Notes on Linear Algebra

November 2, 2020

Recall that a (real) **Vector space** V is a set with an element 0, a "scalar multiplication" map $\mathbb{R} \times V \to V$ and an "addition" map $V \times V \to V$, such that, for any $x, y, z \in V$, any $a, b \in \mathbb{R}$, the followings are true:

(i) $x + y = y + x$

(ii) $(x + y) + z = x + (y + z)$

(iii) $0 + x = 0$

(iv) $1x = x$

(v) $0x = 0$

(vi) $(a + b)x = ax + bx$

(vii) $a(x+y) = ax + ay$

(viii) $(ab)x = a(bx)$

If V is a vector space, any non empty subset $V' \subset V$ which is closed under addition and scalar multiplication is called a **subspace**. The maximum number of linearly independent vectors in a vector space is called its **dimension**.

The set of n dimensional column vectors $\mathbb{R}^n$, under the usual addition and scalar multiplication, is a vector space, and it has dimension n.

The **inner product** on $\mathbb{R}^n$ is defined as $(x, y) = x^T y = \sum_i x_i y_i$. It is easy to check that this inner product satisfies the following properties:

(i) Symmetry: $(x, y) = (y, x)$

(ii) Bilinearity: $(ax + a'x', y) = a(x, y) + a'(x', y)$, $(x, by + b'y') = b(x, y) + b'(x, y')$.

(iii) Positive definiteness: $(x, x) \geq 0$ and $(x, x) = 0$ iff $x = 0$.

Let V be a subspace of $\mathbb{R}^n$. We call a basis of V **orthogonal** if the inner product of distinct basis vectors are all 0, **orthonomal** if in addition, the inner product of any basis vector with itself is 1.

Given any basis $\{x_1, \ldots, x_d\}$ of a subspace $V \subset \mathbb{R}^n$, we can make it into an orthogonal or orthonormal basis via the **Gram-Schmidt process**:

$$y_1 = x_1$$

$$y_i = x_i - \sum_{j<i} ((y_j, x_i)/(y_j, y_j))y_j$$

Then $\{y_i\}$ is an orthogonal basis, and $\{(y_i, y_i)^{-1/2} y_i\}$ is an orthonormal basis.

If V is a subspace of $\mathbb{R}^n$, $x \in \mathbb{R}^n$, we call the **orthogonal projection** of x on V, denoted as $P_V(x)$, the unique vector that satisfies $P_V(x) \in V$ and $(x - P_V(x), y) = 0$ for all $y \in V$.

For any $x' \in V$, $(x - P_V(x), x - P_V(x)) \leq (x - x', x - x')$ and equality happens iff $x' = P_V(x)$.

To calculate $P_V(x)$, we can use either of these formulas:

(i) If $\{x_i\}$ is an orthonormal basis of V, then $P_V(x) = \sum_i (x, x_i) x_i$.

(ii) If $\{x_i\}$ is an orthogonal basis of V, then $P_V(x) = \sum_i ((x, x_i)/(x_i, x_i)) x_i$.

(iii) If $\{x_i\}$ is just a basis of V, let $X = [x_1, \ldots x_d]$ be a $n \times d$ matrix, then

$$P_V(x) = X(X^T X)^{-1} X^T x = \sum_i (\sum_j a_{ij}(x_j, x)) x_i$$

Where $A = [a_{ij}] = [(x_i, x_j)]^{-1}$ is a $d \times d$ matrix.

If one replace (x, y) with $(x, y)_A$ defined as $x^T A y$, where A is a symmetric matrix with all eigenvalues positive, then $(\cdot, \cdot)_A$ still satisfies symmetry, bilinearity and positive definiteness, and all the conclusions about $(\cdot, \cdot)$ above are still valid.