Math 2270-1

Notes of 11/18/19

6.8 Applications of Inner Product Spaces

- Recall our **Definition** (p. 378, textbook): An **inner product** on a vector space V is a function that, to each pair of vectors $\mathbf{u}$ and $\mathbf{v}$ in V, associates a real number $<\mathbf{u},\mathbf{v}>$ and satisfies the following axioms, for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V and all scalars c:

1. $<\mathbf{u},\mathbf{v}>=<\mathbf{v},\mathbf{u}>$.

2. $<\mathbf{u}+\mathbf{v},\mathbf{w}>=<\mathbf{u},\mathbf{w}>+<\mathbf{v},\mathbf{w}>$

3. $<c\mathbf{u},\mathbf{v}>=c<\mathbf{u},\mathbf{v}>$

4. $<\mathbf{u},\mathbf{u}>\geq 0$ and $<\mathbf{u},\mathbf{u}>=0$ if and only if $\mathbf{u}=\mathbf{0}$.

A Vector space with an inner product is called an **inner product space**.

- Associated with an inner product space is the norm

$$\|\mathbf{u}\|=\sqrt{<\mathbf{u},\mathbf{u}>},$$

the concept of orthogonality

$$<\mathbf{u},\mathbf{v}>=0,$$

the Pythagorean Theorem, Projection, Least Squares, the Cauchy-Schwarz Inequality, and the Triangle Inequality.

- We'll look at two applications, one quickly, and one in more detail.

Weighted Least Squares

- Suppose A is an $m \times n$ matrix with $m > n$ and $\mathbf{b}$ is a vector in $\mathbb{R}^m$. Then the linear system

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

  is overdetermined and (usually) does not have a solution. In that case it is often useful to solve instead the Least Squares Problem

$$\|A\mathbf{x} - \mathbf{b}\|^2 = (A\mathbf{x} - \mathbf{b}, A\mathbf{x} - \mathbf{b}) = \min .$$

- We can find the least squares solution $\mathbf{x}$ by for example solving the **normal equations**

$$A^T A\mathbf{x} = A^T \mathbf{b}.$$

- In this approach all equations have the same weight and contribute equally to $\|A\mathbf{x} - \mathbf{b}\|$.
- But suppose that some of the equations in (1) are more important than others. It may be more important to get close agreement in the first equation than in all the others, for example.
- Weighted Least Squares can help in that situation. Define positive weights

$$w_i > 0, \quad i = 1, \ldots, m.$$

Also define the inner product

$$< u, v >= \sum_{i=1}^{m} w_i u_i v_i.$$

- (On Wednesday we considered the example where m is 2 and the weights are 4 and 5.)
- With the diagonal weight matrix

$$W = \begin{bmatrix} w_1 & 0 & 0 & \dots & 0 \\ 0 & w_2 & 0 & \dots & 0 \\ 0 & 0 & w_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & w_m \end{bmatrix}$$

we can write our new inner product as

$$< \mathbf{u}, \mathbf{v} >= (W\mathbf{u}, \mathbf{v}) = (\mathbf{u}, W\mathbf{v}).$$

- The associated norm then is

$$\|\mathbf{x}\|_w = \sqrt{(\mathbf{x}, W\mathbf{x})} = \sqrt{\sum_{i=1}^{m} w_i x_i^2}.$$

- As we saw when we first discussed Least Squares, the residual $A\mathbf{x} - \mathbf{b}$ must be orthogonal to every vector in the column space of A. Writing

$$A = [\, \mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n \,]$$

we get the requirement that

$$< \mathbf{a}_j, A\mathbf{x} - \mathbf{b} >= (\mathbf{a}_j, WA\mathbf{x} - W\mathbf{b}) = 0.$$

- In matrix form this turns into the modified normal equations

$$A^T W A\mathbf{x} = A^T W\mathbf{b}.$$

Fourier Series

- One of the most widely used inner products on a function space V is

$$< f, g >= \int_0^1 f(x)g(x)\mathrm{d}x. \tag{2}$$

- The textbook assumes that V is the linear space of continuous functions, but actually the space could be larger. For example, jump discontinuities would not matter for the integral. A more typical choice of V would be the set of all functions f that are **square integrable**, i.e.,

$$\int_0^1 f^2(x)\mathrm{d}x$$

  is well defined and finite.

- It is easy to verify that (2) defines an inner product (exercise).

Note that (2) looks just like our ordinary inner product, except that the "sum" is an integral. It's one of the major simplifying principles of analysis that integrals behave like sums.

- Suppose we want to approximate a function f by a linear combination of some basis functions

$$\phi_1, \phi_2, \ldots, \phi_n$$

such that

$$\|f-\sum_{i=1}^n \alpha_i\phi_i\|^2 =< f-\sum_{i=1}^n \alpha_i\phi_i, f-\sum_{i=1}^n \alpha_i\phi_i >= \min.$$

- As before we get the requirement that the residual be orthogonal to the approximating space, i.e.,

$$< \phi_i, f - \sum_{j=1}^n \alpha_j\phi_j >= 0, \qquad i = 1, \dots, n.$$

- This gives the linear system

$$A\mathbf{a} = \mathbf{b}$$

where

$$A = \begin{bmatrix} <\phi_1,\phi_1> & <\phi_1,\phi_2> & <\phi_1,\phi_3> & \dots & <\phi_1,\phi_n> \\ <\phi_2,\phi_1> & <\phi_2,\phi_2> & <\phi_2,\phi_3> & \dots & <\phi_2,\phi_n> \\ <\phi_3,\phi_1> & <\phi_3,\phi_2> & <\phi_3,\phi_3> & \dots & <\phi_3,\phi_n> \\ \vdots & \vdots & \vdots & & \vdots \\ <\phi_n,\phi_1> & <\phi_n,\phi_2> & <\phi_n,\phi_3> & \dots & <\phi_n,\phi_n> \end{bmatrix},$$

$$\mathbf{a} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} <\phi_1, f> \\ <\phi_2, f> \\ <\phi_3, f> \\ \vdots \\ <\phi_n, f> \end{bmatrix}.$$

- Example: Compute A for the case that

$$\phi_i(x) = x^{i-1} \quad \text{and} \quad < f, g >= \int_0^1 f(x)g(x)\mathrm{d}x.$$

- The basis functions can be chosen suitably. It would, of course, be particularly nice if they were orthogonal. In that case, the coefficient matrix A would be diagonal.

- Today's main subject is the approximation of periodic functions by periodic functions, leading to the large subject of **Fourier Series** and **Fourier Analysis.**

- **Jean-Baptiste Joseph Fourier, 1768–1830**

- Periodic functions occur in many applications!

- A function f is p-**periodic** (or **periodic of period** p, or **periodic of periodicity** p) if

  $$f(t+p) = f(t)$$

  for all t in the domain of f.

- We usually assume that $p > 0$. Negative p also work, and $p = 0$ is sometimes included for generality. (In that case any function is considered 0-periodic.)

- Example: The ordinary trig functions are 2π-periodic.

If f is periodic of period p then it is also periodic of period pk for any non-zero integer k.

- For example $\sin x$ is 2π-periodic, but also 2000π-periodic.

- By the way, the tangent function is π-periodic.

- We will assume that our function is 2π periodic.
- If we have any period $p \neq 0$ we can make the function 2π-periodic by a linear change of variables:
$$x = \frac{2\pi t}{p}$$
- So let's assume, without loss of generality, that $p = 2\pi$, i.e.,
$$\boxed{f(t + 2\pi) = f(t)}$$
for all real numbers t.
- So what should we use as basis functions?
- It seems natural to choose
$$\phi_i(t) = 1/2, \cos t, \sin t, \cos 2t, \sin 2t, \ldots$$
- The reason for choosing the constant $1/2$, rather than 1, say, as our first basis function will become apparent later. It simplifies the formulas for the coefficients.
- This leads to the choice
$$s(t) = F_n(t) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos kt + b_k \sin kt\right) \tag{3}$$

- This is a partial sum of the **Fourier Series**

$$s(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \big(a_k \cos kt + b_k \sin kt\big) \qquad (4)$$

- We define the **Fourier coefficients** a_k, b_k, by the least squares requirement

$$\int_{-\pi}^{\pi} \big(f(t) - F_n(t)\big)^2 \mathrm{d}t = \min \qquad (5)$$

- The inner product underlying this approach is, of course,

$$< f, g >= \int_{-\pi}^{\pi} f(t)g(t)\mathrm{d}t.$$

- As we discussed, we have to solve the linear system

$$Ac = b$$

where

$$a_{ij} = (\phi_i, \phi_j),$$

$$c = [a_0, a_1, b_1, a_2, b_2, \ldots, a_n, c_n]^T,$$

and

$$b = [(f, \phi_i)].$$

Remarkably, the basis functions ϕ_i are **orthogonal** with respect to our chosen inner product.

- In fact, we get

$$\int_{-\pi}^{\pi} \sin nx \cos mx = \int_{-\pi}^{\pi} \sin nx = \int_{-\pi}^{\pi} \cos nx = 0$$

$$\int_{-\pi}^{\pi} \sin nx \sin mx = \int_{-\pi}^{\pi} \cos nx \cos mx = \begin{cases} \pi & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases}$$

- If time allows, let's do some of the calculations:

- So we get the linear system

$$\begin{bmatrix} \pi & & & & & \\ & \pi & & & 0 & \\ & & \pi & & & \\ & & & \pi & & \\ & 0 & & & \ddots & \\ & & & & & \pi \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ b_1 \\ \vdots \\ a_n \\ b_n \end{bmatrix} = \begin{bmatrix} \int_{-\pi}^{\pi} \frac{1}{2} f(t) \mathrm{d}t \\ f(t) \cos t \mathrm{d}t \\ \vdots \\ \int_{-\pi}^{\pi} f(t) \sin nt \mathrm{d}t \end{bmatrix}.$$

- So our linear system is actually diagonal!

We did what came naturally, and we obtained a diagonal linear system!

- Solving the linear system gives

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(s) \cos ks \mathrm{d}s, \qquad k = 0, 1, 2, \ldots$$
$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(s) \sin ks \mathrm{d}s, \qquad k = 1, 2, 3, \ldots$$

and the Fourier approximation

$$\begin{aligned} f(t) &\approx F_n(t) \\ &= \frac{\int_{-\pi}^{\pi} f(s) \mathrm{d}s}{2\pi} + \\ &\quad + \frac{1}{\pi} \sum_{k=1}^{n} \left(\int_{-\pi}^{\pi} f(s) \cos ks \mathrm{d}s \cos kt + \int_{-\pi}^{\pi} f(s) \sin ks \mathrm{d}s \sin kt \right) \end{aligned}$$

A Sawtooth Function

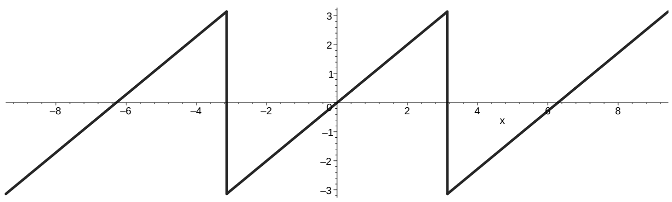

Figure 1. A Sawtooth Function.

- Let's do an example. Consider the sawtooth function f shown in Figure 1. f is 2π-periodic and

$$f(t) = t$$

  in the interval $(-\pi, \pi)$.

- We can easily compute the coefficients of our Fourier Series: the cosine terms are zero because the integrand is odd:

$$a_k = \int_{-\pi}^{\pi} t \cos kt \mathrm{d}t = 0$$

- The sine terms are more complicated. However, integration by parts is straightforward and gives:

$$
\begin{aligned}
b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi}\underbrace{t}_{u}\underbrace{\sin kt}_{v'}\,\mathrm{d}t \\
&= \frac{1}{\pi}\left[-\frac{t}{k}\cos kt\Big|_{-\pi}^{\pi} + \underbrace{\int_{-\pi}^{\pi}\frac{1}{k}\cos kt\mathrm{d}t}_{=0}\right] \\
&= \frac{1}{\pi k}\left[-\pi\cos(k\pi) - \pi\cos(-k\pi)\right] \\
&= \frac{-2\cos k\pi}{k} \\
&= \begin{cases} \frac{2}{k} & \text{if } k \text{ is odd} \\ -\frac{2}{k} & \text{if } k \text{ is even} \end{cases}
\end{aligned}
$$

- The Fourier series of our sawtooth function is, therefore,

$$
f(t) = 2(\sin t - \frac{1}{2}\sin 2t + \frac{1}{3}\sin 3t - \frac{1}{4}\sin 4t + \ldots)
$$

- Figures 2 through 5 show the function f and its Fourier approximation for $n = 5, 10, 20, 100$.
- Here are some observations:
  - The function f is odd, and the Fourier approximation is also odd for each n. It's obvious that this must happen! (why?)
  - The approximated function is discontinuous. For finite n, the Fourier approximation is infinitely differentiable. This is bound

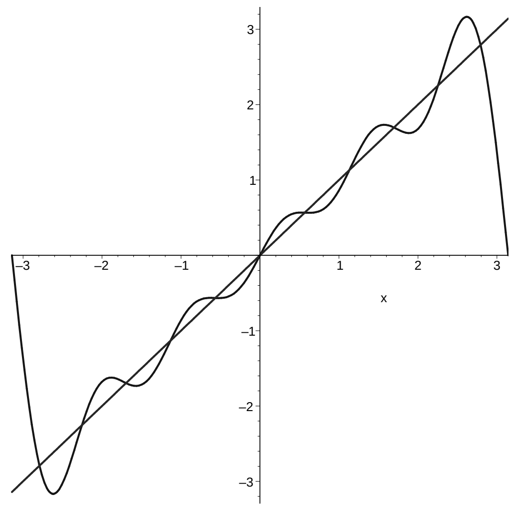

Figure 2. Fourier approximation of sawtooth function, $n = 5$.

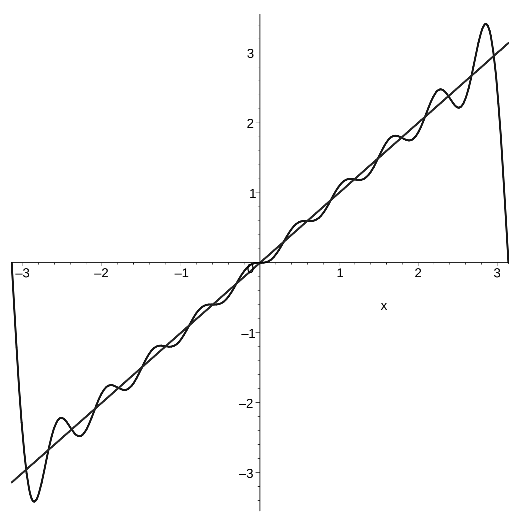

Figure 3. Fourier approximation of sawtooth function, $n = 10$.

to cause some kind of problem. It is apparent from the graphs that there is some oscillation at the discontinuity. It turns out the vertical extent of that oscillation remains constant, while the horizontal extent shrinks, as n goes to infinity. This is

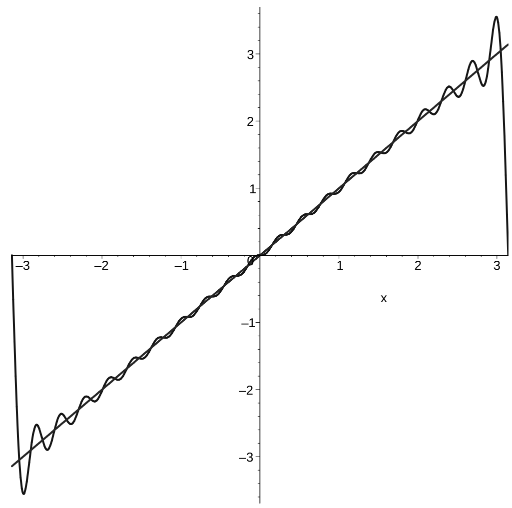

Figure 4. Fourier approximation of sawtooth function, $n = 20$.

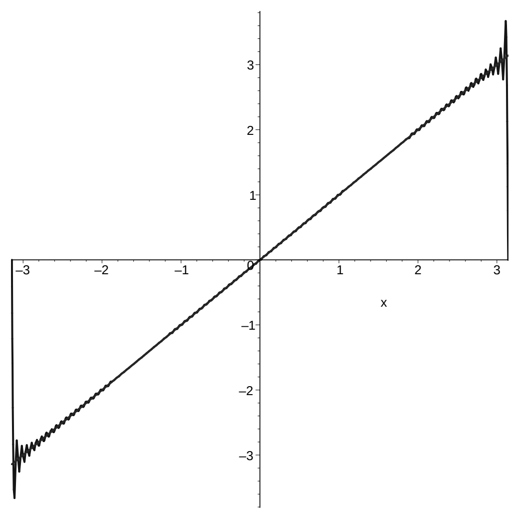

Figure 5. Fourier approximation of sawtooth function, $n = 100$.

the contents of the celebrated **Gibbs Phenomenon**.

– Another effect of the discontinuities in f is that the coefficients go to zero only slowly as n goes to infinity.

Another Example

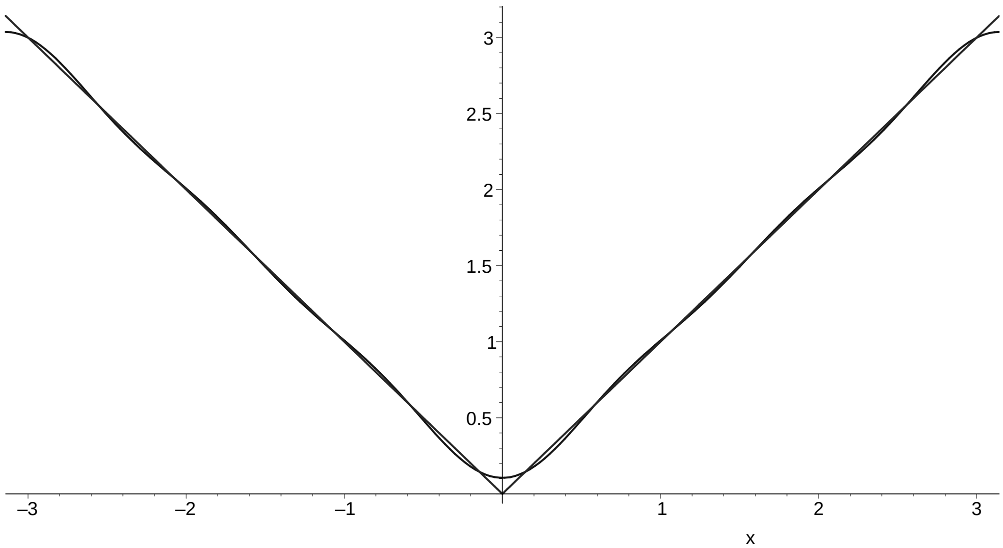

Figure 6. Fourier approximation of triangular wave, $n = 5$.

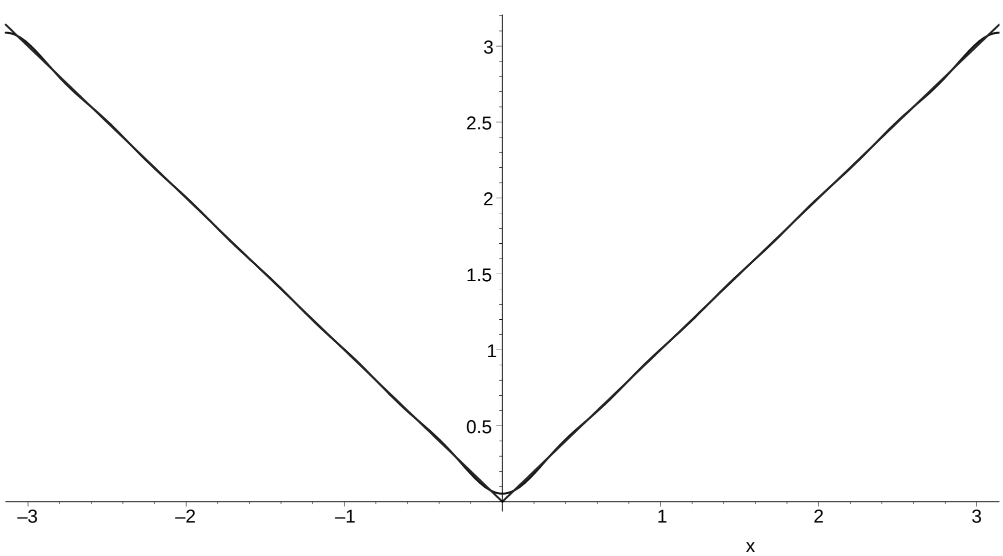

Figure 7. Fourier approximation of triangular wave, $n = 10$.

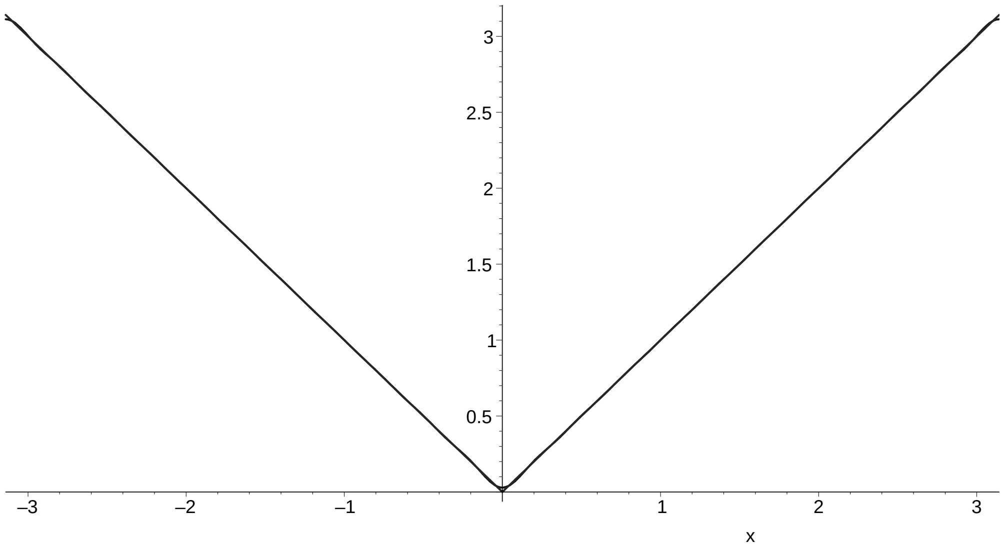

Figure 8. Fourier approximation of triangular wave, $n = 20$.

- Consider the function g that is 2π periodic and that in the interval $[-\pi, \pi]$ satisfies $g(t) = |t|$. It's graph is a **triangular wave** and its Fourier Series is (exercise):

$$
\begin{aligned}
g(t) &= \frac{\pi}{2} - \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\cos(2k-1)x}{(2k-1)^2} \\
&= \frac{\pi}{2} - \frac{4}{\pi} \left(\cos x + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \dots \right)
\end{aligned}
$$

Figures 6 through 8 show the function g and its Fourier approximations for $n = 5, 10, 20$.

The coefficients approach zero much faster, and the Gibbs Phenomenon is absent.