Analytic Geometry

Norm

A norm on a vector space V is a function $||\cdot|| : V \to \mathbb{R}$ such that $\forall \lambda \in \mathbb{R},\ x, y \in V$, we have :

- $||\lambda x|| = |\lambda| ||x||$
- $||x + y|| \leq ||x|| + ||y||$
- $||x|| \geq 0$
- $||x|| = 0 \iff x = 0$

Manhattan Norm (L1 norm)

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

Euclidian Norm

$$||x||_2 = \sum_{i=1}^{n} x_i^2$$

When we say $||x||$ without any subscript ,we usually refer to the Euclidian norm.

General inner product

An inner product is any mapping $< \cdot, \cdot >: V \times V \to W$ such that $\forall\ x, y, z \in V,\ p, q \in \mathbb{R}$, we have

- $< px + qy,\ z >= p < x, z > +q < y, z >$
- $< z,\ px + qy >= p < z, x > +q < z, y >$
- $< x, y >=< y, x >$
- $< 1, 1 >$ is positive.

Suppose $\mathbf{x} = \sum x_i \mathbf{b}_i$ and $\mathbf{y} = \sum y_i \mathbf{b}_i$, then $< \mathbf{x}, \mathbf{y} >= \sum_i x_i < \mathbf{b_i}, \mathbf{y} >= \mathbf{x^T z}$ where $z_i =< \mathbf{b_i}, \mathbf{y} >=< \mathbf{b_i}, \sum_j y_j \mathbf{b_j} >= \sum_j < \mathbf{b_i}, \mathbf{b_j} > y_j = \sum_j \mathbf{r_i}^T \mathbf{y}$ where $\mathbf{r_i}$ is the column vector with the j^{th} entry being $< \mathbf{b_i}, \mathbf{b_j} >$. And $\mathbf{r_i}^T$ is the row vector.

From here, it's easy to see that $< \mathbf{x}, \mathbf{y} >= \mathbf{x^T A y}$ where $A_{ij} =< \mathbf{b_i}, \mathbf{b_j} >$ is the matrix with the i^{th} row being $\mathbf{r_i}^T$.

Since $A_{ij} = < \mathbf{b_i}, \mathbf{b_j} > = < \mathbf{b_j}, \mathbf{b_i} > = \ A_{ji}$, thus $\mathbf{A}$ is a symmetric matrix.

Moreover because $\forall \mathbf{x} \in V,\ < \mathbf{x}, \mathbf{x} >$ is positive, except for $\mathbf{x} = \mathbf{0}$ when it's 0, thus $\forall \mathbf{x} \in V - \{\mathbf{0}\},\ \ \mathbf{x^T A x} > 0$, which is what we call a positive definite matrix.

So finally, an inner product is an operation on $V \subseteq \mathbb{R}^n$ given by $< \mathbf{x}, \mathbf{y} >= \mathbf{x^T A y}$ where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a symmetric, positive definite matrix.

Induced norms

Any definition of a norm that can be expressed using an inner product as $||\mathbf{x}|| = \quad < \mathbf{x}, \mathbf{x} >$ is called an induced norm.

For any general inner product, the Cauch-Schwarz inequality guarantees that

$| < \mathbf{x}, \mathbf{y} > | \leq ||\mathbf{x}|| \; ||\mathbf{y}||$, and thus, there is always a way of defining an angle between two vectors.

Angle between vectors

For any 2 vectors $\mathbf{x}, \mathbf{y} \in V$, if the angle between them is θ , then :

$$\cos(\theta) = \frac{| < \mathbf{x}, \mathbf{y} > |}{||\mathbf{x}|| \; ||\mathbf{y}||}$$

And thus, we also have a notion of orthogonality, which is when $\cos(\theta) = 0$ and thus $< \mathbf{x}, \mathbf{y} > = 0$.

Orthonormal bases

A bases $B = \{\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_n}\}$ is called orthonormal iff $\forall i, j \in \mathbb{N}, \; i, j \leq n, \; i \neq j$, we have $< \mathbf{b_i}, \mathbf{b_i} >= 1$ and $< \mathbf{b_i}, \mathbf{b_j} >= 0$. The inner product for such a the vector space generated by this bases is basically the dot product.

Orthogonal Projections

Suppose U is a vector subspace of $V \subseteq \mathbb{R}^n$ with bases $B = \{\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_m}\}$, then for any $\mathbf{x} \in V$, the projection $\mathbf{x}_U \in U$ is the vector in U with the least distance $||\mathbf{x} - \mathbf{x}_U||$ as given by a norm induced from inner product $< \cdot, \cdot >$.

It can be shown that it is the case when $\mathbf{x} - \mathbf{x}_U$ is perpendicular to every vector in B because then, using the triangle inequality, it gives less distance than any other case.

To find the coordinates of $\mathbf{x}_U$ as expressed in B , say λ, we write the equations

$0 = < \mathbf{b_i}, \mathbf{x} - \mathbf{x}_U >=< \mathbf{b_i}, \mathbf{x} - \mathbf{B}\lambda >= \mathbf{b_i^T} \mathbf{A}(\mathbf{x} - \mathbf{B}\lambda)$. Stacking up these equation for all

i till m, we get : $\mathbf{0} = \mathbf{B^T A}(\mathbf{x} - \mathbf{B}\lambda) \implies \mathbf{B^T A B}\lambda = \mathbf{B^T A x} \implies$
$\lambda = (\mathbf{B^T A B})^{-1}\mathbf{B^T A x} \implies \mathbf{x_U} = \mathbf{B}(\mathbf{B^T A B})^{-1}\mathbf{B^T A x}$

Special case : If the default bases that everything is expressed in is an orthonormal bases , then $\mathbf{A} = \mathbf{I}$ and we basically have $\lambda = (\mathbf{B^T B})^{-1}\mathbf{B^T x}$ and $\mathbf{x_U} = \mathbf{B}(\mathbf{B^T B})^{-1}\mathbf{B^T x}$. This is what we'll use usually. This is basically the least squares solution for the problem $\mathbf{B}\lambda \approx \mathbf{x}$.

Special Special case : If the default bases as well as the bases B is orthonormal. In this case $\mathbf{B^T B} = \mathbf{I}^{m\times m}$ and thus $\lambda = \mathbf{B^T x}$ and $\mathbf{x}_U = \mathbf{B B^T x}$.

Note that here $\mathbf{BB^T}$ is not (always) equal to $\mathbf{I}^{n\times n}$.

Projection of affine spaces

To project a vector is to find the point in a subspace closest to that vector, viewed as a point. Thus, being able to project a point on any general affine space (hyperplane) enables us to do things like SVC and SVM .

The projection of a point $\mathbf{x}$ on a hyperplane $\mathbf{x_0} + U$ is given by $\mathbf{x_0} + (\mathbf{x} - \mathbf{x_0})_U$

Gram-Schmidt Orthogonalization

For a basis $B = \{\mathbf{b_1}, \mathbf{b_2} \ldots, \mathbf{b_n}\}$, define

$\mathbf{u_1} = \mathbf{b_1}$ and $\mathbf{u_k} = \mathbf{b_k} - (\mathbf{b_k})_{\text{span}(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_{k-1}})}$ for $1 < k \leq n$.

The basis $U = \{\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3} \ldots, \mathbf{u_n}\}$ is an orthonormal basis.

Here $(\mathbf{b_k})_{\text{span}(\mathbf{u_1}, \mathbf{u_2}, \mathbf{u_3} \ldots, \mathbf{u_{k-1}})}$ is the projection of $\mathbf{b_k}$ on the span of the $k-1$ elements of U that we have already calculated.