12.1.1 Matrices and Vectors

Definition of **Matrix**. An $M \times N$ *matrix* $\mathbf{A}$ is a two-dimensional array of numbers

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \dots & a_{MN} \end{bmatrix}$$

A matrix can also be written as $\mathbf{A} = (a_{nm})$ where $n = 1, 2, \ldots, N$ and $m = 1, 2, \ldots, M$. Matrices are usually written as boldface, upper-case letters.

Definition of **Matrix Transpose**. The transpose of $\mathbf{A}$, denoted by $\mathbf{A}^t$, is the NxM matrix (a_{mn}) or

$$\mathbf{A}^t = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1M} \\ a_{21} & a_{22} & \dots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NM} \end{bmatrix}$$

Definition of **Vectors**.. A *column vector* is a $Bx1$ matrix and a *row vector* is a $1xB$ matrix, where $B \geq 1$. Column and row vectors are usually simply referred to as *vectors* and they are assumed to be column vectors unless they are explicitly identified as row vector or if it is clear from the context. Vectors are denoted by boldface, lower-case letters.

$$\mathbf{x} = [x_1, x_2, \ldots, x_B]^t = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_B \end{bmatrix}$$

Definition of **Dimension**. The integer B is called the *dimension* of the vector $\mathbf{x}$ in the definition above. The phrase $\mathbf{x}$ is $B - dimensional$ is equivalent to stating that the dimension of $\mathbf{x}$ is B.

Definition of **0 and 1 vectors**.. The vectors

$$\mathbf{0} = (0, 0, \ldots, 0)^t \text{ and } \mathbf{1} = (1, 1, \ldots, 1)^t$$

are called the *Origin* or *Zero Vector* and the *One Vector*, respectively.

Definition of **Vector Addition and Scalar Multiplication**.. The addition of two vectors $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^B$ is defined as: $\mathbf{x}_3 = [x_{11} + x_{21}, x_{12} + x_{22}, \ldots, x_{1B} + x_{2B}]^t$. The word "scalars" is another word for numbers. Scalar multiplication is the product of a number α and a vector $\mathbf{x}$, defined by $\alpha\mathbf{x} = [\alpha x_1, \alpha x_2, \ldots, \alpha x_B]^t$.

Definition of **Matrix Multiplication**. If $\mathbf{A}$ and $\mathbf{B}$ are $M \times N$ and $N \times P$ matrices, then the *matrix product*, $\mathbf{C} = \mathbf{AB}$, is the $M \times P$ matrix defined by

$$c_{mp} = \sum_{n=1}^{N} a_{mn} b_{np}$$

Definition of **Dot, or Inner, Product**. If $\mathbf{x}$ and $\mathbf{y}$ are B-dimensional vectors, then the matrix product $\mathbf{x}^t\mathbf{y}$ is referred to as the *Dot* or *Inner Product* of $\mathbf{x}$ and $\mathbf{y}$.

Note that a linear system of equations can be represented as a matrix multiplication. For example, the system

$$a_{11}x_1 + a_{12}x_2 = b_1$$
$$a_{21}x_1 + a_{22}x_2 = b_2$$

and be written in matrix-vector form as

$$\mathbf{Ax} = \mathbf{b}.$$

Definition of **Diagonal Matrix**. A diagonal matrix $\mathbf{D} = (d_{mn})$ is an NxN matrix with the property that $d_{mn} = 0$ if $m \neq n$.

Definition of **Identity Matrix**. The NxN *Identity Matrix*, denoted by $\mathbf{I}_N$ or just $\mathbf{I}$ if N is known, is the NxN diagonal matrix with $d_{nn} = 1$ for $n = 1, 2, \ldots, N$. Notice that if $\mathbf{A}$ is any NxN matrix, then $\mathbf{AI}_N$ = $\mathbf{I}_N\mathbf{A}$ = $\mathbf{A}$.

Definition of **Inverse Matrix**. Let $\mathbf{A}$ be an NxN square matrix. If there is a matrix $\mathbf{B}$ with the property that $\mathbf{AB}$ = $\mathbf{BA}$ = $\mathbf{I}_N$, then $\mathbf{B}$ is called *the inverse of* $\mathbf{A}$ or $\mathbf{A}$ *inverse*. It is denoted by $\mathbf{A}^{-1}$. The inverse of a square matrix $\mathbf{A}$ does not always exist. If it does exist, then $\mathbf{A}$ is *invertible*.

12.1.2 Vector Spaces

Definition of **Vector, or Linear, Space**. Let B denote a positive integer. A finite-dimensional *Vector, or Linear, Space* with *dimension* B is a collection of $B - dimensional$ vectors, V that satisfies commutative, associative, and distributive laws and with the properties that:

1. For every $\mathbf{x} \in V$ and $\mathbf{y} \in V$, the sum $(\mathbf{x} + \mathbf{y}) \in V$.
2. For every real number s and $\mathbf{x} \in V$, the product $s\mathbf{x} \in V$.
3. For every $\mathbf{x} \in V$, the additive inverse $-\mathbf{x} \in V$, which implies that $\mathbf{0} \in V$.

Example. Graphical Vector Operations. Two- and Three-dimensional vectors can be added graphically using the ”tip-to-tail” method. An example is shown in Figure-GraphVecAdd .

Notation. A B-dimensional vector space is often denoted by $\mathbb{R}^B$.

Definition of **Subspace**. A *Subspace* is a subset of a vector space that is also a vector space. Notice that subspaces of vector spaces always include the origin.

Example. The vectors on the line $\mathcal{L}_s$ = $\{(x_1, x_2)^t \, | x_2 = 2x_1\}$ is a subspace of $\mathbb{R}^2$ of dimension 1. The vectors on the line $\mathcal{L}_n = \{(x_1, x_2)^t \, | x_2 = 2x_1 + 2$ is not a subspace.

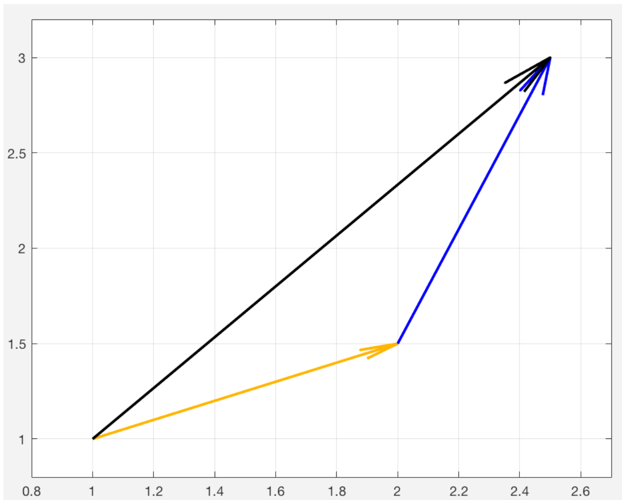

Figure 12.1 Vector Addition: $\mathbf{x}_3 = \mathbf{x}_1 + \mathbf{x}_2$. $\mathbf{x}_3$ is shown in **Black**, $\mathbf{x}_1$ is shown in **Orange** and $\mathbf{x}_2$ is shown in **Blue**.

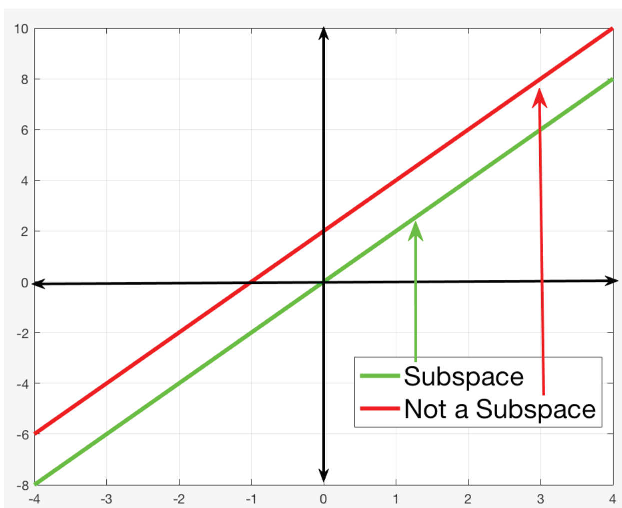

Figure 12.2 Subspace and non-Subspace

Subspaces are often used in algorithms for analyzing image spectrometer data. The motivation is that spectra with certain characteristics might all be contained in a subspace.

Definition of **Linear Combination** . If $\mathbf{y}$, $\{\mathbf{x}_d\}_{d=1}^{D} \in \mathbb{R}^B$ and $\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_D\mathbf{x}_D$ then we say that $\mathbf{y}$ is a *linear combination* of $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_D\}$. The numbers $a_1, a_2, \ldots, a_D$ are called the *coefficients* of the linear combination. Notice that the linear combination can be written in matrix-vector form as:

$$\mathbf{y} = \mathbf{X}\mathbf{a}$$

where $\mathbf{X}$ is the $B \times D$ matrix with $\mathbf{x}_d$ as the d^{th} column and $\mathbf{a}$ is the vector of coefficients.

Definition of **Linear Independence**. If $\mathbb{B}$=$\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_D\}$ is a set of vectors with the property that

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_D\mathbf{x}_D = 0 \implies a_1 = a_2 = \ldots = a_D = 0$$

then the vectors in the set $\mathcal{B}$ are called *Linearly Independent*. Informally, no vector in a set of linearly independent vectors can be written as a linear combination of the other vectors in the set.

Fact.There can be no more than B linearly independent vectors in a $B - dimensional$ vector space. If $\mathcal{B}$ = $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_B\}$ is a set of B linearly independent vectors in a $B - dimensional$ space $\mathcal{V}$, then every $\mathbf{x} \in \mathcal{V}$ can be written uniquely as a linear combination of the vectors in $\mathcal{B}$.

Definition of **Basis**. A *Basis* of a subspace of B-dimensional space, $\mathcal{S}$, is a set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_D\}$ with the property that every vector $\mathbf{x} \in \mathcal{S}$ can be written one and exactly one way as a linear combination of the elements of the basis. In other words, if $\mathbf{x} \in \mathcal{S}$ then there is a unique set of coefficients, $a_1, a_2, \ldots, a_D$ such that $\mathbf{y} = a_1\mathbf{v}_1 + a_2 + \mathbf{v}_2, \cdots + a_D\mathbf{v}_D$. It must be the case that $D \leq B$.

Fact.There are infinitely many bases for any subspace but they all have the same number of elements. The number of elements is called the dimension of the subspace.

Definition of **Standard Basis**.. The standard basis of a D-dimensional subspace is the set of vectors $\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_D\} \subset \mathbb{R}^D$ where the d^{th} element of the d^{th} vector $\mathbf{s}_d$, s_{dd}, is equal to one and the rest of the elements are equal to zero. That is

$$s_{dj} = 1 \text{ if } d = j \text{ and } s_{dj} = 0 \text{ if } d \neq j.$$

Example. . The standard basis for $\mathbb{R}^3$ is

$$\mathbf{s}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \mathbf{s}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } \mathbf{s}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Any 3-Dimensional vector can be represented as a linear combination of the standard basis vectors:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\mathbf{s}_1 + x_2\mathbf{s}_2 + x_3\mathbf{s}_3 = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Fact.If the columns of an N×N matrix $\mathbf{A}$ form a basis for $\mathbb{R}^N$, then A is invertible.

It can be useful to represent vectors in terms of different bases.

Definition of **Change of Basis**. The process of changing from a representation in terms of basis to a representation in terms of a different basis is called a *Change of Basis* transformation.

Calculating a change of basis transformation may require solving a linear system. To see this, let $\mathcal{U}$ = $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_B\}$ and $\mathcal{V}$ = $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_B\}$ be two bases for $\mathbb{R}^B$. Suppose the representation of a vector $\mathbf{x}$ in terms of the basis $\mathcal{U}$ is known to be

$$\mathbf{x} = a_1\mathbf{u}_1 + a_2\mathbf{u}_2 + \cdots + a_B\mathbf{u}_B = \mathbf{Ua}$$

and that the representation of $\mathbf{x}$ in terms of $\mathcal{V}$ is unknown

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_B\mathbf{v}_B = \mathbf{Vc}.$$

More precisely, $\mathbf{a}$ is known but $\mathbf{c}$ is unknown. Then $\mathbf{c} = \mathbf{V}^{-1}\mathbf{Ua}$. In particular, if $\mathcal{U}$ is the standard basis, then $\mathbf{U} = \mathbf{I}$ and $\mathbf{x} = [a_1, a_2, \ldots, a_B]^t$, so $\mathbf{c} = \mathbf{V}^{-1}\mathbf{x}$. For these cases, the matrices $W = \mathbf{V}^{-1}\mathbf{U}$ and $\mathbf{V}^{-1}$ are called the *change of basis matrices*.

12.1.3 Norms, Metrics, and Dissimilarities

Definition of **norm**. A norm produces a quantity computed from vectors that somehow measures the size of the vector. More formally, a norm is a function, $\mathcal{N} : \mathbb{R}^B \to \mathbb{R}_+$ with the properties that $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^B$ and $a \in \mathbb{R}$:

1. $\mathcal{N}(\mathbf{x}) \geq 0$ and $\mathcal{N}(\mathbf{x}) = 0$ if and only if $\mathbf{x} = 0$.
2. $\mathcal{N}(a\mathbf{x}) = a\mathcal{N}(\mathbf{x})$.
3. $\mathcal{N}(\mathbf{x} + \mathbf{y}) \geq \mathcal{N}(\mathbf{x}) + \mathcal{N}(\mathbf{y})$

Notation. The norm is usually written with the vertical bar notation: $\mathcal{N}(\mathbf{x}) = \|\mathbf{x}\|$.

Definition of $\mathcal{L}_p$ **norm, or just** p**-norm**.. The *p-norm* is defined by $\|\mathbf{x}\|_p = \left(\sum_{b=1}^{B} |x_b|^p\right)^{\frac{1}{p}}$.

Definition of **Euclidean Norm**.. The *Euclidean Norm* is the p-norm with $p = 2$. If p is not specified, then it is assumed that $\|\mathbf{x}\|$ denotes the Euclidean norm unless explicitly indicated otherwise. Note that the Euclidean norm can be written as $\|\mathbf{x}\| = \sqrt{\mathbf{x}^t\mathbf{y}}$.

Definition of ∞**-norm**.. The ∞*-norm* is $\|\mathbf{x}\|_\infty = \max_{b=1}^{B}(x_b)$.

Definition of **Unit norm**. A vector $\mathbf{x}$ is said to have *unit norm* if $\|\mathbf{x}\| = 1$. Note that if $\mathbf{x}$ is any vector then $\frac{\mathbf{x}}{\|\mathbf{x}\|}$ has unit norm.

Fact.If $\mathbf{x}$ and $\mathbf{y}$ are vectors, then $\mathbf{x}^t\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos\theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$. If $\mathbf{x}$ and $\mathbf{y}$ are normalized, then the inner product of the vectors is equal to the cosine of the angle between the vectors

$$\frac{\mathbf{x}^t\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \left(\frac{\mathbf{x}}{\|\mathbf{x}\|}\right)^t \left(\frac{\mathbf{y}}{\|\mathbf{y}\|}\right) = \cos\theta.$$

Definition of **Distance Function or Metric**. A *distance* or *metric* is a function defined on pairs of vectors,

$$d : \mathbb{R}^B \text{x} \mathbb{R}^B \to \mathbb{R}_+$$

and has the following properties:

1. $d(\mathbf{x}, \mathbf{y}) \geq 0$ and $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$
2. $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$

3. For any $\mathbf{z}$, $d(\mathbf{x},\mathbf{y}) \le d(\mathbf{x},\mathbf{z}) + d(\mathbf{z},\mathbf{y})$

Definition of $\mathcal{L}_p$ **distance**. $d_p(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_p$.

Note that the Euclidean distance squared can be written as

$$d_2(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|^2 = (\mathbf{x}-\mathbf{y})^t(\mathbf{x}-\mathbf{y}) = \|\mathbf{x}\| + \|\mathbf{y}\| - 2\mathbf{x}^t\mathbf{y}.$$

Therefore, if $\mathbf{x}$ and $\mathbf{y}$ are unit vectors, then measuring Euclidean distance between vectors is equivalent to measuring the cosine of the angle between the two vectors:

$$d_2(\mathbf{x},\mathbf{y})^2 = \|\mathbf{x}-\mathbf{y}\|^2 = (\mathbf{x}-\mathbf{y})^t(\mathbf{x}-\mathbf{y}) = 1 + 1 - 2\mathbf{x}^t\mathbf{y} = 2 - 2\cos\theta.$$

Therefore, $\cos\theta = 1 - 0.5\|\mathbf{x}-\mathbf{y}\|^2$.

Definition of **Orthogonality and Orthonormal Bases**. Two non-zero vectors $\mathbf{x}$ and $\mathbf{y}$ are called *orthogonal* if $\mathbf{x}^t\mathbf{y} = 0$. Note that orthogonal is another word for perpendicular since (for non-zero vectors) $\mathbf{x}^t\mathbf{y} = \|\mathbf{x}\|\|\mathbf{y}\|\cos\theta = 0$ only happens if $\cos\theta = 0$. An *orthonormal basis* is a basis $\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_B\}$ with the property that $\mathbf{u}_i\mathbf{u}_j = \delta_{ij}$ where δ_{ij} is the Kronecker delta function.

Definition of **Projections**. The projection of a vector $\mathbf{x}$ onto a vector $\mathbf{y}$ is given by

$$Proj(\mathbf{x},\mathbf{y}) = \mathbf{x}^t\frac{\mathbf{y}}{\|\mathbf{y}\|}.$$

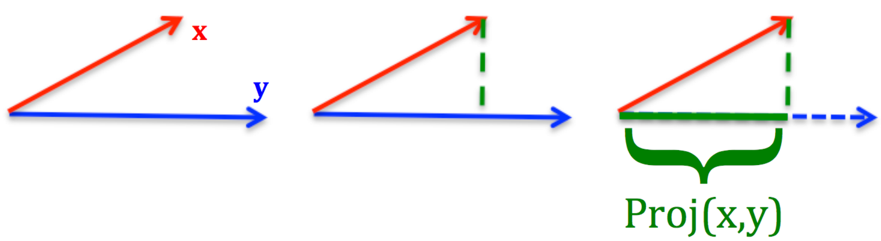

Figure 12.3 Projection of a vector $\mathbf{x}$ onto a vector $\mathbf{y}$