LINEAR ALGEBRA

Louis de Thanhoffer de Volcsey, Ph.D.

Contents

---

0 Overview

1 the Normal equation

In this section, we'll refresh the reader on projections onto images of linear maps: The starting point of this construction is the following lemma:

Lemma 1.1. *Let $W \subset V$ be a subspace of the inner product space V. Let $v \in V$ and $u \in W$. Then the following are equivalent:*

1. $(v - u) \perp W$

2. $\operatorname{argmin}_{w \in W} ||v - w|| = u$

Proof. Let $w \in W$.
Assuming $u \in W$ satisfies the first condition, we have $v - u \perp u - w$ and by the Pythagorean theorem, we have:

$$||v - w||^2 = ||v - u||^2 + ||u - w||^2 \geq ||v - u||^2$$

Proving the second condition.
Conversely, we assume $u \in W$ satisfies the second condition and apply the following trick:
Consider the function:

$$\phi : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto ||v - u + tw||^2$$

Since $||v - u||^2$ is minimal, ϕ has a minimum at $t = 0$. Moreover, ϕ is differentiable so that $\phi'(0) = 0$. It's also easy to see that

$$\phi(t) = ||v - u||^2 + 2 \cdot t \langle v - u \,, w \rangle + t^2 ||w||^2$$

Hence $0 = \phi'(0) = 2 \langle v - u \,, w \rangle$, and $v - u \perp w$ as required X

Lemma 1.2. *Let $W \subset V$ be a subspace of a finite-dimensional inner product space. Then the map*

$$\pi : V \longrightarrow W : v \mapsto \mathrm{argmin}_{w \in W} ||v - w||$$

is well defined

Proof. By the above lemma we need to show that for any $v \in V$, there exists a unique $\pi(v) \in W$ such that $v - \pi(v) \perp W$. To this end, we look at the following map $f \in W^*$

$$f : W \longrightarrow \mathbb{R} : w \mapsto \langle v \, , w \, \rangle$$

Since W^* is in turn an inner product space, f can be written in the form $\langle \pi(v) \, , - \rangle$ for some unique $\pi(v) \in W$. The result now follows X

The lemma above motivates the following definition:

Definition 1.3. Let $W \subset V$ be a subspace of a finite dimensional inner product space. Then the unique map $\pi : V \longrightarrow W$ defined by

$$\pi(v) \stackrel{\text{def}}{=} \mathrm{argmin}_{w \in W} ||v - w||$$

is *the projection of V onto W*

Corollary 1.4. *Let $W = \mathrm{im}(f) \oplus \mathrm{im}(f)^{\perp}$ and $\pi_{\mathrm{im}(f)} : W \longrightarrow \mathrm{im}(f)$ be the canonical projection. Then $\pi_{\mathrm{im}(f)}$ coincides with the projection onto the subspace $\mathrm{im}(f) \subset W$ in the sense of Definition 1.3*

Proof. This follows immediately from 1.1 X

Next, we consider a linear map $f \in \mathrm{Hom}_{\Bbbk}(U, V)$ and let W be the subspace $\mathrm{im}(f)$. One can give a more explicit description of the projection $\pi : V \longrightarrow W$:

Lemma 1.5. *Let $f \in \mathrm{Hom}(U, V)$ and and $v \in V$ Then the following are equivalent:*

1. *$f(u)$ is the projection of v onto the subspace $\mathrm{im}(f) \subset V$*
2. *The vector $u \in U$ satisfies $(f^* \circ f)(u) = f^*(v)$*

Proof. $f(u)$ is the projection of v onto $\mathrm{im}(f)$ if and only if $\langle v - f(u) \, , f(u') \, \rangle$ for any $u' \in U$ by Lemma 1.1. Now,

$$\langle v - f(u) \, , f(u') \, \rangle = \langle f^*(v) - f^* f(u) \, , u' \, \rangle$$

This last expression is 0 if and only if $f^*(v) - f^* f(u) = 0$ since the inner product is nondegenerate X

The above lemma justifies the following definition:

Definition 1.6. Let $f \in \mathrm{Hom}(V, W)$ and $w \in W$.
We say that $v \in V$ satisfies the normal equation if and only if

$$(f^* \circ f)(v) = f^* w$$

In this terminology, we can restate lemma 1.5 as follows:

Lemma 1.7. *Let $f \in \mathrm{Hom}(V, W)$ and $w \in W$ Then the following are equivalent:*

1. $w = \mathrm{argmin}_{u \in V} ||w - f(u)||$
2. *v is a solution to the normal equation $(f^* \circ f)(v) = f^* w$*

It is finding the solutions to this equation that we are interested in. It turns out that one can give an explicit description of them using the so-called *Moore-Penrose pseudo-inverse.* Since this construction seems to be a little less well covered in standard linear algebra literature, we'll discuss in detail below:

2 the (Moore-Penrose) Pseudo-inverse

In this section, we will let V, W be finite-dimensional vector spaces and $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$.
It is well-known that f does not have an inverse in general. There is however a natural generalization of the notion of inverse which can be defined for *any* map: a *pseudo-inverse.* More precisely, if f either has a nonzero kernel or if the image of f is not the whole of W, then the inverse of f will not exist. One natural way to remediate this issue is to consider complements for both subspaces and write

$$V \stackrel{\text{def}}{=} \ker(f) \oplus U_V \text{ and } W \stackrel{\text{def}}{=} \mathrm{im}(f) \oplus U_W$$

It's easy to see that restricting f to appropriate subspaces now does produce an invertible map as follows:

Lemma 2.1. *the map $f : U_V \longrightarrow \mathrm{im}(f)$ is an isomorphism.*

We'll denote the inverse of f on U_V by $f^\sharp : \mathrm{im}(f) \longrightarrow U_V$. A pseudo-inverse is now the natural lift of $f^\sharp$ to the whole of W:

Lemma 2.2. *There exists a unique map $f^\sharp : W \longrightarrow U_V$ making the following diagram commute:*

$$\begin{tikzcd}
W \arrow[d, "\pi_{\mathrm{im}(f)}"'] \arrow[dr, "f^\sharp"] & \\
\mathrm{im}(f) \arrow[r, "f^\sharp"'] & U_V
\end{tikzcd}$$

Proof. The commutativity of the diagram means that for $u \in U_V$, we have

$$f^\sharp(w) \stackrel{\text{def}}{=} u \iff f^\sharp(\pi_{\mathrm{im}(f)}(w)) = u \iff \pi_{\mathrm{im}(f)}(w) = f(u)$$

Where the second equivalence follows from the fact that $f^\sharp$ is the inver of f on U_V.
The claim will thus follow if we show that the above assignment is indeed a well-defined linear map. To this end assume that $u, u' \in U_V$ satisfy $f(u') = \pi_{\mathrm{im}(f)}(w) = f(u)$.
Then $u - u' \in \ker(f)$, hence $u - u' \in \ker(f) \cap U_V$ in particular. Now since $\ker(f) \oplus U_V = V$, we have $u - u' = 0$, so that $u = u'$, showing the well-definedness.
We leave the linearity to the reader. X

It will be helpful to note that the map $f^\sharp \in \mathrm{Hom}(W, V)$ can also be characterized by $\mathrm{im}(f^\sharp) \subset U_V$ and $f \circ f^\sharp = \pi_{\mathrm{im}(f)}$.
To give the map $f^\sharp$ a name, we first let $\Lambda(f)$ denote the set

$$\Lambda(f) \stackrel{\text{def}}{=} \{(U_V, U_W) |\ \ker(f) \oplus U_V = V \text{ and } \mathrm{im}(f) \oplus U_W = W\}$$

and conclude from Lemma 2.2 that there is a assignment:

$$\Phi : \Lambda(f) \longrightarrow \mathrm{Hom}_{\mathbb{R}}(W, V) : (U_V, U_W) \mapsto f^\sharp$$

where $f^\sharp \in \mathrm{Hom}_{\mathbb{R}}(W, V)$ is the unique map satisfying

$$f \circ f^\sharp = \pi_{\mathrm{im}(f)} \text{ and } \mathrm{im}(f^\sharp) \subset U_V$$

Let's denote the image of Φ by $\Pi(f)$. Summarizing the discussion, we make the following:

Definition 2.3. Let $(U_V, U_W) \in \Lambda(f)$. Then the pseudo-inverse of (U_V, U_W, f) is the map $\Phi(f)$.
We say that $g \in \mathrm{Hom}_{\mathbb{R}}(W, V)$ is a pseudo-inverse to f if $g \in \Pi(f)$

We can give a slightly different description of pseudo-inverses by describing them on the 2 components in the decomposition $\mathrm{im}(f) \oplus U_W = W$:

Lemma 2.4. *Let* (U_V, U_W) *in* $\Lambda(f)$*. Then the following are equivalent:*

1. $f^\sharp$ *is the pseudo-inverse to* (U_V, U_W, f)
2. $f^\sharp|_{\mathrm{im}(f)}$ *is the inverse to* $f : U_V \longrightarrow \mathrm{im}(f)$ *and* $f^\sharp|_{U_W} = 0$

Proof. Since the pseudo=inverse to (U_V, U_W, f) is unique, it suffices to show that the pseudo-inverse indeed satisfies the conditions of (2). The fact that $f^\sharp|_{\mathrm{im}(f)}$ is the inverse of $f|_{U_V}$ follows from

$$(f \circ f^\sharp)|_{\mathrm{im}(f)} = \big(\pi_{\mathrm{im}(f)}\big)|_{\mathrm{im}(f)} = \mathrm{Id}\,|_{\mathrm{im}(f)}$$

Moreover, if $w \in U_W$, then $\pi_{\mathrm{im}(f)}(w) = 0$ since $\mathrm{im}(f) \oplus U_W$. Hence $f^\sharp(w) = f^\sharp(\pi_{\mathrm{im}(f)}(w)) = 0$ by Lemma 2.2 X

Our next order of business is to give an explicit description of the set $\Pi(f)$ of pseudo-inverses to f. We begin by showing that we can describe the complements U_V and U_W solely by using the maps f and $f^\sharp$:

Lemma 2.5. *Let* $f^\sharp$ *be the pseudo-inverse to* (U_V, U_W, f)*. Then* $U_V = \mathrm{im}(f^\sharp)$ *and* $U_W = \ker(f^\sharp)$

Proof. We have $\mathrm{im}(f^\sharp) \subset U_V$ by Definition 2.3 . Moreover, $f^\sharp$ is a composition of surjections and hence itself surjective, proving the first claim.
To prove the second claim, note that the second condition of Lemma 2.4 immediately implies that $U_W \subset \ker(f^\sharp)$. We can also show the other inclusion by assuming that $w \in W$ satisfies $f^\sharp(w) = 0$, in which case $\pi_{\mathrm{im}(f)}(w) = f(f^\sharp(w)) = f(0) = 0$, implying that w lies in the component U_W of the decomposition $\mathrm{im}(f) \oplus U_W = W$ as required X

Taking the above lemma one step further allows us to describe the set $\Pi(f)$ of pseudo-inverses as promised:

Lemma 2.6. *Let* $f \in \mathrm{Hom}(V, W)$*. Then the following are equivalent:*

1. $g \in \Pi(f)$
2. $(f \circ g)|_{\mathrm{im}(f)} = \mathrm{Id}$ *and* $(g \circ f)|_{\mathrm{im}(g)} = \mathrm{Id}$

Proof. Let g be a pseudo-inverse to f and define $U_V \stackrel{\text{def}}{=} \mathrm{im}(g)$ and $U_W \stackrel{\text{def}}{=} \ker(f)$. Then Lemma 2.5 shows that g is in fact the pseudo-inverse to the triple $\big(U_V, U_W, f\big)$. Now, since $g|_{\mathrm{im}(\,f)}$ is the inverse to $f|_{U_V}$ by Lemma 2.4, we have $(f \circ g)|_{\mathrm{im}(f)} = \mathrm{Id}$ and $(g \circ f)|_{\mathrm{im}(g)} = (g \circ f)|_{U_V} = \mathrm{Id}$.
Conversely, assume that g satisfies the conditions in (2).
We begin by showing that $\big(\mathrm{im}(g), \ker(g)\big) \in \Lambda(f)$. Let's show that $\mathrm{im}(f) \oplus \ker(g) = W$ by way of example. Indeed, first note that $\mathrm{im}(f) \cap \ker(g) = 0$, as any w in this intersection must satisfy $w = (f \circ g)(w) = f(0) = 0$. Moreover, if we write $w = \big(w - f(g(w))\big) + f(g(w))$, we see that trivially $f(g(w)) \in \mathrm{im}(f)$ and

$$g(w - f(g(w))) = g(w) - (g(f(g(w)) = g(w) - g(w) = 0$$

so that $\big(w - f(g(w))\big) \in \ker(g)$. This indeed shows that $\mathrm{im}(f) \oplus \ker(g) = W$. The proof of $\mathrm{im}(g) \oplus \ker(f) = V$ is completely analogous, allowing us to conclude that $(\mathrm{im}(g), \ker(g)) \in \Lambda(f)$.
It now remains to show that g is indeed a pseudo-inverse to the triple $(\mathrm{im}(g), \ker(f), f)$. By Lemma 2.4, it suffices to show that $g|_{\mathrm{im}(f)}$ is the inverse to $f|_{\mathrm{im}(g)}$ and that $g|_{\ker(g)} = 0$. The first claim follows immediately from the fact that g is a left inverse to $f : \mathrm{im}(g) \longrightarrow W$ and the second claim is trivial. X

In order to summarize the previous 2 lemmas, we introduce the following assignment, which is well-defined by Lemma 2.5

$$\Psi : \Pi(f) \longrightarrow \Lambda(f) : g \mapsto \big(\mathrm{im}(g), \ker(g)\big)$$

We now have:

Lemma 2.7. *Let $f \in \mathrm{Hom}(V, W)$. Then:*

- $\Pi(f) = \{g \in \mathrm{Hom}(W, V) \mid (f \circ g)|_{\mathrm{im}(f)} = \mathrm{Id} \text{ and } (g \circ f)|_{\mathrm{im}(g)} = \mathrm{Id}\}$
- *The assignments Φ and Ψ define 1:1 correspondences between $\Lambda(f)$ and $\Pi(f)$*

Proof. The first claim simply restates Lemma 2.6. To prove the second, we note that $\Psi \circ \Phi = \mathrm{Id}$ by Lemma 2.5. Moreover, Φ is surjective by definition, implying that $\Phi \circ \Psi = \mathrm{Id}$ as well X

We finish our discussion of pseudo-inverses by discussing a special choice of pseudo-inverse in $\Pi(f)$ that one can make if the vector spaces V and W are equipped with inner products. Indeed, recall the following standard result:

Lemma 2.8. *Let $U \subset V$ be a subspace of a finite dimensional inner product space. Then $U \oplus U^{\perp} = V$*

This leads us to the following Definition:

Definition 2.9. Let V, W be finite-dimensional inner product spaces and let $f \in \mathrm{Hom}_{\mathbb{R}}(V, W)$. Then the *Moore-Penrose pseudo-inverse* is the pseudo-inverse to the triple $(\ker(f)^{\perp}, \mathrm{im}(f)^{\perp}, f)$.
We will denote it by f^{+}

It turns out that we can give a very satisfying description of Moore-Penrose pseudo-inverses:

Lemma 2.10. *Let V, W be finite-dimensional inner product spaces and $f \in \mathrm{Hom}(V, W)$. Then the following are equivalent:*

1. *g is the Moore-Penrose pseudo-inverse f^{+} to f*
2. *g is a pseudo-inverse to f and $g \circ f$ and $f \circ g$ are self-adjoint linear maps*
3. *f and g satisfy $f \circ g \circ f = f$, $g \circ f \circ g = g$, $(g \circ f)^* = g \circ f$ and $(f \circ g)^* = f \circ g$*

Proof. The equivalence (2) $\iff$ (3) is simply a restatement of Lemma 2.7.
We now prove (2) $\implies$ (1):
Assume that g is a pseudo-inverse to f and that $g \circ f$ and $f \circ g$ are both self-adjoint. then Lemma 2.5 implies that g is the pseudo-inverse to the triple $\big(\mathrm{im}(g), \ker(f), f\big)$. The claim will thus follow if we show that $\mathrm{im}(g) = \ker(f)^{\perp}$ and $\ker(g) = \mathrm{im}(f)^{\perp}$. By way of example, we will prove the former equality: First note that since $\mathrm{im}(g) \oplus \ker(f) = V$, it suffices to show that $\mathrm{im}(g) \perp \ker(f)$. Indeed, for $w \in W$ and $v \in \ker(f)$, we have:

$$\langle v\,, g(w)\,\rangle = \langle v\,, (g \circ f)(g(w))\,\rangle = \langle (g \circ f)^*(v)\,, g(w)\,\rangle = \langle (g \circ f)(v)\,, g(w)\,\rangle = \langle g(0)\,, g(w)\,\rangle = 0$$

The proof of $\ker(g) = \mathrm{im}(f)^{\perp}$ is analogous.
Finally, we show (1) $\implies$ (2):
Assume that g is the Moore Penrose pseudo-inverse to f. Ie g is the pseudo-inverse to the triple $(\ker(f)^{\perp}, \mathrm{im}(f)^{\perp}, f)$.
We will show that $(f \circ g)$ is self-adjoint and leave the other claim to the reader. To this end, let $v, v' \in V$.
Then

$$\begin{aligned}\langle v\,, g(f(v'))\,\rangle &= \Big\langle v - g(f(v)) + g(f(v)), g(f(v')) - v' + v' \Big\rangle \\ &= \Big\langle v - g(f(v)), g(f(v')) \Big\rangle + \Big\langle g(f(v)), g(f(v')) - v' \Big\rangle + \Big\langle g(f(v)), v' \Big\rangle\end{aligned}$$

Now, since $f \circ g \circ f = f$, we conclude that $v - g(f(v))$ and $g(f(v')) - v'$ lie in $\ker(f)$. Moreover, since $\ker(f) = \mathrm{im}(g)^{\perp}$, we conclude that

$$\langle v - g(f(v))\,, g(f(v'))\,\rangle = \langle g(f(v))\,, g(f(v')) - v'\,\rangle = 0$$

So that

$$\langle v \,, g(f(v')) \,\rangle = \langle g(f(v)) \,, v' \,\rangle$$

implying that $f \circ g = (f \circ g)^*$. The equality $g \circ f = (g \circ f)^*$ is completely analogous. X

As mentioned in the introduction of this section, our main motivation for studying the Moore-Penrose pseudo-inverse, is to provide a description of the projection of a vector onto the image of a linear map. We begin with the following preparatory lemma:

Lemma 2.11. *Let V, W be finite-dimensional inner product spaces and $f \in \mathrm{Hom}(V, W)$. Let $v \in V$ and $w \in W$. Finally denote the Moore-Penrose inverse of f by f^+. Then the following are equivalent:*

1. *$f(v)$ is the projection of w onto the subspace* $\mathrm{im}(f)$
2. *v satisfies the normal equation $(f^* \circ f)(v) = f^*(w)$*
3. *v lies in the affine subspace $f^+(w) + \ker(f)$*

Proof. The equivalence of (1) $\iff$ (2) is simply a restatement of Lemma 1.5.
To show the equivalence of (1) $\iff$ (3), we first note that $f(f^+w)) = \pi_{\mathrm{im}(f)}$, where $\pi_{im(f)}$ is the projection onto the subspace $\mathrm{im}(f) \subset W$ by Lemma 1.4. This shows that the vector $f^+(w) \in V$ indeed satisfies the condition (1). Next, assume (1), so that $v \in V$ satisfies $f(v) = \pi_{\mathrm{im}(f)}(v)$ and write $v = f^+(w) + v'$. Then

$$f(v) = \pi_{\mathrm{im}(f)(w)} \iff f(f^+(w) + v') = \pi_{\mathrm{im}(f)}(w) \iff \pi_{\mathrm{im}(f)}(w) + f(v') = \pi_{\mathrm{im}(f)}(w) \iff v' \in \ker(f)$$

This proves the claim X

This lemma has an interesting corollary which allows us to write the Moore-Penrose even more explicitly which will play an important role later on:

Corollary 2.12. *Let V be a finite dimensional vector space and W a finite dimensional inner product space. Let $f \in \mathrm{Hom}(V, W)$ be injective and choose* any *inner product on V. Then*

$$f^+ = (f^* \circ f)^{-1} \circ f^*$$

Proof. Since f is injective (so that $\ker(f) = 0$), f^+ is the pseudo-inverse to the triple $(V, \mathrm{im}(f)^\perp, f)$ by Definition 2.9. It follows immediately that this condition is independent of the inner product on V. To prove the formula, simply note that $f^* \circ f$ is invertible if f is injective and apply the second criterium of Lemma 2.11 X

3 Principal Component Analysis

3.1. the Principal component basis

Throughout this section V will denote a fd. Euclidean space with inner product $\langle - \,, - \rangle$. We will also consider a finite subset $\Delta \subset V$of *data*. The goal of this section is to exhibit an orthonormal basis for V which is fits the data *suitably*. We begin by defining the principal components of Δ, a specific choice of lines which span an orthonormal basis particularly compatible with Δ.
To make our exposition clearer, we introduce the notation $\langle \Delta \,, u \,\rangle \stackrel{\mathrm{def}}{=} \{\langle x \,, u \,\rangle\}_{x \in \Delta} \in \mathbb{R}^\Delta$. Recall from Probability (TODO) that Δ defines the random variable $1_\Delta : V \longrightarrow \mathbb{R}$ whose pushforward probability is the sample probability on $\mathbb{R}$. Moreover, we can also consider the following random variable

$$X_u \stackrel{\mathrm{def}}{=} V \longrightarrow \mathbb{R} : x :\mapsto \langle x \,, y \,\rangle \cdot 1_\Delta$$

Whose pushforward probability is the explained probability. More generally, we will prefix the diferent probabilistic concepts pertaining to this RV with the word *explained*.

We begin by constructing the first principal component, whose existence is guaranteed by the following theorem:

Theorem 3.1. *Let $\overline{\Delta} = \frac{1}{|\Delta|}\sum_{x\in\Delta} x$ and let L be the set of lines through $\overline{\Delta}$ in V. Then there is a unique line $\ell \in L$ minimizing*

$$\sum_{x\in\Delta} ||x - \pi_\ell(x)||^2$$

Moreover, if we write $\ell = \overline{\Delta} + \mathbb{R}u$ with $||u|| = 1$, then the above three quantities coincide:

- $\sum_{x\in\Delta} ||x - \pi_\ell(x)||^2$
- *the explained variance of X_u*
- *The largest eigenvalue of $M^t \cdot M$ where $\phi : V \longrightarrow \mathbb{R}^\Delta$ is given by $\phi(u) = \langle \Delta\,, u\,\rangle$*

Definition 3.2. The principal component of Δ is the line $\ell \in L$ defined in the above theorem

Lemma 3.3. *Let $\ell = \mathbb{R}u$ where $||u|| = 1$. Then the following are equivalent:*

- *The mean square error $\sum_{x\in\Delta} ||x - \pi_\ell(x)||^2$ is minimal*
- $\sum_{x\in\Delta} \langle x\,, u\,\rangle^2$ *is minimal*

3.2. Application: Dimension Reduction

Proof. This is an easy exercise in bilinear forms X

Remark 3.4. *Note that if $\overline{\Delta} = 0$ (an assumption commonly made in ML), then the latter quantity is the* explained variance *of Δ in our terminology.*

For the next lemma we recall that the spectral theorem tells us that a symmetric endomorphism always has basis of eigenvectors which can be chosen to be orthonormal.

Lemma 3.5. *Assume $\overline{\Delta} = 0$. Consider the linear map*

$$\phi : V \longrightarrow \mathbb{R}^\Delta : u \mapsto \langle \Delta\,, u\,\rangle$$

Let $u_1, \ldots u_n$ be an orthormal basis of eigenvectors for $\phi^t \circ \phi \in \mathrm{End}_{\mathbb{R}}(V)$. Then

- *The eigenvalue for u_i is $\lambda_i \stackrel{\text{def}}{=} \sum_{x\in\Delta} \langle x\,, u_i\,\rangle$*
- $\sup_{u\in B(0,1)} \sum_{x\in\Delta} \langle x\,, u\,\rangle = \max_i\{\lambda_i\}$

Proof. Let $u \in V$. Then the explained variance of Δ with respect to u is

$$\sum_{x\in\Delta} \langle x\,, u\,\rangle^2 = ||\phi(u)||^2_{\mathbb{R}^\Delta} = \langle \phi(u)\,, \phi(u)\,\rangle = \big\langle u\,, (\phi^t \circ \phi)(u)\,\big\rangle$$

Writing $u \stackrel{\text{def}}{=} \sum \alpha_i u_i$ wrt the orthonormal basis of eigenvectors then yields

$$\sum_{x\in\Delta} \langle x\,, u\,\rangle^2 = \big\langle u\,, (\phi^t \circ \phi)(u)\,\big\rangle = \Big\langle \sum \alpha_i u_i\,, (\phi^t \circ \phi)(\sum \alpha_i u_i)\,\Big\rangle = \Big\langle \sum \alpha_i u_i\,, (\sum \lambda_i \alpha_i u_i)\,\Big\rangle = \sum_i \lambda_i \alpha_i^2$$

In particular if $u = u_i$, we obtain the first claim. To prove the second, we assume additionally that $u \in B(0,1)$ and let λ_n be the largest eigenvalue. Then

$$\sum_{x\in\Delta} \langle x\,, u\,\rangle^2 = \sum_i \lambda_i \alpha_i^2 \le \lambda_n \sum \alpha_i^2 = \lambda_n$$

And this value indeed gets reached by u_n X

proof of Theorem 3.1. First we claim that wlog we can assume that $\overline{\Delta} = 0$. Indeed
Next, by Lemma 3.3, we need to maximize the explained variance $\sum_{x \in \Delta} \langle x, u \rangle^2$ which by lemma 3.5 coincides with the largest eigenvalue. The line ℓ is now spanned by any corresponding eigenvector. X

Since the endomorphism $\mathcal{L} \stackrel{\text{def}}{=} \phi^t \circ \phi$ plays a crucial role in our discussion, we give a more explicit description:

Lemma 3.6. *The map $\mathcal{L}^t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is given by the matrix whose (i, j)-component is $\langle x_i, u_j \rangle$.*

Proof. Let M be said matrix. Then the j-th column is given by $M \cdot e_j$ (where e_j is of course the j^{th} standard basisvector for $\mathbb{R}^n$). X

Definition 3.7. the principal component basis of Δ is the orthonomormal basis of eigenvectors for the map $\mathcal{L}$.

3.3. Application: Dimension Reduction

One can use PCA to find a way to reduce the dimension of the dataspace:

Definition 3.8. The reduction of order n of the dataset $\Delta \subset V$ is the image of the map

Definition 3.9. The pc reconstruction of order n of the dataset Δ is