Math 5600 Spring 2021 Notes of 2/22/21

• Why interpolate or approximate?
  – We may have exact or approximate data at certain points and want to estimate function values at other points.
  – A computer can directly evaluate only rational functions (based on the 4 basic arithmetic operations). Other functions are evaluated by evaluating an approximation that is accurate within the given computer accuracy.
  – Evaluating functions, integrals, or derivatives, solving DEs, all based on interpolation or approximation one way or the other.
  – Another application area is the design of surfaces, like the shape of a vehicle.
  – Computer Graphics.

Continuous Least Squares

• So far we have approximated functions by interpolation.
• We get a linear combination of basis functions:

\[ f(x) \approx \sum_{i=0}^{n} \alpha_i \phi_i(x) = p(x). \]

• Interpolation means that the coefficients \( \alpha_i \) are determined by the requirement that

\[ p(x_i) = y_i, \quad i = 0, \ldots, n. \]

• There are alternatives!
• The basis functions need not be polynomial either.
• We can pick them suitable for the problem, for example, they might be periodic or exponential, they might have singularities (to model explosions, say), or they might have a certain asymptotic behavior.
• How do we quantify $f(x) \approx p(x)$?

• There are several, indeed, infinitely many, ways of quantifying the quality of the approximation.

• For example

\[
\int_a^b |f(x) - p(x)| \, dx = \min
\]

\[
\max_{a \leq x \leq b} |f(x) - p(x)| = \min
\]

\[
\int_a^b (f(x) - p(x))^2 \, dx = \min
\]

• The three approaches are illustrated in the last three problems of the current homework.

• The last option is the most popular and is referred to as **continuous linear Least Squares**.
  
  – **continuous** means we have infinitely many points (the graph of $f$) instead of finitely many points. (The alternative to continuous least squares is **discrete** least squares.)
  
  – **linear** means that—as we shall soon see—the problem leads to a linear system of equations, as opposed to a nonlinear one. It does not mean that the approximating function $p$ is linear!
  
  – **squares** because we integrate the square of something.
  
  – **least** because we minimize something.

• Here are some examples for problems leading to **nonlinear** least squares:

\[
f(x) \approx \alpha \sin(\beta x + \gamma)
\]

\[
f(x) \approx \alpha e^{\beta x}
\]

\[
f(x) \approx \frac{\alpha_0 + \alpha_1 x}{\beta_0 + x}
\]
Let's see how to solve the linear problem.

Start with an example:

\[ \phi_i(x) = x^i, \quad \int_0^1 (f(x) - \sum_{j=0}^{n} \alpha_j x^j)^2 \, dx = \min. \]

\[ F(d_0, \ldots, d_n) = \int_0^1 (f(x) - \sum_{j=0}^{n} \alpha_j x^j)^2 \, dx = \min \]

\[ \frac{\partial}{\partial d_i} F = -\int_0^1 2 (f(x) - \sum_{j=0}^{n} \alpha_j x^j) x^i \, dx = 0 \]

\[ \sum_{j=0}^{n} \alpha_j \int_0^1 x^j x^i \, dx = \int_0^1 f(x) x^i \, dx \]

\[ \int_0^1 x^j x^i \, dx = \frac{x^{i+j+1}}{i+j+1} \bigg|_0^1 = \frac{1}{i+j+1} \]

Math 5600 Spring 2021 Notes of 2/22/21 page 3
• The \((n + 1) \times (n + 1)\) matrix \(H_n\) of the linear system

\[
\begin{bmatrix}
1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n+1} \\
\frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+2} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\frac{1}{n+1} & \frac{1}{n+2} & \frac{1}{n+3} & \cdots & \frac{1}{2n+1}
\end{bmatrix}
\begin{bmatrix}
\alpha_0 \\
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix}
= 
\begin{bmatrix}
\int_0^1 f(x)dx \\
\int_0^1 xf(x)dx \\
\vdots \\
\int_0^1 x^n f(x)dx
\end{bmatrix}
\]

is called the **Hilbert Matrix**

• Let’s list some of its properties:

- **Symmetric**

\[
\begin{bmatrix}
1 & 1 \\
1 & 1 \\
\end{bmatrix}
= \begin{bmatrix}
1 & -1 \\
-1 & 1
\end{bmatrix} \begin{bmatrix}
1 & 1 \\
1 & 3
\end{bmatrix}
= -6
\]

\(x^T A x < 0\)

- \(p(x) = \sum_{j=0}^{n} \alpha_j x^j \neq 0\)

\[
0 < \int_0^1 p^2(x)dx = a^T H_n a
\]

\(a = \begin{bmatrix}
\alpha_0 \\
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix}\)

\(\Rightarrow\) **minimum**

Math 5600 Spring 2021  Notes of 2/22/21  page 4
Let’s redo the problem in general

\[ F(\alpha_0, \alpha_1, \ldots, \alpha_n) = \int_a^b \left( f(x) - \sum_{j=0}^n \alpha_j \phi_j(x) \right)^2 dx = \min. \]

Differentiating gives the equations

\[ \frac{\partial F}{\partial \alpha_i} = -2 \int_a^b \left( f(x) - \sum_{j=0}^n \alpha_j \phi_j(x) \right) \phi_i(x) dx = 0, \quad i = 0, \ldots, n \]

which can be rewritten as

\[ \sum_{j=0}^n \alpha_j \int_a^b \phi_i(x) \phi_j(x) dx = \int_a^b f(x) \phi_i(x) dx, \quad i = 0, \ldots, n \]

We don’t have to write our approximating polynomial in its power form. For example, choosing

\[ \phi_0(x) = 1 \]
\[ \phi_1(x) = 2x - 1 \]
\[ \phi_2(x) = 6x^2 - 6x + 1 \]

gives the matrix

\[ \left[ \int_0^1 \phi_i(x) \phi_j(x) dx \right]_{i,j=0,1,2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{5} \end{bmatrix} \]

That’s much nicer!

Does \[ \int_a^b \phi_i(x) \phi_j(x) dx \] remind you of anything?
• One of the great simplifying principles of Analysis is that Integrals behave like sums (unless they don’t).

• An inner product define on a linear space $V$ is a function that associates a number $(f, g)$ with any two vectors (functions) $f, g$ in $V$ such that the following properties hold:

\[(i) \quad (f, g) = (g, f)\]
\[(ii) \quad (\alpha f, g) = \alpha (f, g)\]
\[(iii) \quad (f + g, h) = (f, h) + (g, h)\]
\[(iv) \quad (f, f) \geq 0\]
\[(v) \quad (f, f) = 0 \implies f = 0\]

for all scalars $\alpha$ and vectors $f, g$ in $V$.

• A vector space with an inner product is called an inner product space.

• An example of an inner product space is the set of all continuous functions on the interval $[a, b]$ with the inner product

\[(f, g) = \int_{a}^{b} f(x)g(x)dx.\]

• Other examples for inner products.

\[(f, g) = \int_{a}^{b} w(x)f(x)g(x)dx, \quad w(x) > 0\]
\[(f, g) = \int_{a}^{b} f(x)g(x) + f'(x)g'(x)dx\]
\[(f, g) = \int_{a}^{b} f(x)g(x)dx + f(c)g(c)\]

• Associated with an inner product is a norm

\[\|f\| = \sqrt{(f, f)}.\]
• The associated linear approximation problem is

\[ \| f - \sum_{j=0}^{n} \alpha_j \phi_j(x) \|^2 = \min \]

• Minimizing that norm gives rise to a linear system. Let’s redo our problem in general.

• We want to solve

\[
F(\alpha_0, \ldots, \alpha_n) = \left( f - \sum_{j=0}^{n} \alpha_j \phi_j(x), f - \sum_{j=0}^{n} \alpha_j \phi_j(x) \right) = (f, f) - 2 \sum_{j=0}^{n} \alpha_j (\phi_j, f) + \sum_{i=0}^{n} \sum_{j=0}^{n} \alpha_i \alpha_j (\phi_i, \phi_j) = \min
\]

• Differentiating and setting to zero, as usual, gives

\[
\frac{\partial F}{\partial \alpha_i} = -2(\phi_i, f) + 2 \sum_{j=0}^{n} \alpha_j (\phi_j, \phi_j) = 0, \quad i = 0, 1, 2, \ldots, n.
\]

• This is the same system as before.

\[
\begin{bmatrix}
(\phi_0, \phi_0) & (\phi_0, \phi_1) & \ldots & (\phi_0, \phi_n) \\
(\phi_1, \phi_0) & (\phi_1, \phi_1) & \ldots & (\phi_1, \phi_n) \\
\vdots & \vdots & \ddots & \vdots \\
(\phi_n, \phi_0) & (\phi_n, \phi_1) & \ldots & (\phi_n, \phi_n)
\end{bmatrix}
\begin{bmatrix}
\alpha_0 \\
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix}
= \begin{bmatrix}
(\phi_0, f) \\
(\phi_1, f) \\
\vdots \\
(\phi_n, f)
\end{bmatrix}
\]

• Properties of that system?

symmetric

positive definite
It would be nice if the system was diagonal too!

Two vectors $f$ and $g$ in an inner product space are said to be orthogonal if

$$(f, g) = 0.$$ 

We want an orthogonal basis!

How can we obtain one?

- even better
  
  $$
  (\phi_i, \phi_j) = 0 \quad i \neq j \\
  (\phi_i, \phi_i) = 1
  $$

- even $A = I$

Gram-Schmidt

$$(\phi_i, \phi_j) = \delta_{i,j}$$

$$
\Rightarrow \quad p = \sum_{i=0}^{n} (f, \phi_i) \phi_i
$$