

ECE551 - Homework 1

Khoi-Nguyen Mac

September 3, 2016

1 Geometry of orthogonal transformations in Euclidean spaces

1.1

$U \in \mathbb{R}^{N \times N}$ is an orthogonal matrix, therefore:

$$U^\top U = UU^\top = I.$$

By definition,

$$\begin{aligned}\|x\|^2 &:= x^\top x = \langle x, x \rangle, \quad x \in \mathbb{R}^N \\ \Rightarrow \|Ux\| &= \langle Ux, Ux \rangle = (Ux)^\top Ux = x^\top U^\top Ux = x^\top x = \|x\|^2 \\ \Rightarrow \|Ux\| &= \|x\|, \quad \because \text{norms are non-negative.}\end{aligned}$$

1.2

By definition,

$$\begin{aligned}\langle x, y \rangle &:= y^\top x = \sum_{i=0}^{N-1} x_i y_i, \quad x, y \in \mathbb{R}^N \\ \Rightarrow \langle Ux, Uy \rangle &= (Uy)^\top Ux = y^\top U^\top Ux = y^\top x = \langle x, y \rangle.\end{aligned}$$

1.3

If $M > N$ and $U^\top U = I$, then $\text{rank}(U) = N \Rightarrow U^\top U = I_N$. Since $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^N$,

$$\begin{aligned}x^\top U^\top Ux &= x^\top I_N x = x^\top x \\ y^\top U^\top Ux &= y^\top I_N x = y^\top x.\end{aligned}$$

Hence, 1.1 and 1.2 hold.

1.4

If $M < N \Rightarrow \text{rank}(U)$ is at most M .

$$U^\top U \neq I_N.$$

Hence, 1.1 and 1.2 do not hold.

2 Some basic properties of inner product spaces

Notice that:

$$\langle a, b + c \rangle = \overline{\langle b + c, a \rangle} = \overline{\langle b, a \rangle} + \overline{\langle c, a \rangle} = \langle a, b \rangle + \langle a, c \rangle.$$

2.1 The Cauchy-Schwarz inequality

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

Proof: If $\langle x, y \rangle = 0$, then theorem holds because $\|x\|$ and $\|y\| \geq 0$.

Suppose that $x \neq 0$ and $y \neq 0$. Let $z \in \mathbb{C}$, such that:

$$z = \frac{\langle x, y \rangle}{\|y\|^2} = \frac{\overline{\langle y, x \rangle}}{\|y\|^2}.$$

We have

$$\begin{aligned} 0 &\leq \|x - zy\|^2 = \langle x - zy, x - zy \rangle = \langle x, x - zy \rangle - \langle zy, x - zy \rangle \\ &= \langle x, x \rangle - \langle x, zy \rangle - \langle zy, x \rangle + \langle zy, zy \rangle \\ &= \|x\|^2 - \bar{z}\langle x, y \rangle - z\langle y, x \rangle + z\bar{z}\|y\|^2 \\ &= \|x\|^2 - \frac{\langle y, x \rangle \langle x, y \rangle}{\|y\|^2} - \frac{\langle x, y \rangle \langle y, x \rangle}{\|y\|^2} + \frac{\langle x, y \rangle \langle y, x \rangle}{\|y\|^2} \|y\|^2 \\ &= \|x\|^2 - \frac{\langle x, y \rangle \langle y, x \rangle}{\|y\|^2} \\ &= \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2}. \end{aligned}$$

Therefore,

$$\begin{aligned} 0 &\leq \|x\|^2 - \frac{|\langle x, y \rangle|^2}{\|y\|^2} \\ \Leftrightarrow \|x\|^2 &\geq \frac{|\langle x, y \rangle|^2}{\|y\|^2} \\ \Leftrightarrow \|x\|^2 \|y\|^2 &\geq |\langle x, y \rangle|^2 \\ \Leftrightarrow \|x\| \|y\| &\geq |\langle x, y \rangle|. \end{aligned}$$

□

The equality occurs iff $x = \alpha y$, for some scalar α .

Proof: By substituting $x = \alpha y$, we have

$$|\langle x, y \rangle| = |\langle \alpha y, y \rangle| = |\alpha| |\langle y, y \rangle| = |\alpha| \|y\|^2 = |\alpha| \|y\| \|y\| = \|x\| \|y\|.$$

□

2.2 The triangle inequality

$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2$ with equality iff $y = \alpha x$

Proof:

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x + y \rangle + \langle y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle \end{aligned}$$

By Cauchy-Schwarz inequality:

$$\begin{aligned} \langle x, y \rangle &\leq \|x\| \|y\| \\ \Rightarrow \langle y, x \rangle &\leq \|x\| \|y\| \\ \Rightarrow \langle x, y \rangle + \langle y, x \rangle &\leq 2\|x\| \|y\|. \end{aligned}$$

Therefore,

$$\begin{aligned} \|x + y\|^2 &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2 \\ \Rightarrow \|x + y\| &\leq \|x\| + \|y\| \quad \because \text{norms are non-negative.} \end{aligned}$$

If $y = \alpha x$ then the equality of Cauchy-Schwarz theorem occurs. Therefore,

$$\begin{aligned} \langle x, y \rangle + \langle y, x \rangle &= 2\|x\| \|y\| \\ \Rightarrow \|x + y\|^2 &= \|x\|^2 + \|y\|^2 + 2\|x\| \|y\| = (\|x\| + \|y\|)^2 \\ \Rightarrow \|x + y\| &= \|x\| + \|y\|. \end{aligned}$$

□

2.3 Parallelogram law

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

Proof:

$$\begin{aligned}
LHS &= \langle x + y, x + y \rangle + \langle x - y, x - y \rangle \\
&= (\|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle) + (\|x\|^2 + \|y\|^2 - \langle x, y \rangle - \langle y, x \rangle) \\
&= 2(\|x\|^2 + \|y\|^2) = RHS.
\end{aligned}$$

□

3 Least-squares approximation with orthonormal bases

3.1

We notice that $\langle x, \varphi_i \rangle \varphi_i$ is the orthogonal projection of x onto φ_i . We want to show that $\sum_{\varphi_i \in \hat{\mathcal{B}}} \langle x, \varphi_i \rangle \varphi_i$ is the orthogonal projection of x onto the subspace $\text{span}(\hat{\mathcal{B}}) = \hat{V}$:

$$\langle \hat{x}, \varphi_i \rangle = \langle \langle x, \varphi_i \rangle \varphi_i, \varphi_i \rangle = \langle x, \varphi_i \rangle \langle \varphi_i, \varphi_i \rangle = \langle x, \varphi_i \rangle, \forall \varphi_i \in \hat{\mathcal{B}}.$$

$$\begin{aligned}
&\langle x, \varphi_i \rangle = \langle \hat{x}, \varphi_i \rangle, \forall \varphi_i \in \hat{\mathcal{B}} \\
&\Leftrightarrow \langle x, \varphi_i \rangle - \langle \hat{x}, \varphi_i \rangle = 0, \forall \varphi_i \in \hat{\mathcal{B}} \\
&\Leftrightarrow \langle x - \hat{x}, \varphi_i \rangle, \forall \varphi_i \in \hat{\mathcal{B}} \\
&\Leftrightarrow x - \hat{x} \perp \hat{V}
\end{aligned}$$

Let $z = x - \hat{x}$, where $\hat{x} = \sum_{\varphi_i \in \hat{\mathcal{B}}} \langle x, \varphi_i \rangle \varphi_i$. Thus, $z \perp \hat{V}$.

Since any vectors in \hat{V} can be written as $f(\alpha) = \sum_{\varphi_i \in \hat{\mathcal{B}}} \alpha_i \varphi_i$ (α is a vector and α_i 's are scalars,) we have:

$$\begin{aligned}
\|x - f(\alpha)\|^2 &= \|x - \hat{x} + \hat{x} - f(\alpha)\|^2 = \|z + \hat{x} - f(\alpha)\|^2 \\
&= \langle z + \hat{x} - f(\alpha), z + \hat{x} - f(\alpha) \rangle \\
&= \|z\|^2 + \|\hat{x} - f(\alpha)\|^2 + \langle z, \hat{x} - f(\alpha) \rangle + \langle \hat{x} - f(\alpha), z \rangle
\end{aligned}$$

We can see that $\hat{x} \in \hat{V}$ and $f(\alpha) \in \hat{V}$, therefore $\hat{x} - f(\alpha) \in \hat{V}$. Since $z \perp \hat{V} \Rightarrow z \perp \hat{x} - f(\alpha) \Rightarrow \langle z, \hat{x} - f(\alpha) \rangle = \langle \hat{x} - f(\alpha), z \rangle = 0$.

Therefore, $\|x - f(\alpha)\|^2 = \|z\|^2 + \|\hat{x} - f(\alpha)\|^2$. Since $\|\hat{x} - f(\alpha)\|^2 \geq 0$,

$$\begin{aligned}
&\|x - f(\alpha)\|^2 \geq \|z\|^2 \\
&\Leftrightarrow \|x - f(\alpha)\|^2 \geq \|x - \hat{x}\|^2 \\
&\Leftrightarrow \|x - f(\alpha)\| \geq \|x - \hat{x}\|
\end{aligned}$$

3.2

We know that any inner products can define the valid norm

$$\|x\| = \sqrt{\langle x, x \rangle}.$$

Therefore $\|x - f(\alpha)\|^2 = \langle z + \hat{x} - f(\alpha), z + \hat{x} - f(\alpha) \rangle$, with other inner products. The expansion

$$\langle z + \hat{x} - f(\alpha), z + \hat{x} - f(\alpha) \rangle = \|z\|^2 + \|\hat{x} - f(\alpha)\|^2 + \langle z, \hat{x} - f(\alpha) \rangle + \langle \hat{x} - f(\alpha), z \rangle$$

also uses the properties of general inner products. Hence, the proof holds for other kinds of inner products, instead of only standard Euclidean one.

4 Signal Sets and Spaces

4.1

For $S = \mathbb{R}$, we have

$$\mathbb{R}^I = \{v \mid v : I \rightarrow \mathbb{R}\}.$$

We know that \mathbb{R} is closed under addition and scalar multiplication (\mathbb{R} is a vector space.) Therefore, for $t \in I$ and $u, v \in \mathbb{R}^I$,

$$\begin{aligned} u[t], v[t] &\in \mathbb{R} \\ \Rightarrow u[t] + v[t] &\in \mathbb{R} \\ \Rightarrow (u + v)[t] &:= u[t] + v[t] \in \mathbb{R}. \end{aligned}$$

For scalar $\alpha \in \mathbb{R}$,

$$\begin{aligned} \alpha u[t] &\in \mathbb{R} \\ \Rightarrow (\alpha u)[t] &:= \alpha u[t] \in \mathbb{R} \end{aligned}$$

4.2

- (i) For complex-valued sequences indexed by the integers, signal values live in complex space and indices live in integer space, i.e.

$$\mathbb{C}^I = \{v \mid v : \mathbb{Z} \rightarrow \mathbb{C}\}.$$

Since \mathbb{C} is also a vector space, this signal set is linear. The proof is similar as in 4.1.

Zero vector in \mathbb{R}^I is defined as $\{a_i\}_{i \in I}$, where $a_i = 0, \forall i \in I$.

- (ii) For 8-bit RGB color (three channels) digital photos of dimension $W \times H$, we can choose the value set to be $\mathcal{B}_8 = \{\overline{x_0 x_1 \dots x_7}\}, x_i = \{0, 1\}, i = \{0, 7\}$, denoting the set of all binary sequences with length of 8. The indices are chosen as set of 3 natural numbers, corresponding to the width, height, and channel, i.e.

$$\mathcal{B}_8^I = \{v \mid v : \mathbb{N}^{W \times H \times 3} \rightarrow \mathcal{B}_8\}.$$

This set is not linear because it is not closed under addition, e.g. $(11111111)_2 + (00000001)_2 = (100000000)_2 \notin \mathcal{B}_8$. We can also see that, in general, \mathcal{B}_k with finite k is not a linear space.

- (iii) For 32-bit floating point buffers containing 1 second of stereo audio at 48KHz, we can choose the value space as \mathcal{B}_{32} , with similar definition as \mathcal{B}_8 . Since the audio is 1-second long with 48KHz, there are 48k samples. Therefore, the signal set can be described as:

$$\mathcal{B}_{32}^I = \{v \mid v : \mathbb{N}^{48K \times 2} \rightarrow \mathcal{B}_{32}\}.$$

We can see that \mathcal{B}_{32} is not a linear, therefore the signal set is not linear.

4.3

We can consider signals with indices $I_1 \subset I_2$ are truncated version of signals with indices I_2 , where the signals in $I_d = I_2/I_1$ are reduced to 0. So \mathbb{R}^{I_1} is a subset of \mathbb{R}^{I_2} . Since \mathbb{R} is a subspace, \mathbb{R}^{I_1} is closed under addition and scalar multiplication. Hence, \mathbb{R}^{I_1} is a subspace of \mathbb{R}^{I_2} .

4.4

Linearity Let $u, v \in \mathbb{R}^I$ and α, β be scalars

$$T(\alpha u + \beta v)_k = T(\alpha u)_k + T(\beta v)_k = \alpha(Tu)_k + \beta(Tv)_k.$$

Hence, T is linear.

Invertibility Let $Tu = v$. T can be seen as a permutation matrix that maps the k entry of v to i_k entry of u . Therefore T has full rank and $\text{rank}(T) = N$ (there are N indices). Hence, T is invertible (full rank matrices are invertible.)

4.5

We see that $u[i], v[i] \in \mathbb{R}, \forall i \in I$. We need to prove that the defined inner product satisfies the three axioms:

Conjugate symmetry:

$$\overline{\langle v, u \rangle}_I = \overline{\sum_{i \in I} u[i]v[i]} = \sum_{i \in I} \overline{u[i]v[i]} = \sum_{i \in I} u[i]v[i] = \langle u, v \rangle_I.$$

Linearity in the first argument:

$$\langle \alpha u, v \rangle = \sum_{i \in I} \alpha u[i]v[i] = \alpha \sum_{i \in I} u[i]v[i] = \alpha \langle u, v \rangle.$$

Let $u_1, u_2 \in \mathbb{R}^I$,

$$\langle u_1 + u_2, v \rangle = \sum_{i \in I} (u_1[i] + u_2[i])v[i] = \sum_{i \in I} u_1[i]v[i] + \sum_{i \in I} u_2[i]v[i] = \langle u_1, v \rangle + \langle u_2, v \rangle.$$

Positive-definiteness:

$$\langle u, u \rangle = \sum_{i \in I} u[i]u[i] = \sum_{i \in I} (u[i])^2 \geq 0$$

Let u be a zero vector, i.e. $u[i] = 0, \forall i \in I$,

$$\langle u, u \rangle = \sum_{i \in I} (u[i])^2 = \sum_{i \in I} 0 = 0.$$

Hence, $\langle u, v \rangle_I := \sum_{i \in I} v[i]u[i]$ is an inner-product in \mathbb{R}^I .

4.6

For an arbitrary $t \in I$, let $e_t = \{\epsilon_\tau\}_{\tau \in I}$, s.t.

$$\epsilon_\tau = \begin{cases} 1, & \tau = t \\ 0, & \text{otherwise.} \end{cases}$$

It is obvious that $\{e_\tau\}$ are linearly independent, as $\sum_\tau \alpha_\tau e_\tau = 0 \Leftrightarrow \alpha_\tau = 0, \forall \tau$. Therefore $\{e_\tau\}$ is a basis. Since $e_t \subset \mathbb{R}^I$, we have

$$\langle u, e_t \rangle_I = \sum_{\tau \in I} u[\tau] \epsilon_\tau = u[t].$$

which satisfies the criteria. Hence, the defined e_t is the standard basis (or reproducing kernel).

5 Computer Exercise

This section include the implementation using Python 2.7. The real code can be found in `main.py`.

5.1

```
1 def create_grid_index(N):
2     """ Create the stereo 2D integer square grid index set
3
4     Args:
5         N: limit for the grid
6
7     Returns:
8         I: set of tuples in the form of {(c,i,j)}, where c in {'L','R'}
9           and 0<=i,j<=N-1
10    """
11    r = range(N)
12    I = {(c,i,j) for c in ['L','R'] for i in r for j in r}
13    return I
14
15 I = create_grid_index(2)
16 print 'Grid index:\n', I, '\n'
```

5.2

```
1 u = Vector(indices=I, valfun=rand)
2 v = Vector(indices=I, valfun=ones)
3 print '2*v + u\n', 2*v+u
4 print 'u + v\n', u+v
5 print '10*u\n', 10*u
```

5.3

```
1 def inner_product(u, v):
2     """ Return the inner product of 2 vectors
3
4     Args:
5         u, v: vectors with identical indices
6
7     Returns:
8         prod: inner product of u and v
9     """
10    assert u.indices == v.indices, 'Vectors have different indices'
11
12    prod = 0.0
13    for idx in u.indices:
14        prod += u[idx] * v[idx]
15    return prod
```



```

16
17
18 def norm(u):
19     """Return the norm of a vector
20
21     Args:
22         u: vector
23
24     Returns:
25         norm of u, defined as sqrt(<u,u>), where <.,.> is the inner
26         product
27         """
28     return np.sqrt(inner_product(u, u))
29
30
31 print '||u|| = ', norm(u)
32 print '||v|| = ', norm(v)
33 print '<u,v> = ', inner_product(u,v), '\n'

```

5.4

```

1 def standart_vec(I, t):
2     """Find a standard basis (or reproducing kernel) of a vector
3
4     Args:
5         I: index set
6         t: arbitrarily chosen index
7
8     Returns:
9         et: standard basis, such that u[t] = <u,et>
10        """
11     assert t in I, 'Undefined index'
12     et = Vector(indices=I, valfun=zeros)
13     et[t] = 1.0
14     return et
15
16
17 t = random.sample(I, 1)[0] # randomly choose an index
18 et = standart_vec(I, t)
19 print 'Standard vector et\n', et
20 print 'u[t] - <u, et> =', u[t] - inner_product(u, et), '\n'

```

Bonus Assume that we do not change the order of index set, the mapping $T : \mathbb{R}^I \rightarrow \mathbb{R}^N$ is implemented as:

```

1 def linear_map(I, inp):
2     """Linear mapping from Vector to list and vice versa
3
4     Args:
5         I: index set
6         inp: input variable, either a Vector or a list
7
8     Returns:

```

```

9         out: if inp is a Vector then out is its corresponding list;
10             if imp is a list then out is its corresponding Vector
11     """
12     if isinstance(inp, Vector):
13         out = []
14         for i in I:
15             out.append(inp[i])
16     elif isinstance(inp, list):
17         out = Vector(indices=I, valfun=zeros)
18         k = 0
19         for i in I:
20             out[i] = inp[k]
21             k += 1
22     else:
23         print 'Not convertible'
24         return
25     return out

```

5.5

```

1 def conv_matrix(p, L):
2     """Compute the convolution matrix of vector p with a vector of
3     length L
4
5     Args:
6         p: array of number
7         L: length of the array that p is convolved with
8
9     Returns:
10         Tp: convolution matrix (Toeplitz matrix)
11     """
12     Tp = np.zeros((len(p)+L-1, L))
13     for j in range(L):
14         Tp[j:j+len(p),j] = p
15     return Tp

```

```

1 p = [1,-2,1]
2 q = [1,2,3,3,2,1]
3 L = len(q)
4 Tp = conv_matrix(p, L)
5 print 'convolution matrix Tp\n', Tp
6 print 'Tp*q - convolve(p, q) = ', np.dot(Tp, q) - np.convolve(p, q)

```