Applied Linear Algebra in Data Analysis

Vector spaces: A set of vectors that is closed under vector scaling and vector addition. E.g. $\mathbb{R}^n$, $\mathbb{C}^n$. *A vector space will always contain the zero vector.*

Subspace: A subset of a vector space $\mathcal{V}$ which is also a vector space.

Span of a set (of vectors): The set of all linear combinations of a set of vectors $\mathcal{S} = \{\mathbf{s}_i\}_{i=1}^p$ from the vector space $\mathcal{V}$.

$$\text{span}\,\mathcal{S} = \left\{\sum_{i=1}^{p} \alpha_i \mathbf{s}_i \;\middle|\; \alpha_i \in \mathbb{R}\right\} \subseteq \mathcal{V}$$

Linear independence: A set $\mathcal{S}$ is linearly independent if and only if, $\sum_{i=1}^p \alpha_i \mathbf{s}_i = \mathbf{0} \implies \alpha_i = 0, \forall i$. *If the set has* $\mathbf{0}$*, then the set is linearly dependent.*

Basis: A set of vectors $\mathcal{B}$ is a basis for a vector space $\mathcal{V}$ if and only if, $\mathcal{B}$ is linearly independent and span $\mathcal{B} = \mathcal{V}$. The elements of $\mathcal{B}$ are called basis vectors of $\mathcal{V}$. *There are infinitely many bases for a vector space. Every vector in* $\mathcal{V}$ *can be written as a* ***unique*** *linear combination of the basis vectors.*

Dimension: The number of basis vectors in a basis of a vector space $\mathcal{V}$ is called the dimension of $\mathcal{V}$.

Inner product: $\mathbf{x}^\top \mathbf{y} = \sum_{i=1}^n x_i y_i,\ \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Norm: Measure of the length of a vector. $\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}},\ \mathbf{x} \in \mathbb{R}^n$. $\|\mathbf{x}\|_2^2 = \mathbf{x}^\top\mathbf{x}$.

Cauchy-Schwarz inequality: $|\mathbf{x}^\top\mathbf{y}| \leq \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$.

Orthogonality: Two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are orthogonal if and only if, $\mathbf{x}^\top\mathbf{y} = 0$.

Orthonormal basis: A basis $\mathcal{B} = \{\mathbf{b}_i\}_{i=1}^n$ is orthonormal if and only if, $\mathbf{b}_i^\top \mathbf{b}_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta function.

Linear function: A function $f : \mathbb{R}^n \to \mathbb{R}$ that satisfies superposition. All linear functions f can be represented as $f(\mathbf{x}) = \mathbf{w}^\top\mathbf{x}$, where $\mathbf{w} \in \mathbb{R}^n$.

Matrix post-multiplication by a column vector: $\mathbf{A} \in \mathbb{R}^{n\times m}, \mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} = \mathbf{A}\mathbf{b} = \sum_{i=1}^m b_i \mathbf{a}_i$.

Matrix pre-multiplication by a row vector: $\mathbf{A} \in \mathbb{R}^{n\times m}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{c} = \mathbf{b}^\top\mathbf{A} = \sum_{i=1}^n b_i \tilde{\mathbf{a}}_i^\top$.

Matrix multiplication: $\mathbf{C} = \mathbf{A}\mathbf{B}$, $\mathbf{C} \in \mathbb{R}^{n\times m}$, $\mathbf{A} \in \mathbb{R}^{n\times p}$, $\mathbf{B} \in \mathbb{R}^{p\times m}$. Four views of matrix multiplication:

- *Inner product view*: $c_{ij} = \tilde{\mathbf{a}}_i^\top \mathbf{b}_j$
- *Column view*: $\mathbf{c}_i = \mathbf{A}\mathbf{b}_i$
- *Row view*: $\tilde{\mathbf{c}}_i^\top = \tilde{\mathbf{a}}_i^\top \mathbf{B}$
- *Outer product view*: $\mathbf{C} = \sum_{i=1}^p \mathbf{a}_i \tilde{\mathbf{b}}_i^\top$

Outer product: $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\mathbf{A} = \mathbf{x}\mathbf{y}^\top \in R^{n\times n}$. *Columns of* $\mathbf{A}$ *are scaled* $\mathbf{x}$*, and rows of* $\mathbf{A}$ *are scaled* $\mathbf{y}^\top$*.*

Rank of a matrix: The rank of a matrix $\mathbf{A} \in \mathbb{R}^{n\times m}$ is the number of linearly independent columns or rows of $\mathbf{A}$. rank $\mathbf{A} = \min(n, m)$.

Matrix inverse: When $\mathbf{A}$ is sqyare matrix, and is full rank, $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ and $\mathbf{A}^{-1}$ is unique.

Solutions to $\mathbf{A}\mathbf{x} = \mathbf{b}$: $\mathbf{A} \in \mathbb{R}^{n\times m}$.

- $\mathbf{b} \notin$ span $\mathbf{A} \implies$ No solution
- $\mathbf{b} \in$ span $\mathbf{A}$ and rank $\mathbf{A} = m \implies$ Unique solution
- $\mathbf{b} \in$ span $\mathbf{A}$ and rank $\mathbf{A} < m \implies$ Infinite solutions

Four fundamental subspaces of a matrix: $\mathbf{A} \in \mathbb{R}^{n\times m}$ and rank $\mathbf{A} = r$.

- *Column space*: $\mathcal{C}(\mathbf{A}) = \text{span}\ \{\mathbf{a}_i\}_{i=1}^m \subseteq \mathbb{R}^n$, $\dim \mathcal{C}(\mathbf{A}) = r$.
- *Row space*: $\mathcal{C}(\mathbf{A}^\top) = \text{span}\ \left\{\tilde{\mathbf{a}}_i^\top\right\}_{i=1}^n \subseteq \mathbb{R}^m$, $\dim \mathcal{C}(\mathbf{A}^\top) = r$.
- *Null space*: $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^m \mid \mathbf{A}\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^m$, $\dim \mathcal{N}(\mathbf{A}) = m - r$.
- *Left null space*: $\mathcal{N}(\mathbf{A}^\top) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}^\top\mathbf{x} = \mathbf{0}\} \subseteq \mathbb{R}^n$, $\dim \mathcal{N}(\mathbf{A}^\top) = n - r$.

Orthogonal subspaces: Two subspace $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^n$ are orthogonal if and only if, $\mathbf{u}^\top\mathbf{v} = 0, \forall \mathbf{u} \in \mathcal{U}, \forall \mathbf{v} \in \mathcal{V}$. $\mathcal{U} \perp \mathcal{V}$.

Sum of two subspaces: $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^n$. $\mathcal{U} + \mathcal{V} = \{\mathbf{u} + \mathbf{v} \mid \mathbf{u} \in \mathcal{U}, \mathbf{v} \in \mathcal{V}\} \subseteq \mathbb{R}^n$. $\mathcal{U} + \mathcal{V}$ is a subspace of $\mathbb{R}^n$.

Complementary subspaces: $\mathcal{U}, \mathcal{V} \subseteq \mathbb{R}^n$. $\mathcal{U} \cap \mathcal{V} = \{\mathbf{0}\}$ and $\mathcal{U} + \mathcal{V} = \mathbb{R}^n$.

Orthogonal complements: $\mathcal{U}, \mathcal{V}$ are complementary subspace of $\mathbb{R}^n$. If $\mathcal{U} \perp \mathbf{V}$, then $\mathcal{U}, \mathcal{V}$ are orthogonal complements. $\mathcal{U}^\perp = \mathcal{V}$ and $\mathcal{V}^\perp = \mathcal{U}$.

Orthogonal projection onto a subspace $\mathcal{S}$: If $\{\mathbf{u}_i\}_{i=1}^m$ is an orthonormal basis for $\mathcal{S}$ with $\mathbf{u}_i \in \mathbb{R}^n$, then the orthogonal projection of $\mathbf{x}$ onto $\mathcal{S}$ is $\mathbf{P}_{\mathcal{S}}\mathbf{x} = \left(\sum_{i=1}^m \mathbf{u}_i\mathbf{u}_i^\top\right)\mathbf{x} = \sum_{i=1}^m \left(\mathbf{u}_i^\top\mathbf{x}\right)\mathbf{u}_i$.

Components of a vector: Let $\mathcal{U}, \mathcal{V}$ be complementary subspaces of $\mathbb{R}^n$. Then $\mathbf{x} \in \mathbb{R}^n$ can be uniquely expressed as, $\mathbf{x} = \mathbf{x}_{\mathcal{U}} + \mathbf{x}_{\mathcal{V}}$, where $\mathbf{x}_{\mathcal{U}} \in \mathcal{U}$ and $\mathbf{x}_{\mathcal{V}} \in \mathcal{V}$. If $\mathcal{U}^\perp = \mathcal{V}$, then $\mathbf{x}_{\mathcal{U}}^\top\mathbf{x}_{\mathcal{V}} = 0$.

Gram-Schmidt orthogonalization: Let $\{\mathbf{a}_i\}_{i=1}^m$ be a set of linearly independent vectors in $\mathbb{R}^n$. Then, $\{\mathbf{u}_i\}_{i=1}^m$ is an orthonormal basis for span $\{\mathbf{a}_i\}_{i=1}^m$, where $\mathbf{u}_i = \frac{\mathbf{a}_i - \sum_{j=1}^{i-1}\left(\mathbf{u}_j^\top\mathbf{a}_i\right)\mathbf{u}_j}{\left\|\mathbf{a}_i - \sum_{j=1}^{i-1}\left(\mathbf{u}_j^\top\mathbf{a}_i\right)\mathbf{u}_j\right\|_2}$.

QR factorization: Let $\mathbf{A} \in \mathbb{R}^{n\times m}$ with rank $\mathbf{A} = m$. Then, $\mathbf{A} = \mathbf{Q}\mathbf{R}$, where $\mathbf{Q} \in \mathbb{R}^{n\times m}$ is an orthogonal matrix and $\mathbf{R} \in \mathbb{R}^{m\times m}$ is an upper triangular matrix.