Week 12, November 15

Seewoo Lee

1 Discussion notes

- Summary of Chapter 6, 7: almost all the facts that you learned in Math 54 about the standard inner product also hold for general inner products.
- Theorems that you need to know:
  - 6.12: Pythagorean theorem: for orthogonal u, v, $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.
  - 6.14: Cauchy–Schwarz inequality, $|\langle u, v\rangle| \le \|u\|\|v\|$. Equality holds when u, v are parallel, i.e. $u = \lambda v$ for some $\lambda \in F$.
  - 6.17: Triangle inequality, $\|u+v\| \le \|u\| + \|v\|$. This essentially tells you that any inner product defines a *distance* on the vector space (see below).
  - 6.21: Parallelogram equality.
  - 6.32: Gram-Schmidt: make "orthonormal" basis out of arbitrary basis.
  - 6.38: Schur's theorem, any operator on a complex inner product space has an upper-triangular matrix with respect to some orthonormal basis.
  - 6.42: Riesz representation theorem, any linear functional $\varphi : V \to F$ on a finite dimensional inner product space is a form of inner product, i.e. there exists $v \in V$ such that $\varphi(u) = \langle u, v\rangle$ for all $u \in V$.
- Inner product over $F = \mathbb{R}$ is a *non-degenerate symmetric bilinear form*. When $F = \mathbb{C}$, inner product is *not* symmetric, but *hermitian*: $\langle u, v\rangle = \overline{\langle v, u\rangle}$. Especially, it is linear on the first argument but not for the second argument: $\langle u, cv\rangle = \overline{c}\langle u, v\rangle$.

  Given inner product, it always defines a *distance* on V:

  $$d(u, v) := \|u - v\| = \langle u - v, u - v\rangle^{1/2}.$$

  (the reason is triangle inequality, Theorem 6.17). It means that you can say *how close* to vectors are. For example, you can even say about how close two polynomials (or more generally some continuous functions) are. Note that there are a lot of possible distances on the space of polynomials, but only few of them arise from inner products. In fact, parallelogram equality (Theorem 6.21) tells you how to test this. (See also additional problem 3.)

- When there is an inner product on a vector space, you can always find *orthonormal* basis, not just a basis. Having orthonormal basis is good for the following reason:
  - Assume that a basis $\beta = (v_1, \ldots, v_n)$ is given. To express an arbitrary vector $v \in V$ as a linear combination of v_i's ($v = \sum_{i=1}^n c_i v_i$), one needs to solve a system of linear equation. But when β is an orthonormal basis, we can easily find c_i: in fact, we have

    $$v = \langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + \cdots + \langle v, v_n \rangle v_n.$$

    (If you are a CS major, think about computational complexity for these two approaches to find coordinates.)
  - Now, consider the following situation: let V be a (possibly infinite dimensional) vector space and $W \subset V$ be a subspace with a *orthonormal* basis $\beta = (w_1, \ldots, w_m)$. We can ask the following question: what is the vector $w^* \in W$ that is *closest* to v among the vectors in w (in other words, $\|v - w^*\| \leq \|v - w\|$ for all $w \in W$)? Since $w \in W$, we can write it uniquely as $w = \sum_{i=1}^n c_i w_i$, and the goal is to find the (optimal) coefficients c_i's. Using the fact that β is an orthonormal basis, we get a very nice answer for this:

    $$w^* = \langle v, w_1 \rangle w_1 + \langle v, w_2 \rangle w_2 + \cdots + \langle v, w_n \rangle w_n.$$

    Why this matters? If we consider the case when V is a space of continuous functions on some interval $[a, b]$ and W is a subspace of *special functions* (for example, polynomial or trigonometric functions), than the above question is equivalent to find a function $g(x) \in W$ that gives a best approximation of given $f(x) \in V$, i.e. minimizing the distance

    $$\|f - g\| = \left(\int_a^b (f(x) - g(x))^2 \mathrm{d}x \right)^{1/2}.$$

    When we have an orthonormal basis $(g_1, \ldots, g_n)$ of W, then the answer is given by the previous formula, which is $f \approx g^* = \sum_{i=1}^n c_i g_i$ where

    $$c_i = \langle f, g_i \rangle = \int_a^b f(x) g_i(x) \mathrm{d}x.$$

For example, Exercise 6B.4 tells you that, when $[a, b] = [-\pi, \pi]$, the trigonometric functions

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots$$

form an orthonormal basis of the space of trigonometric functions (of the form $\sum_{i=0}^{n} a_i \cos(ix) + \sum_{j=1}^{n} b_j \sin(jx)$, and one can use it to approximate a given continuous function with trigonometric functions on $[-\pi, \pi]$. There's a whole theory for such approximations, which is called *Fourier theory*. One can also consider the case when $W = \mathcal{P}_n(\mathbb{R})$ - see the additional problem 2.

- (6A.5, 6A.21, 6A.26) Using the basic but important identity: $\|u + v\|^2 = \langle u + v, u + v\rangle = \|u\|^2 + \langle u, v\rangle + \langle v, u\rangle + \|v\|^2$.
- (6A.6, 6A.14) Cauchy–Schwarz applications. Caution: when you are dealing with complex vector space, we have $\langle u, v\rangle = \overline{\langle v, u\rangle}$, but not $\langle u, v\rangle = \langle v, u\rangle$.
- (6B.8, 6B.11) 6B.8 is very important! You can do 6B.11 manually by setting $p(x) = ax^2 + bx + c$ and $q(x) = dx^2 + ex + f$, and find d, e, f such that the equation holds for *any* a, b, c. But I strongly recommend you to use the basis you found in 6B.8. (In fact, I recommend you to do in both ways, and realize that using 6B.8 is much simpler.)
- (6B.18) Proof uses Riesz representation theorem. But also think about its geometric meaning: for a fixed u_i, how the set of vectors v satisfying $\langle v, u_i\rangle = 1$ looks like? Try when $\dim V = m = 2$.

2 Problems

- Recommended problems: 6A.5, 6A.14, 6B.3, 6B.4, 6B.8 (must!!), 6B.11, 6B.19
- Additional problems:

  1. (More general inner product) Let $w : [-1, 1] \to \mathbb{R}$ be a continous function which is strictly positive on $[-1, 1]$. Define a map $\langle -, -\rangle_w : \mathcal{P}(\mathbb{R}) \times \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ via
  $$\langle p, q\rangle_w := \int_{-1}^{1} p(x)q(x)w(x)\mathrm{d}x.$$
  Prove that $\langle -, -\rangle_w$ is an inner product.

2. (Legendre polynomial)

   (a) Find monic polynomials P_0, P_1, P_2, P_3 such that

   - $\deg P_n(x) = n$,
   - $P_n(1) = 1$,
   - they form an orthogonal basis (not necessarily orthonormal) of $\mathcal{P}_3(\mathbb{R})$ under the inner product

   $$\langle p, q\rangle := \int_{-1}^{1} p(x)q(x)\mathrm{d}x.$$

   Answers can be found here.

   (b) Using the polynomials above, find a degree ≤ 3 polynomial $p(x)$ that minimizes

   $$\|p(x) - \sin(\pi x)\| := \left(\int_{-1}^{1} (p(x) - \sin(\pi x))^2\right)^{\frac{1}{2}} \mathrm{d}x.$$

3. (Distance function that does not come from inner products) Consider the function $d : \mathcal{P}_n(\mathbb{R}) \times \mathcal{P}_n(\mathbb{R}) \to \mathbb{R}$ defined as

   $$d(p, q) := \int_{-1}^{1} |p(x) - q(x)|\mathrm{d}x.$$

   (a) Prove that the above d satisfies the triangle inequality: for any $p, q, r \in \mathcal{P}_n(\mathbb{R})$, $d(p, q) + d(q, r) \geq d(p, r)$.

   (b) Prove that there is *no* inner product $\langle -, -\rangle$ on $\mathcal{P}_n(\mathbb{R})$ that gives $\langle p - q, p - q\rangle^{1/2} = d(p - q)$ for all p, q. (Hint: proof by contradiction, and use the parallelogram identity).