T)
\end{aligned}$$

Coordinate Transforms

Let V be a vector space with bases for $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_n\}$

- For $v \in V$, α^v and β^v are the coordinate representations of v against A and B
- $\beta(A,B)$ is the $n \times n$ matrix with columns as the coordinate representations of the elements of A against the basis B.
- $\beta^v = \beta(A,B)\alpha^v$

Coordinate Transforms

Two matrices P and Q are *similar* if there is S so that $Q = SPS^{-1}$

- Let V have bases A and B and let $T : V \to V$ be linear.
- Let $M(T,A)$ and $M(T,B)$ be the matrix representations of T against A and B. Then

$$M(T,B) = \beta(A,B)M(T,A)\beta(A,B)^{-1}.$$

Coordinate Transforms

Let M be a $n \times n$ representing $T : \mathbb{R}^n \to \mathbb{R}^n$ against the canonical basis.

- Suppose M has n linearly independent eigenvectors $\Xi = \{\xi_1, \ldots, \xi_n\}$ and associated eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$.
- Let S be the matrix whose columns are the ξ_i
- Let Λ be the matrix representation of T against Ξ
- $M = S\Lambda S^{-1}$, so $\Lambda = S^{-1}MS$.
- Λ is a diagonal matrix, and the diagonal elements correspond to the eigenvalues of M.

Coordinate Transforms

The product $S\Lambda S^{-1}$ is an eigenvalue decomposition of M

- When an eigenvalue decomposition exists, the matrix is said to be diagonalizable, i.e. similar to a diagonal matrix

- Not all matrices are diagonalizable

Theorem 3.3 If $M \in \mathbb{R}^{n\times n}$ had n distinct eigenvalues then M is diagonalizable

Matrix Structure and Invariant Subspaces

We need 3 definitions:

- 1. A matrix M is upper triangular if $M_{ij} = 0$ whenever $i > j$
- 2. A complex matrix M is *unitary* if $M^* M = I_n = MM^*$, where M^* is the Hermitian (conjugate) transpose. A real matrix M is *orthogonal* if $M^T M = I_n = MM^T$
- 3. A matrix M is *block diagonal* if it can be written as as diagonal matrix of square matrices, i.e.

$$M = \begin{pmatrix} M_1 & 0 & 0 & \cdots & 0 \\ 0 & M_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & M_m \end{pmatrix} \equiv \bigoplus_{i=1}^{m} M_i \tag{1}$$

Decompositions

Definition Jordon block: a Jordan block, J, is as follows: for given λ in $\mathbb{C}$ or $\mathbb{R}$, and $n \in \mathbb{N}$, let

$$J(\lambda,n) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

If $\lambda = a + bi$ with $b \neq 0$, define

$$C(\lambda) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Decompositions

Theorem 3.4 Let $M \in \mathbb{R}^{n\times n}$ with eigenvalues $\{\lambda_1, \dots, \lambda_n\}$

- 1. M is similar to a Jordon block diagonal matrix, i.e. a matrix of the form

$$\bigoplus_{k=1}^{N} J(\lambda_{m_k}, n_k) = \begin{pmatrix} J(\lambda_{m_1}, n_1) & & & \\ & J(\lambda_{m_2}, n_2) & & \\ & & \ddots & \\ & & & J(\lambda_{m_N}, n_N) \end{pmatrix}.$$

    - Note that there may be multiple Jordon blocks associated to the same eigenvalue, which is the reason for the strange m_k subscript on the λs.

Decompositions

Theorem 3.4 (continued)

- 2. If λ is not real and $J(\lambda, m)$ is an associated Jordon block then there is necessarily a Jordon block of the form $J(\overline{\lambda}, m)$. Furthermore, the direct sum $J(\lambda, m) \oplus J(\overline{\lambda}, m)$ can be replaced with a $2m$ block of the form

$$J^{\mathbb{R}}(\lambda, 2m) = \begin{pmatrix} C(\lambda) & I_2 & & & \\ & C(\lambda) & I_2 & & \\ & & \ddots & \ddots & \\ & & & C(\lambda) & I_2 \\ & & & & C(\lambda) \end{pmatrix}.$$

Decompositions

Theorem 3.5 Let $M \in \mathbb{R}^{n\times n}$ with eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$

- 1. $M = QTQ^*$ where T is upper triangular with the eigenvalues of M on the diagonal, and Q is unitary.

- 2. $M = ZT^{\mathbb{R}}Z^T$ where $T^{\mathbb{R}}$ is upper block triangular and Z is orthogonal. The diagonal elements of $T^{\mathbb{R}}$ correspond to the eigenvalues of M. The real eigenvalues of M correspond to 1×1-blocks, and conjugate pairs of non-real eigenvalues of M $(\lambda, \bar{\lambda})$ correspond to 2×2 blocks of the form $C(\lambda)$.

Thus a Schur decomposition provides for a nested sequence of invariant subspaces together with a matrix representation that acts on these subspaces recursively.

Definiteness

Given a matrix $M \in \mathbb{R}^{n \times n}$, we may view M as a map from $\mathbb{R}^n \oplus \mathbb{R}^n$ to $\mathbb{R}$, by sending (v, w) to $w^T M v$

- for fixed w, the map $v \to w^T M v$ is a linear functional, and for fixed v, the map $w \to w^T M v$ is a linear functional

- M is a bilinear form

Definition Tensor: A multilinear form as a list with a multi-index

Definiteness

Some Definitions:

- A matrix M is *symmetric* if $M = M^T$.

- A matrix is *positive definite* if $v^T M v > 0$ for all non-zero $v \in \mathbb{R}^n$.

- A matrix is *positive semi-definite* if $v^T M v \geq 0$ for all non-zero $v \in \mathbb{R}^n$.

Theorem 3.4 If M is symmetric and positive semi-definite then the eigenvalues of M are real and non-negative.