1 Mathematics Refresher for Machine Learning

1.1 Sets

A set is a well-defined collection of distinct objects (possibly infinite or uncountable). E.g:

- $\{1, 2, 3\}, \{a, e, i, o, u\}, \{\pi, e\}$
- Integers $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
- Positive Integers $\mathbb{Z}_{++} = \{1, 2, 3, \ldots\}$
- Real Numbers $\mathbb{R}$

If x is an elements of set Z we write x. E.g: $x \in \mathbb{R}$ means x is a real number. Set builder notation:

- Positive reals $\mathbb{R}_{++} = \{x \in \mathbb{R} : x > 0\}$

1.2 Sets: empty set, cardinality, intersection, union

The empty set is the set with no elements = $\{\}$ The cardinality of a set is the number of elements in the set. E.g: $X = \{1, 2, 3\}, |X| = \#X = 3$ The intersection of two sets is the set containing all common elements. If $A = \{1, 2, 3\}$ and $B = \{3, 4, 5\}$ then the intersection $A \cap B = \{3\}$

$$\mathbb{Z} \cap \mathbb{R} = \mathbb{Z}.$$

The union of two sets is the set containing all elements that occur in either set. If $A = \{1, 2, 3\}$ and $B = \{3, 4, 5\}$, then the union $A \cup B = \{1, 2, 3, 4, 5\}$

$$\mathbb{Z} \cup \mathbb{R} = \mathbb{R}.$$

1.3 Sets: subsets

A is a subset of B if all the elements of A are also contained in B. Written as $A \subset B$

$$\{1, 2\} \subset \{1, 2, 3\}.$$

$$\mathbb{Z} \subset \mathbb{R}.$$

$$\mathbb{Z}_{++} \subset \mathbb{Z}_{+} \subset \mathbb{Z} \subset \mathbb{R}.$$

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}.$$

1.4 Vectors

Vectors, in general, an abstract mathematical notation, but for the purpose of this module can be thought of as an ordered list of numbers E.g: Column vectors:

$$x = \begin{pmatrix} 1 \\ 3 \end{pmatrix} y = \begin{pmatrix} 3 \\ 1 \\ 8 \end{pmatrix}.$$

To say that a vector x is real values with D dimensions, we write $x \in \mathbb{R}^D$ E.g: $x \in \mathbb{R}^2, y \in \mathbb{R}^3$ Can write column vectors more compactly using parentheses $x = (13)$ The elements of a vector are usually denoted using subscripts E.g: if $x = (145)$ then $x_1 = 1, x_2 = 4, x + 3 = 5$ The transpose of a column vector is a row vector

$$x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} x^T = [123].$$

This gives us another way to write column vectors compactly: $x = [x_1 x_2 x_3]^T \in \mathbb{R}^3$

1.5 Adding and Scaling Vectors

To add two row or column vectors of the same dimension, just add their components

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_D \end{pmatrix} y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_D \end{pmatrix} x + y = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_D + y_D \end{pmatrix}.$$

You cannot add a row vector to a column vector or vice versa (unless the dimension is 1) You cannot add vectors of different dimensions. To multiply a vector $x \in \mathbb{R}^D$ by a scalar $\alpha \in \mathbb{R}$, just multiply each component.

$$\alpha x = [\alpha x_1 \alpha x_2 \ldots \alpha x_D]^T.$$

1.6 Dot product

The dot product between two vectors $x, y \in \mathbb{R}^D$ is computed by multiplying the corresponding components of x and y and adding up the products.

$$x \cdot y = x_1 y_1 + x_2 y_2 + \ldots = \sum_{i=1}^{N} x_i y_i.$$

The dot product is also called the inner product or scalar product. Alternative notations:

- $x^T y$ (dot product as a a matrix multiplication)
- (x, y) or $(x|y)$ (bracket notation, common in physics)
- xy implicit notation

1.7 Norms

Most common norm is the Euclidean norm or L_2 norm, denoted $|| \ x \ ||_2$ or just $|| \ x \ ||$. Gives the length of a vector.

$$|| \ x \ || = \sqrt{x^T x} = \sqrt{\sum_i x_i^2}.$$

The squared Euclidean norm $|| x ||^2$ is often useful:

$$|| x ||^2 = x^T x = x \cdot x.$$

Another common norm is the L_1 norm, which is the sum of absolute values of x

$$|| x ||_1 = \sum_i | x_i | .$$

1.8 Properties of Norms

A norm is any function p from vectors to $\mathbb{R}$ that satisfies:

1. $p(\alpha x) = | \alpha | p(x)$ for $\alpha \in \mathbb{R}$
2. $p(x + y) \leq p(x) + p(y)$ (Triangle inequality)
3. $p(x) = 0 \iff x = 0$

We also have $p(-x) = p(x)$ E.g: $|| 5x || = 5 || x ||$ for any norm $|| \cdot ||$

1.9 Norms and Distance Metrics

Any vector space V and norm p can be used to induce a distance metric (define a metric space) by defining the distance metric to be $d(x, y) = p(x, y)$ E.g: The space of D dimensional real valued vectors $\mathbb{R}^D$ and the L_2 norm give the Euclidean space with metric $d(x, y) = || x - y ||_2$ In 2D this gives the familiar Euclidean distance between two points x and y:

$$d(x, y) = || x - y ||_2 = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

1.10 Matrices