

# **STA290**

## **Winter 2014**

*Selected presentation materials*

# Contents

<b>1</b>	<b>Abstract vector spaces</b>	<b>3</b>
1.1	Coordinate free projections . . . . .	3
1.2	Projections for Gaussians . . . . .	5

# Chapter 1

## Abstract vector spaces

### 1.1 Coordinate free projections

In this chapter we consider some abstract vector space, call it  $V$ , which has a norm  $\|\cdot\|$  induced by an inner product  $\langle\cdot,\cdot\rangle$  (i.e.  $\|v\|^2 := \langle v, v \rangle$ ). The elements of  $V$  are called vectors but since we are considering abstract spaces we do not necessarily have access to the “coordinates” of each vector. What we do have is the ability to construct an orthonormal basis of  $V$ , usually denoted something like  $\phi_1, \phi_2, \dots, \phi_d$ , where  $d$  is the dimension of  $V$ . Then each vector in  $v$  has a decomposition in terms of these basis vectors. This note explores to computing things like norms and inner products and taking projections with these basis vectors.

The idea is that  $V$  will denote something like points in space. In physics, points in space should exist independently of what coordinate system one uses. In particular, the notion of distance and inner product exist *before* we attach a coordinate system to this space. The following chapter talks about how to work with such vectors. The main story is that one can choose a particular orthonormal basis and then the coefficients of the basis decomposition behave exactly as regular coordinates in  $\mathbb{R}^d$ . Moreover, if one wants to project a vector in  $V$  to a linear subspace  $M \subset V$  then by choosing the basis vectors appropriately one can easily perform the desired projection by simply truncating the basis representation.

**Claim 1 (Coefficients are coordinates).** *Let  $\phi_1, \dots, \phi_d$  be an orthonormal basis for  $V$ . Then any  $v \in V$  has a unique representation*

$$v = \sum_{k=1}^d c_k \phi_k, \text{ where } c_k = \langle v, \phi_k \rangle.$$

Moreover, these basis coefficients behave exactly like coordinates in regular Euclidean space so that if  $v = \sum_{k=1}^n c_k \phi_k$  and  $w = \sum_{k=1}^n d_k \phi_k$ , then

$$\langle v, w \rangle = \sum_{k=1}^d c_k d_k.$$

In particular,  $\|v\|^2 = \sum_{k=1}^d c_k^2$ .

Now suppose  $M \subset V$  is an  $m$ -dimensional linear subspace (the zero vector should be in here). We can define  $M^\perp$  to be the set of all vectors in  $V$  which are orthogonal to every vector in  $M$ . In this case we can write  $M \oplus M^\perp = V$  to signify that every  $v \in V$  has a unique decomposition  $v_M + v_M^\perp$  where  $v_M \in M$  and  $v_M^\perp \in M^\perp$ . Often, in statistics, one needs to project a vector  $v \in V$  to a subspace  $M$ . This operation is denoted  $P_M v$  and is technically defined as  $P_M v := \operatorname{argmin}_{w \in M} \|w - v\|^2$ . The following theorem shows that with a judicious choice of your orthonormal basis this projection is easily calculated

**Claim 2 (Projections are easy).** *If one constructs the orthonormal basis  $\phi_1, \dots, \phi_d$  of  $V$  in such a way that*

$$M = \operatorname{span}\{\phi_1, \dots, \phi_m\}$$

*then for any vector  $v = \sum_{k=1}^d c_k \phi_k$  the projection to  $M$  is computed by truncating the decomposition to  $m$ :*

$$P_M v = \sum_{k=1}^m c_k \phi_k. \tag{1.1}$$

Moreover, since  $P_M v$  must be orthogonal to  $P_{M^\perp} v$ , and  $P_{M^\perp} v = v - P_M v$ , we have that

$$P_M v \perp (v - P_M v).$$

The following is a simple consequence of the above claim, but it will be important later so we state it here

**Corollary 1.** *If  $M$  is a linear subspace of  $V$  and  $v, w \in V$  then*

$$\langle v, P_M w \rangle = \langle P_M v, P_M w \rangle = \langle P_M v, w \rangle.$$

## 1.2 Projections for Gaussians

If we are working in an abstract vector space  $V$ , what do we even mean by a Gaussian vector in  $V$ ? A random vector  $Y$  is said to be Gaussian if  $\langle v, Y \rangle$  is a Gaussian random variable for all  $v \in V$ . Now we want some notion of  $Y \sim \mathcal{N}(0, \sigma^2 I_d)$ . To be able to say  $Y \sim \mathcal{N}(0, \sigma^2 I_d)$  we simply require that there exists a orthonormal basis representation  $Y = \sum_{k=1}^d d_k \psi_k$  where  $d_1, \dots, d_d$  are iid  $\mathcal{N}(0, \sigma^2)$ .

A fundamental fact about independent mean zero Gaussian random variables is that they are invariant under rotations. In particular if  $W \sim \mathcal{N}(0, \sigma^2 I_d)$  where  $W$  is a random vector in Euclidean space then  $UW \sim \mathcal{N}(0, \sigma^2 I_d)$  for any orthogonal rotation matrix ( $U$  is a rotation if  $I = U^t U$ ). In fact, independent mean zero Gaussians make up the only multivariate distribution with independent coordinates that is rotationally invariant. You can think of left multiplication by a rotation matrix as simply changing basis. Therefore if  $Y$  is an abstract random vector that satisfies  $Y \sim \mathcal{N}(0, \sigma^2 I)$ , then *any* basis decomposition of  $Y$  should give iid  $\mathcal{N}(0, \sigma^2)$  coefficients.

**Claim 3.** *Suppose  $Y$  is an abstract Gaussian random vector in  $V$  which satisfies  $Y \sim \mathcal{N}(0, \sigma^2 I)$ . If  $\phi_1, \dots, \phi_d$  is a orthonormal basis of  $V$  then*

$$Y = \sum_{k=1}^d c_k \phi_k \text{ implies } c_1, \dots, c_d \text{ are iid } \mathcal{N}(0, \sigma^2).$$

Once we have the above theorem the following corollary is easy to prove.

**Corollary 2.** *Suppose  $\phi_1, \dots, \phi_d$  is a orthonormal basis of  $V$  and  $Y \sim \mathcal{N}(0, \sigma^2 I_d)$ . Suppose  $M$  is an  $m$ -dimensional linear subspace of  $V$  which is spanned by  $\phi_1, \dots, \phi_m$ . If  $Y = \sum_{k=1}^d c_k \phi_k$  then the following three statements are true:*

- (i)  $\|P_M Y\|^2 = c_1^2 + \dots + c_m^2 \sim \sigma^2 \chi_m^2$ ;
- (ii)  $P_M Y$  is independent of  $Y - P_M Y$ ;
- (iii)  $\text{var}(\langle v, Y \rangle) = \text{var}(v_1 c_1 + \dots + v_d c_d) = \sigma^2 \|v\|^2$ .