Contents

1 Vectors

Interpretations of Vectors

- **Vector**: an ordered list of numbers.
- Possible notations: $\vec{v} = \boldsymbol{v}$ are most common.
- **Dimensionality**: the number of the elements in a vector.
- **Geometric vector**: an object with a magnitude and direction.
- Standard position: when the vector beings at the origin.
- Vectors must have same dimensionality for addition and subtraction.
- **Unit vector**: a vector with a norm (length) of 1. Notation: $\hat{u} = \dfrac{\boldsymbol{u}}{|\boldsymbol{u}|}$

Vector Multiplication

- **Scalar**: scales each element in a vector, does not change direction. Generally represented with greek letters.
- **Dot product**: a single number that provides information about the relationship between two vectors. Must have same dimensionality.
- Notation for dot product: $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{a}^T\boldsymbol{b} = \langle \boldsymbol{ab} \rangle = \sum a_i b_i$
- *Algebraic* dot product properties:
  - *Associative:* $\boldsymbol{a}^T(\boldsymbol{b}^T\boldsymbol{c}) \neq (\boldsymbol{a}^T\boldsymbol{b})^T\boldsymbol{c}$
  - *Distributive:* $\boldsymbol{a}^T(\boldsymbol{b}+\boldsymbol{c}) = \boldsymbol{a}^T\boldsymbol{b} + \boldsymbol{a}^T\boldsymbol{c}$
  - *Commutative:* $\boldsymbol{a}^T\boldsymbol{b} = \boldsymbol{b}^T\boldsymbol{a}$
  - Vector magnitude/length: $||\boldsymbol{v}|| = \sqrt{\boldsymbol{v}^T\boldsymbol{v}}$
- *Geometric* dot product properties:
  - Magnitudes of vectors scaled by angle between them. i.e. $\vec{a} = |a||b|\cos(\theta_{ab})$
  - Geometric and algebraic are really the same. The above equation can be rewritten as the algebraic vector length, i.e. $\boldsymbol{a}^T\boldsymbol{b} = \cos(\theta_{ab})|a||b|$
- Dot product features based on θ:
  - If $\cos(\theta) > 0$ then $\alpha > 0$
  - If $\cos(\theta) < 0$ then $\alpha < 0$
  - If $\cos(\theta) = 0$ then $\alpha = 0$; termed **Orthogonal**

  - If $\cos(\theta) = 1$ then $\alpha = |a||b|$

- *Hadamard vector multiplication*: elementwise multiplication of two vectors of equal dimensionality.
- **Outer product**: $\boldsymbol{v}\boldsymbol{w}^T = n \times m$; a matrix resulting from the product of vectors with dimensions n and m.
- **Cross product**: defined only between two 3D vectors; produces another 3D vector that is perpendicular to both original vectors, or normal, to the plane containing them.
- *Complex conjugate*: the inverse sign of imaginary componenet of a number.
- **Hermitian transpose**: or conjugate transpose, a transpose of a vector or matrix containing imaginary numbers using the complex conjugate.
- Notation for Hermitian transpose on a matrix: $\boldsymbol{M}^H$ or $\boldsymbol{M}^*$

Vector Space

- **Dimension**: in linear algebra dimension represnets a new element of infomation about a vector. Geometrically, each dimension represent a new direction.
- **Fields**: a set of numbers on which a set of arithmetic are valid.
- Field notations:
  - $\mathbb{R}$: real numbers.
  - $\mathbb{C}$: complex numbers.
  - $\mathbb{K}$: real complex numbers.
  - $\mathbb{Z}$: integers.
  - $\mathbb{N}$: poisite integers.
  - $\mathbb{Q}$: rational numbers.
  - Dimensionality in fields is written with superscripts, e.g. $\mathbb{R}^N$
- **Subspace**: The set of all vectors that can be created by a linear combination of some vectors or a set of vectors, i.e. $\lambda\boldsymbol{v},\ \lambda \in \mathbb{R}$
  - More formal definition: a vector subspace (V) must be closed under addition and scalar multiplication, and must contain the zero vector.
  - $\forall \boldsymbol{v}, \boldsymbol{w} \in V;\ \forall \lambda, \alpha \in \mathbb{R};\ \lambda\boldsymbol{v} + \alpha\boldsymbol{w} \in V$
  - Geometric: all possible scaled versions of vector produces a line for a single vector, and a 2d plane between two linear independent vectors.

- **Subset**: a set of points that satisfies some conditions; doesn't need to include toe origin, doesn't need to be closed, and can have boundaries.
- **Span**: all possible linear combinations of vectors within a subspace.
- **Linear Independence**: a property of an entire set of vectors that is true only if no vector in the set can be written as a linear combination of the others.
  - Geometric perspective: a set of M vectors is independent if each vector points in a geometric dimension not reachable using other vectors in the set.
  - Any set of $M > N$ vectors in $\mathbb{R}^N$ is dependent.
  - Any set of $M \leq N$ vectors in $\mathbb{R}^N$ *could be* independent.
- **Basis**: a combination of span and independence. More formally: if every element of a vector space V over a field is a linearly independent subset of V that spans V.

2 Introcution to Matrices

Terminology and Dimensionality

- Matrix notation:
  - **M** for reference to entire matrix.
  - m$_{M,N}$ represnets an individual element in a matrix.
  - **Block matrix**: a matrix that includes matrices itself; useful for matrices with higher level structure and can have computational benifits. e.g:

$$M = \begin{bmatrix} D & 0 \\ 1 & D \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 4 \end{bmatrix}$$

  - **Diagonal**: the diagonal elements of a matrix going form top left to bottom right.
    - *Off-diagonal*: elements not in the diagonal elements.
    - Works for nonsquare matrices.
  - Rows (M) first, then (N) columns when describing matrices.
  - **Dimensionality**: more open to interpretations:
    - $\mathbb{R}^{MN}$: simply the product, or total number of elements.
    - $\mathbb{R}^{M\times N}$: more explicit, not the product. $M \times N$ could be different from $N \times M$
    - $C(M) \in \mathbb{R}^M$: *column space* that is the collection of column vectors.
    - $C(M) \in \mathbb{R}^N$: *row space* that is the collection of row vectors.
    - Ambiguity opens up flexibility, but makes terminology more context dependent.
  - **Tensor**: higher dimensional cubes, won't be discussed much in early linear algebra.
- Types of common matrices:
  - **Square**: $M \times M$
    - **Symmetric**: a square matrix that is symmetric across the diagonal.
    - **Skew-symmetric**: diagonal must be zero, and numbers must mirror signs across diagonal.

- **Identity**: square matrix with 1s across the diagonal and zeros everywhere elses.
  - Notation: I_x.
  - x indicates size of matrix; no subscript means it is relevent size to matrix it is being applied to.
- **Rectangular**: non-square, $M \times N$
- **Zero**: matrix consisting of only zeros.
- **Diagonal**: all off diagonal elements are zero.
  - Identity matrix is a special case of diagonal.
- **Triangular**:
  - *Upper*: all elements below diagonal are zero.
  - *Lower*: all elements above diagonal are zero.
- **Concatenated**: two matrices with same number of rows concatenated. Often a line is placed in the product matrix to indicate concatenation point.

Basic Matrix Arithmetic

- Addition/subtraction is defined only for two matrices with same number of elements.
  - Simply a-dd/subtract corresponding elements.
  - Commutative: $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$
  - Associative: $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$
- **Shifting**:
  - Adding a scaled version of identity matrix.
  - Formula: $\mathbf{A} + \lambda I = \mathbf{C}$
  - Geometrically infats matrix, pushing it towards a sphere.
  - Tends to regularize the matrix.
- **Matrix-Scalar multiplication**:
  - Elementwise multiplication by the scalar.
  - Commutative: $\delta\mathbf{MA} = \mathbf{M}\delta\mathbf{A} = \mathbf{MA}\delta$
  - Distributive: $\delta(\mathbf{MA}) = \delta\mathbf{M} + \delta\mathbf{A}$
  - Distributive nature makes it a linear opperation.

- **Matrix transpose**:
  - $(\mathbf{A}^{M\times N})^T = \mathbf{A}^{N\times M}$
  - First column becomes first row, and vice versa.
  - $\mathbf{A}^{TT} = \mathbf{A}$
  - *Symmetric*: $\mathbf{A} = \mathbf{A}^T$
  - *Skew-symmetric*: $\mathbf{A} = \mathbf{-A}^T$
- **Complex matrices**: similar to vectors, a matrix containing at least one non-zero imaginary component.
  - Hermitian transpose only changes the sign (the conjugate) of the complex parts, not the real numbers.
- **Diagonal and trace**: extracts the diagonal as a vector.
  - Defined both for square and rectangular.
  - Not the same as *diagonalizing* a matrix.
  - **Trace**: the sum of the vector extracted from a matrix and is defined only for square matrices.
  - Diagonal formula: $\mathbf{v}_i = \mathbf{A}_{i,i} \;\; i = \{1, 2, ..., min(m, n)\}$
  - Trace formula: $tr(\mathbf{A}) = \sum \mathbf{A}_{i,i}$
- **Broadcasting**: expand, or repeat, a vector until traditional arithmetic is valid between a matrix and the vector.

3 Matrix Multiplications

Standard Matrix Multiplication

- Order of multiplication often matters and is generally not communitive.
    - i.e. $\mathbf{AB} \neq \mathbf{BA}$.
- Rules for validity:
    - Multiplication is valid only when the inner dimensions match.
    - Resulting matrix size corresponds to the outer dimensions.
    - i.e. $\mathbf{M} \times \mathbf{N} \cdot \mathbf{N} \times \mathbf{K} = \mathbf{M} \times \mathbf{K}$
- Methods of multiplication:
    - **Element perspective**: an order collection of dot products; each element is a dot product between rows of the left and columns of the right matrix.
    - **Layer perspective**: outer products summed between columns of the left matrix and rows of the right matrix.
        - or the sum of *rank 1 matrices*.
    - **Column perspective**: columns of the left matrix scaled then summed by the elements of the right matrix to make a column in the product matrix.
        - or a *linear weighted combination* of the columns in the left, weighted by the columns in the right.
    - **Rows perspective**: rows of the left matrix scales the rows of right matrix then summed to make a row in the product matrix.

Diagonal Matrix Multiplication

- $\mathbf{AD} \rightarrow$ Post-multiplying by a diagonal.
- $\mathbf{DA} \rightarrow$ Pre-multiplying by a diagonal.
- Post-multiplication results in the dense matrix's columns being weighted by the diagonal elements, while pre-multiplication scales the rows.

Order of Operations

- $(\mathbf{LIVE})^T = \mathbf{E}^T\mathbf{V}^T\mathbf{I}^T\mathbf{L}^T$
- Matrices multipied then transposed is equal to the same matrices transposed then multiplied in reverse direction.

Matrix-Vector Multiplication

- The result is always vector when multiplying by a matrix by a vector.
- The product vector always has the same orientation as the intput vector.
- The length of the vector is determined by the length of the matrix.
- $\mathbf{Aw} \rightarrow$ a weighted combination of the columns of $\mathbf{A}$.
- $\mathbf{w}^T\mathbf{A} \rightarrow$ a weighted combination of the rows of $\mathbf{A}$.
- Symmetric matrices have the same result.

Additive and Multiplicative Matrix Identities

- Pre or post multiplication of an identity matrix gives you the same matrix, i.e. $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$
  - This is the multiplicative indentity, which is often just called the identity matrix.
- Adding a zeros matrix is the additive identity, i.e. $\mathbf{A} + \mathbf{0} = \mathbf{A}$
- Methods of creating symmetric matrices from non-symmetric matrices:
  - Additive method: $\mathbf{S} = (\mathbf{A} + \mathbf{A}^T)/2$
    - Works only if $\mathbf{A}$ is a square matrix.
  - Multiplicative method:
    - $\mathbf{A}^T\mathbf{A} = \mathbf{S}$ that is $n \times n$
    - $\mathbf{AA}^T = \mathbf{S}$ that is $m \times m$
    - Always produces a symmetric square matrix.

Multiplication of Symmetric Matrices

- The product of symmetric matrices is not itself symmetric.
- A 2×2 matrix with constant diagonals will produce symmetric matrix.

Frobenius Dot Product

▹ There are three ways to calculate the Frobenius dot product:

1. Hadamard multiplation then sum all elements.
   - **Hadamard multiplation**: element wise multiplation; valid only with matrices of equal size.
     - A diagonal matrix multiplied with its self has the same product using Hadamard or standard matrix multiplation.
2. Vectorize both matrices then compute the vector dot product.
   - **Vectorizating a matrix**: creating a vector of matrix by concatenating column wise of a matrix.
3. Transpose and trace of the product.
   - Notation: $\langle \mathbf{A}, \mathbf{B} \rangle_F = tr(\mathbf{A}^T\mathbf{B})$
   - **Trace**: the sum of the elements in the main diagonal.
   - The most efficient method of calculating the Frobenius dot product.