Linear Algebra Review

Leila Wehbe

January 29, 2013

Table of contents

Metric

Given a space $\mathcal{X}$, then $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}_0^+$ is a metric is for all $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$ in $\mathcal{X}$ if:

- $d(\mathbf{x}, \mathbf{y}) = 0$ is equivalent to $\mathbf{x} = \mathbf{y}$
- $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$
- $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$

Example of a metric

Euclidean Distance:

Given $\mathcal{X} = \mathbb{R}^n$, $d(\mathbf{x}, \mathbf{y}) := (\sum_{i=1}^{n} (x_i - y_i)^2)^{\frac{1}{2}}$

- $d(\mathbf{a}, \mathbf{b}) = 0$ is equivalent to $\mathbf{a} = \mathbf{b}$
- $d(\mathbf{a}, \mathbf{b}) = d(\mathbf{b}, \mathbf{a})$
- $d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{c}) + d(\mathbf{c}, \mathbf{b})$ (this is the triangle inequality)

Vector Space

A vector space is a space $\mathcal{X}$ such that for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ and for all $\alpha \in \mathbb{R}$:

- $\mathbf{x} + \mathbf{y} \in \mathcal{X}$
- $\alpha \mathbf{x} \in \mathcal{X}$

Examples of vector spaces

Real Numbers: given $x, y \in \mathbb{R}$, and $\alpha \in \mathbb{R}$:

- $x + y \in \mathbb{R}$
- $\alpha x \in \mathbb{R}$

$\mathbb{R}^{\mathbf{n}}$: given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and $\alpha \in \mathbb{R}$:

- $\mathbf{x} + \mathbf{y} \in \mathbb{R}^n$
- $\alpha \mathbf{x} \in \mathbb{R}^n$

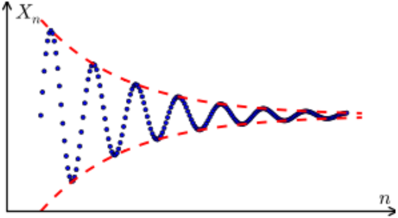
Examples of vector spaces

Polynomials: given $f(x) = \sum_{i=0}^{n} a_i x^i$ and $g(x) = \sum_{i=0}^{n} b_i x^i$, and $\alpha \in \mathbb{R}$:

- $f(x) + g(x) = \sum_{i=0}^{n} (a_i + b_i) x^i$, i.e. polynomial of order n
- $\alpha f(x) = \sum_{i=0}^{n} \alpha a_i \ x^i$, i.e. polynomial of order n

Cauchy Series

Given a space $\mathcal{X}$, a Cauchy series is a series $x_i \in \mathcal{X}$ for which for every $\epsilon > 0$ there exist an n_0 such that for all $m, n \geq n_0$, $d(\mathbf{x}_m, \mathbf{x}_n) \leq \epsilon$



Completeness

A space $\mathcal{X}$ is complete if the limit of every Cauchy series $\in \mathcal{X}$.

For example, $(0, 1)$ is not complete but $[0, 1]$ is.

The set $\mathbb{Q}$ of rational numbers is not complete: you can construct a sequence that converges to $\sqrt{2}$ but $\sqrt{2}$ is not in $\mathbb{Q}$.

Norm

Given a vector space $\mathcal{X}$, a norm is a mapping $||.|| : \mathcal{X} \to \mathbb{R}_0^+$ that satisfies, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ and for all $\alpha \in \mathbb{R}$:

- $||\mathbf{x}|| = 0$ if and only if $\mathbf{x} = 0$
- $||\alpha \mathbf{x}|| = |\alpha| ||\mathbf{x}||$
- $||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$ (triangle inequality)

A norm is also a metric: $d(\mathbf{x}, \mathbf{y}) := ||x - y||$

Banach Space

A Banach Space is a complete vector space $\mathcal{X}$ together with a norm $||.||$.

ℓ_p^m **Spaces**: $\mathbb{R}^m$ with the norm $||\mathbf{x}|| := \left(\sum_{i=1}^{m} |x_i|^p\right)^{\frac{1}{p}}$

ℓ_p **Spaces**: These are subspaces of $\mathbb{R}^{\mathbb{N}}$ with $||\mathbf{x}|| := \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}}$

Function Spaces $L_p(\mathcal{X})$: Over $\mathcal{X}$, $||f|| := \left(\int_{\mathcal{X}} |f(x)|^p dx\right)^{\frac{1}{p}}$.

Dot Product

Given a vector space $\mathcal{X}$, a dot product is a mapping $\langle . \, , \, . \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ that satisfies, for all $\mathbf{x}, \mathbf{y}$ and $\mathbf{z} \in \mathcal{X}$ and for all $\alpha \in \mathbb{R}$:

- Symmetry: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
- Linearity: $\langle \mathbf{x}, \alpha \mathbf{y} \rangle = \alpha \langle \mathbf{x}, \mathbf{y} \rangle$
- Additivity: $\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle$

Hilbert Space

A Hilbert Space is a complete vector space $\mathcal{X}$ together with a dot product $\langle . \, , \, . \rangle$.

The dot product automatically generates a norm: $||\mathbf{x}|| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Hilbert spaces are special cases of Banach spaces.

Examples of Hilbert Spaces

Euclidean spaces and the standard dot product for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{m} x_i y_i$$

Function spaces ($L_2(\mathcal{X})$): functions on $\mathcal{X}$ with $f : \mathcal{X} \to \mathbb{C}$ for all $f, g \in \mathcal{F}$, with the dot product: $\langle f, g \rangle = \int_X \overline{f(x)} g(x) dx$

ℓ_2 series of real numbers (infinite), $\in \mathbb{R}^N$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{\infty} x_i y_i$$

Matrices

A matrix $M \in \mathbb{R}^{m \times n}$ corresponds to a linear map from $\mathbb{R}^m$ to $\mathbb{R}^n$.

A symmetric matrix $M \in \mathbb{R}^{m \times m}$ satisfies $M_{ij} = M_{ji}$.

An anti-symmetric matrix $M \in \mathbb{R}^{m \times m}$ satisfies $M_{ij} = -M_{ji}$.

Rank: Denote by I the image of $\mathbb{R}^m$ under M. $\text{rank}(M)$ is the smallest number of vectors that span I.

Matrices: orthogonality

A matrix $M \in \mathbb{R}^{m \times m}$ is orthogonal if $M^T M = \mathbf{I}$. This means $M^T = M^{-1}$.

An orthogonal matrix consists of mutually orthogonal rows and columns.

Matrix Norms

The norm of a linear operator between two Banach spaces $\mathcal{X}$ and $\mathcal{Y}$:
$||A|| := \max_{x\in\mathcal{X}} \frac{||Ax||}{||x||}$

- $||\alpha A|| = \max_{\mathbf{x}\in\mathcal{X}} \frac{||\alpha Ax||}{||x||} = |\alpha|||A||$
- $||A + B|| = \max_{\mathbf{x}\in\mathcal{X}} \frac{||(A+B)x||}{||x||} \leq \max_{x\in\mathcal{X}} \frac{||Ax||}{||x||} + \max_{x\in\mathcal{X}} \frac{||Bx||}{||x||} = ||A|| + ||B||$
- $||A|| = 0$ implies $\max_{x\in\mathcal{X}} \frac{||Ax||}{||x||}$ and thus $A\mathbf{x} = 0$ for all $\mathbf{x}$, i.e. $A = 0$.

Matrix Norms

Frobenius norm: (in analogy with vector norm)

$$||M||_{Frob}^2 = \sum_{i=1}^{m} \sum_{j=1}^{m} M_{ij}^2$$

Eigen Systems

Given M in $\mathbb{R}^{m \times m}$, then $\lambda \in \mathbb{R}$ is an eigenvalue and $\mathbf{x} \in \mathbb{R}^m$ is an eigenvector if:

$M\mathbf{x} = \lambda\mathbf{x}$

Eigen Systems, symmetric matrices

For symmetric matrices all eigenvalues are real and the matrix is fully diagonalizable (i.e. m eigenvectors).

All eigenvectors with different eigenvalues are mutually orthogonal:
Proof, for two eigenvectors $\mathbf{x}$ and $\mathbf{x}'$ with respective eigenvalues λ and λ':
$\lambda \mathbf{x}^T \mathbf{x}' = (M\mathbf{x})^T \mathbf{x} = \mathbf{x}^T (M^T \mathbf{x}') = \mathbf{x}^T (M\mathbf{x}') = \lambda' \mathbf{x}^T \mathbf{x}'$ so $\lambda' = \lambda$ or $\mathbf{x}^T \mathbf{x} = 0$.

We can decompose $M = O^T \Lambda O$.

Eigen Systems, symmetric matrices

We also have the operator norm:

$$
\begin{aligned}
||M||^2 &= \max_{x \in \mathbb{R}^m} \frac{||Mx||^2}{||x||^2} \\
&= \max_{x \in \mathbb{R}^m and ||x||=1} ||Mx||^2 \\
&= \max_{x \in \mathbb{R}^m and ||x||=1} \mathbf{x}^T M^T M \mathbf{x} \\
&= \max_{x \in \mathbb{R}^m and ||x||=1} \mathbf{x}^T O \Lambda O^T O \Lambda O^T \mathbf{x} \\
&= \max_{x \in \mathbb{R}^m and ||x'||=1} \mathbf{x}'^T \Lambda^2 \mathbf{x}' \\
&= \max_{i \in [m]} \lambda_i^2
\end{aligned}
$$

Eigen Systems, symmetric matrices

Frobenius norm:

$$\begin{aligned} ||M||^2_{Frob} &= tr(MM^T) = tr(O\Lambda O^T O\Lambda O^T) \\ &= tr(\Lambda O^T O\Lambda O^T O) = tr(\Lambda^2) = \sum_{i=1}^{m} \lambda_i^2 \end{aligned}$$

Matrices: Invariants

Trace: $tr(M) = \sum_{i=1}^{m} M_{ii}$.

$tr(AB) = tr(BA)$.

For symmetric matrices:

$$tr(M) = tr(O^T \Lambda O) = tr(\Lambda O O^T) = tr(\Lambda) = \sum_{i=1}^{m} \lambda_i$$

Determinant:

$$det(M) = \prod_{i=1}^{m} \lambda_i$$

Positive Matrices

A Positive Definite Matrix is a matrix $M \in \mathbb{R}^{m \times m}$ for which for all $\mathbf{x} \in \mathbb{R}^m$:

$\mathbf{x}^T M \mathbf{x} > 0$ if $\mathbf{x} \neq 0$

This matrix has only positive eigenvalues:
$\mathbf{x}^T M \mathbf{x} = \lambda \mathbf{x}^T \mathbf{x} = \lambda ||\mathbf{x}|| > 0$
Induced norm: $||\mathbf{x}||_M^2 = \mathbf{x}^T M \mathbf{x}$

Singular Value Decomposition

Want to find similar thing for arbitrary matrix $M \in \mathbb{R}^{m \times n}$ where $m \geq n$:

$$M = U \Lambda O$$

$$U \in \mathbb{R}^{m \times n}, \;\; U^T U = \mathbf{I}$$
$$O \in \mathbb{R}^{n \times n}, \;\; O^T O = \mathbf{I}$$
$$\Lambda = diag(\lambda_1, \lambda_2, ... \lambda_n)$$