MATH2040A Week 10 Tutorial Notes

1 Orthogonal Complement

Let V be an inner product space. The *orthogonal complement* $S^\perp$ of a set $S \subseteq V$ is the set of all vectors that are orthogonal to each vector in S, that is, $S^\perp = \{ v \in V \mid \forall w \in S, \langle v, w \rangle = 0 \}$.

Basic properties of orthogonal complement are:

- $S^\perp$ is always a subspace
- if $U \subseteq V$ is a subspace, $U \cap U^\perp = \{0\}$
- if $U \subseteq V$ is a *finite dimensional* subspace, $V = U \oplus U^\perp$ (hence, *complement*)
- if $S_1 \subseteq S_2$, then $S_1^\perp \supseteq S_2^\perp$. In particular, $S^\perp = \operatorname{Span}(S)^\perp$
- $(S^\perp)^\perp \supseteq S$, and if U is a *finite dimensional* subspace, $(U^\perp)^\perp = U$ [1]

In particular, if V is finite dimensional, $\dim(V) = \dim(U) + \dim(U^\perp)$. (Compare with dimension theorem)

1.1 Computation of Orthogonal Complement

Given a finite set of vectors $S = \{ w_1, \ldots, w_m \}$ in a finite dimensional space V, how to find $S^\perp$? One way to compute its orthogonal complement is to

1. take a basis $\beta = \{ v_1, \ldots, v_n \}$ of V
2. for a generic vector $v = \sum c_j v_j \in V$, compute $\langle v, w_j \rangle$ for each j. Each yields a linear equation $\langle v, w_j \rangle = 0$ on the coefficients $c_1, \ldots, c_n$
3. solve the linear system of all m equations. $U^\perp$ is then the space of vectors $\sum c_j v_j$ with these solutions as the coefficients

Of course, there are many approaches to do this, and many ways to simplify these computations.

1.2 Optimal Approximation

Let $U \subseteq V$ be a finite dimensional subspace, so $V = U \oplus U^\perp$. As noted in lecture, the following holds:

Theorem 1.1. *Let $v \in V$ with decomposition $v = u + w$ according to the direct sum $V = U \oplus U^\perp$ with $u \in U$, $w \in U^\perp$. Then u is the unique optimal approximation (with respect to the norm) of v in U in the sense that $\| v - u \| = \| w \| < \| v - x \|$ for all $x \in U \setminus \{u\}$.*

Furthermore, if $\{ e_1, \ldots, e_n \}$ is an orthonormal basis of U, then $u = \sum \langle v, e_j \rangle e_j$.

The *orthogonal projection operator* is then $P_U(v) = \sum \langle v, e_j \rangle e_j \in U$ for $v \in V$. It is easy to see that

- $P_U^2 = P_U$ and so is a projection
- (textbook Sec. 6.3 Q9) $(P_U)^* = P_U$
- (Bessel inequality) $\| v \|^2 = \| w \|^2 + \| u \|^2 \geq \| u \|^2 = \sum | \langle v, e_j \rangle |^2$, with equality holds if and only if $v \in U$ (Parseval identity)

[1] Counter-examples exist for *infinite dimensional* spaces.

A typical application of orthogonal projection is to find approximations of a function in various function spaces. For example, consider $V = C^0([0,1])$ equipped with the (usual) inner product $\langle\, f,\, g\,\rangle = \int_0^1 f(t)g(t)\,\mathrm{d}t$, $v(x) = e^x$ for $x \in [0,1]$, and $U = \mathrm{Span}(\,\{\,1, x, x^2\,\}\,)$. The optimal approximation $P_U(v)$ of v in U can be computed with the following steps:

1. take a basis $\beta = \{\,1, x, x^2\,\}$ of U
2. apply Gram–Schmidt process on β to obtain an orthogonal basis $\alpha = \{\,1, x - \frac{1}{2}, x^2 - x + \frac{1}{6}\,\}$
3. construct $u = \frac{\langle\, v,\, 1\,\rangle}{\|\,1\,\|^2} \cdot 1 + \frac{\langle\, v,\, x-1/2\,\rangle}{\|\,x-1/2\,\|^2} \cdot (x - 1/2) + \frac{\langle\, v,\, x^2-x+1/6\,\rangle}{\|\,x^2-x+1/6\,\|^2} \cdot (x^2 - x + 1/6) \approx 0.84x^2 + 0.85x + 1.01$.

The most difficult part in this computation is to handle the orthogonal basis correctly.

2 Adjoint Operator

I was not aware that the lecture has not finished discussing adjoint operators yet. I will talk about adjoint operators in the later tutorial sessions.

3 Exercises

1. (See also textbook Sec. 6.2 Q14)

   Let W_1 and W_2 be subspaces of an inner product space V. Show that $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$. Furthermore, if V is finite dimensional, show that $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

   Solution:

   (a) Since $W_1, W_2 \subseteq W_1 + W_2$, by property of orthogonal complement we have $(W_1 + W_2)^\perp \subseteq W_1^\perp$ and $(W_1 + W_2)^\perp \subseteq W_2^\perp$, so $(W_1 + W_2)^\perp \subseteq W_1^\perp \cap W_2^\perp$.

   Let $v \in W_1^\perp \cap W_2^\perp$, $w \in W_1 + W_2$. Then $v \in W_1^\perp$ and $v \in W_2^\perp$.
   By definition, there exists $w_1 \in W_1$ and $w_2 \in W_2$ such that $w = w_1 + w_2$.
   Then $\langle\, v,\, w_1\,\rangle = \langle\, v,\, w_2\,\rangle = 0$. This implies that $\langle\, v,\, w_1 + w_2\,\rangle = 0$.
   As w is arbitrary, $v \in (W_1 + W_2)^\perp$. As v is arbitrary, $W_1^\perp \cap W_2^\perp \subseteq (W_1 + W_2)^\perp$.

   Combined, $W_1^\perp \cap W_2^\perp = (W_1 + W_2)^\perp$.

   (b) Since V is finite dimensional, $(W_1^\perp)^\perp = W_1$, $(W_2^\perp)^\perp = W_2$, and $((W_1^\perp + W_2^\perp)^\perp)^\perp = W_1^\perp + W_2^\perp$.
   By the previous part, $W_1 \cap W_2 = (W_1^\perp)^\perp \cap (W_2^\perp)^\perp = (W_1^\perp + W_2^\perp)^\perp$, so $(W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp$.

   ### Note

   We now have many pairs of concepts that have some form of duality when coupled together (at least in the case of finite dimensional spaces): subspace sum and intersection, nullspace and range, a subspace and its orthogonal complement, and (in upcoming lectures) a linear map and its adjoint.

2. Let V be an inner product space, and $U, W \subseteq V$ be two finite dimensional subspaces with corresponding orthogonal projection P_U, P_W.
   Show that $P_U P_W = 0$ if and only if $U \perp W$, that is $\langle\, u,\, w\,\rangle = 0$ for all $u \in U, w \in W$.

   Solution: Let $\beta_U = \{\,e_1, \ldots, e_n\,\}$ and $\beta_W = \{\,f_1, \ldots, f_m\,\}$ be orthonormal bases of U, W respectively.

   Suppose $P_U P_W = 0$.
   Let $u \in U, w \in W$. Then there exists scalars $c_1, \ldots, c_n \in F$ such that $u = \sum c_j u_j$.
   As $P_U P_W = 0$, we have $0 = P_U P_W w = P_U w = \sum \langle\, w,\, e_j\,\rangle\, e_j$. Since β_U is a basis, $\langle\, w,\, e_j\,\rangle = 0$ for all j.

This implies $\langle u, w \rangle = \sum c_j \langle e_j, w \rangle = 0$.
As u, w are arbitrary, $U \perp W$.

Suppose $U \perp W$.
Let $v \in V$. Then $P_W v \in W$, so there exists $d_1, \ldots, d_m \in F$ such that $P_W v = \sum d_k f_k$.
As $U \perp W$, we have $\langle f_k, e_j \rangle = 0$ for all j, k.
This implies that $P_U P_W v = P_U(\sum d_k f_k) = \sum_k d_k \sum_j \langle f_k, e_j \rangle e_j = 0$.
As v is arbitrary, $P_U P_W = 0$.

Therefore, $P_U P_W = 0$ if and only if $U \perp W$.

Note

Since $U \perp W$ if and only if $W \perp U$, we have $P_U P_W = 0$ if and only if $P_W P_U = 0$.

3. Let V be a *complex* inner product space, $U \subseteq V$ be a subspace, $v \in V$. Suppose for all $u \in U$, $\langle v, u \rangle + \langle u, v \rangle \leq \langle u, u \rangle$. Show that $v \in U^\perp$.

Solution: Suppose on the contrary that $v \notin U^\perp$. Then there exists $u_0 \in U$ such that $\langle v, u_0 \rangle \neq 0$. In particular, $u_0 \neq 0$.

Let $\langle v, u_0 \rangle = r + im$ with $r, m \in \mathbb{R}$, so r, m are the real and the imaginary part of $\langle v, u_0 \rangle$ respectively.
If $r = 0$, we must have $m \neq 0$, so $iu_0 \in U$ and $\langle v, iu_0 \rangle = -i \langle v, u_0 \rangle = m$ with a nonzero real part. So by considering iu_0 instead, we may always assume that $r \neq 0$.
If $r < 0$, we have $-u_0 \in U$ and $\langle v, -u_0 \rangle = -\langle v, u_0 \rangle = (-r) + i(-m)$ with a positive real part. So by considering $-u_0$ instead, we may always assume that $r > 0$.

Let $\lambda = \frac{r}{\| u_0 \|^2} > 0$. Then $\lambda u_0 \in U$, so by assumption $2\lambda r = \langle v, \lambda u_0 \rangle + \langle \lambda u_0, v \rangle \leq \langle \lambda u_0, \lambda u_0 \rangle = \lambda^2 \| u_0 \|^2 = \lambda r$. Contradiction arises as $\lambda, r > 0$.

This implies that $v \in U^\perp$.

Note

By the proof, we can also see that $\langle u, u \rangle$ in the assumption can be replaced with $c \langle u, u \rangle$ with any given $c > 0$. In another word, if $\frac{\langle v, u \rangle + \langle u, v \rangle}{\langle u, u \rangle}$ has a uniform upper bound on $u \in U \setminus \{0\}$, $v \in U^\perp$ (in which case this quantity is uniformly zero).

4. Let V be an inner product space, $U \subseteq V$ be a finite dimensional subspace with an ordered basis $\beta = \{ v_1, \ldots, v_n \}$ *not necessarily orthonormal*, $x \in V$, $y = P_U(x) \in U$ be the orthogonal projection of x onto U. Let $G \in F^{n \times n}$ be the *Gram matrix / cross-product matrix* of β defined by $G_{jk} = \langle v_k, v_j \rangle$ for each j, k[2].
Find the β-coordinate $[y]_\beta$ of y with the column vector $a = \begin{pmatrix} \langle x, v_1 \rangle & \ldots & \langle x, v_n \rangle \end{pmatrix}^\mathsf{T}$ and the Gram matrix G of β.

Solution: Let $\alpha = \{ e_1, \ldots, e_n \}$ be an orthonormal basis of U.
By definition, $y = P_U(x) = \sum \langle x, e_j \rangle e_j$. In another word, $[y]_\alpha = \begin{pmatrix} \langle x, e_1 \rangle & \ldots & \langle x, e_n \rangle \end{pmatrix}^\mathsf{T}$.
Let $c = [y]_\beta = \begin{pmatrix} c_1 & \ldots & c_n \end{pmatrix}^\mathsf{T}$ with $c_1, \ldots, c_n \in F$.
Then on $R = [\mathrm{Id}_U]^\alpha_\beta$, $[y]_\alpha = [\mathrm{Id}]^\alpha_\beta [y]_\beta = R[y]_\beta$, which implies that $\langle x, e_j \rangle = ([y]_\alpha)_j = (R[y]_\beta)_j =$

[2] Depending on convention, some defines the Gram matrix as $G'_{jk} = \langle v_j, v_k \rangle$, and the two definitions differ by a complex conjugate $G'_{jk} = \overline{G_{jk}}$. For real inner product spaces, the two definitions are identical.

$\sum_k R_{jk} c_k$ for each j.
By definition, $R_{jk} = ([\mathrm{Id}]_\beta^\alpha)_{jk} = \langle\, v_k,\, e_j \,\rangle$ and $v_k = \sum R_{jk} e_j$ for each k. This implies that for each k,

$$a_k = \langle\, x,\, v_k \,\rangle = \sum_j \overline{R_{jk}} \langle\, x,\, e_j \,\rangle = \sum_j \sum_l \overline{R_{jk}} R_{jl} c_l = \sum_l \sum_j (R^*)_{kj} R_{jl} c_l = (R^* R c)_k$$

With direct computation, we also have

$$(R^* R)_{jk} = \sum_l \overline{R_{lj}} R_{lk} = \sum_l \overline{\langle\, v_j,\, e_l \,\rangle} \langle\, v_k,\, e_l \,\rangle = \left\langle\, v_k,\, \sum_l \langle\, v_j,\, e_l \,\rangle e_l \,\right\rangle = \langle\, v_k,\, v_j \,\rangle = G_{jk}$$

so $G = R^* R$, $a = Gc$, $[y]_\beta = c = G^{-1} a$.

Note

If β is orthonormal, by the definition of orthogonal projection we have $y = P_U(x) = \sum \langle\, x,\, v_j \,\rangle v_j$ and so $[y]_\beta = \begin{pmatrix} \langle\, x,\, v_1 \,\rangle & \dots & \langle\, x,\, v_n \,\rangle \end{pmatrix}^\mathsf{T} = a$. This exercise simply states that a correction factor G^{-1} is needed to compensate for the non-orthonormality.

As noted by one of the students in the tutorial session, you can just do the algebra and obtain the same result, without going through the change of coordinate matrix R from β to an orthonormal basis α.

On the other hand, the argument here that uses R should give you some insight on the Gram matrix. For example, if α is obtained from applying Gram–Schmidt process on β, we have $v_k \in \mathrm{Span}(\, \{\, e_1, \ldots, e_k \,\}\,)$ for each k, so $R_{jk} = \langle\, v_k,\, e_j \,\rangle = 0$ whenever $j > k$, and thus R is upper triangular. In this case, $G = R^* R$ is the Cholesky decomposition of G. Furthermore, since R is the change of coordinate matrix, it is invertible (and so is R^*), and thus G is also invertible.

We can use the Gram matrix on other computations as well. Another example is the following result:

Lemma. *Let V be finite dimensional with basis $\beta = \{\, v_1, \ldots, v_n \,\}$ and corresponding Gram matrix G, and $A \in F^{n\times m}$, $U = \mathrm{Span}(\, \{\, u_1, \ldots, u_m \,\}\,)$ with $u_j = \sum A_{kj} v_k$ for each j.
Then $\left\{\, [v]_\beta \;\middle|\; v \in U^\perp \,\right\} = \mathsf{N}\,(\, A^* G \,)$.*

The proof of this lemma is a simple computation. There are a few more examples concerning Riesz vector and adjoint operators, which we will discuss in later tutorial sessions.

While using Gram matrix (sometimes) allows circumventing the need of explicitly computing for an orthonormal basis (e.g. by Gram–Schmidt process), evaluating G and G^{-1} can be computationally tough, and having an orthonormal basis is usually desirable, so it may be better (and occasionally easier) to just construct an orthonormal basis directly.