MATH 304

Linear Algebra

Lecture 26: Orthogonal polynomials. Review for the final exam.

Problem. Approximate the function $f(x) = e^x$ on the interval $[-1, 1]$ by a quadratic polynomial.

The best approximation would be a polynomial $p(x)$ that minimizes the distance relative to the uniform norm:

$$\|f - p\|_\infty = \max_{|x|\le 1} |f(x) - p(x)|.$$

However there is no analytic way to find such a polynomial. Another approach is to find a *"least squares"* approximation that minimizes the integral norm

$$\|f - p\|_2 = \left(\int_{-1}^{1} |f(x) - p(x)|^2 \, dx \right)^{1/2}.$$

The norm $\|\cdot\|_2$ is induced by the inner product

$$\langle g, h\rangle = \int_{-1}^{1} g(x)h(x)\,dx.$$

Therefore $\|f - p\|_2$ is minimal if p is the orthogonal projection of the function f on the subspace $\mathcal{P}_3$ of quadratic polynomials.

Suppose that p_0, p_1, p_2 is an orthogonal basis for $\mathcal{P}_3$. Then

$$p(x) = \frac{\langle f, p_0\rangle}{\langle p_0, p_0\rangle} p_0(x) + \frac{\langle f, p_1\rangle}{\langle p_1, p_1\rangle} p_1(x) + \frac{\langle f, p_2\rangle}{\langle p_2, p_2\rangle} p_2(x).$$

Orthogonal polynomials

$\mathcal{P}$: the vector space of all polynomials with real coefficients: $p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$.
Basis for $\mathcal{P}$: $1, x, x^2, \ldots, x^n, \ldots$

Suppose that $\mathcal{P}$ is endowed with an inner product.

Definition. **Orthogonal polynomials** (relative to the inner product) are polynomials $p_0, p_1, p_2, \ldots$ such that $\deg p_n = n$ (p_0 is a nonzero constant) and $\langle p_n, p_m \rangle = 0$ for $n \neq m$.

Orthogonal polynomials can be obtained by applying the *Gram-Schmidt orthogonalization process* to the basis $1, x, x^2, \ldots$:

$$p_0(x) = 1,$$

$$p_1(x) = x - \frac{\langle x, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0(x),$$

$$p_2(x) = x^2 - \frac{\langle x^2, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0(x) - \frac{\langle x^2, p_1 \rangle}{\langle p_1, p_1 \rangle} p_1(x),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$p_n(x) = x^n - \frac{\langle x^n, p_0 \rangle}{\langle p_0, p_0 \rangle} p_0(x) - \cdots - \frac{\langle x^n, p_{n-1} \rangle}{\langle p_{n-1}, p_{n-1} \rangle} p_{n-1}(x),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

Then $p_0, p_1, p_2, \ldots$ are orthogonal polynomials.

Theorem (a) Orthogonal polynomials always exist.

(b) The orthogonal polynomial of a fixed degree is unique up to scaling.

(c) A polynomial $p \neq 0$ is an orthogonal polynomial if and only if $\langle p, q \rangle = 0$ for any polynomial q with $\deg q < \deg p$.

(d) A polynomial $p \neq 0$ is an orthogonal polynomial if and only if $\langle p, x^k \rangle = 0$ for any $0 \leq k < \deg p$.

Proof of statement (b): Suppose that P and R are two orthogonal polynomials of the same degree n. Then $P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ and $R(x) = b_n x^n + b_{n-1} x^{n-1} + \cdots + b_1 x + b_0$, where $a_n, b_n \neq 0$. Consider a polynomial $Q(x) = b_n P(x) - a_n R(x)$. By construction, $\deg Q < n$. It follows from statement (c) that $\langle P, Q \rangle = \langle R, Q \rangle = 0$. Then

$$\langle Q, Q \rangle = \langle b_n P - a_n R, Q \rangle = b_n \langle P, Q \rangle - a_n \langle R, Q \rangle = 0,$$

which means that $Q = 0$. Thus $R(x) = (a_n^{-1} b_n)\, P(x)$.

Example. $\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\,dx.$

Note that $\langle x^n, x^m\rangle = 0$ if $m+n$ is odd.
Hence $p_{2k}(x)$ contains only even powers of x while $p_{2k+1}(x)$ contains only odd powers of x.

$$p_0(x) = 1,$$
$$p_1(x) = x,$$
$$p_2(x) = x^2 - \frac{\langle x^2, 1\rangle}{\langle 1, 1\rangle} = x^2 - \frac{1}{3},$$
$$p_3(x) = x^3 - \frac{\langle x^3, x\rangle}{\langle x, x\rangle}x = x^3 - \frac{3}{5}x.$$

$p_0, p_1, p_2, \ldots$ are called the **Legendre polynomials**.

Instead of normalization, the orthogonal polynomials are subject to **standardization**.

The standardization for the Legendre polynomials is $P_n(1) = 1$. In particular, $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{1}{2}(3x^2 - 1)$, $P_3(x) = \frac{1}{2}(5x^3 - 3x)$.

Problem. Find $P_4(x)$.

Let $P_4(x) = a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$.
We know that $P_4(1) = 1$ and $\langle P_4, x^k \rangle = 0$ for $0 \leq k \leq 3$.

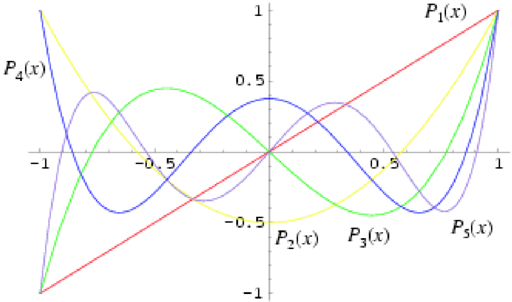
$P_4(1) = a_4 + a_3 + a_2 + a_1 + a_0$,
$\langle P_4, 1 \rangle = \frac{2}{5}a_4 + \frac{2}{3}a_2 + 2a_0$, $\langle P_4, x \rangle = \frac{2}{5}a_3 + \frac{2}{3}a_1$,
$\langle P_4, x^2 \rangle = \frac{2}{7}a_4 + \frac{2}{5}a_2 + \frac{2}{3}a_0$, $\langle P_4, x^3 \rangle = \frac{2}{7}a_3 + \frac{2}{5}a_1$.

$$\begin{cases} a_4 + a_3 + a_2 + a_1 + a_0 = 1 \\ \frac{2}{5}a_4 + \frac{2}{3}a_2 + 2a_0 = 0 \\ \frac{2}{5}a_3 + \frac{2}{3}a_1 = 0 \\ \frac{2}{7}a_4 + \frac{2}{5}a_2 + \frac{2}{3}a_0 = 0 \\ \frac{2}{7}a_3 + \frac{2}{5}a_1 = 0 \end{cases}$$

$$\begin{cases} \frac{2}{5}a_3 + \frac{2}{3}a_1 = 0 \\ \frac{2}{7}a_3 + \frac{2}{5}a_1 = 0 \end{cases} \implies a_1 = a_3 = 0$$

$$\begin{cases} a_4 + a_2 + a_0 = 1 \\ \frac{2}{5}a_4 + \frac{2}{3}a_2 + 2a_0 = 0 \\ \frac{2}{7}a_4 + \frac{2}{5}a_2 + \frac{2}{3}a_0 = 0 \end{cases} \iff \begin{cases} a_4 = \frac{35}{8} \\ a_2 = -\frac{30}{8} \\ a_0 = \frac{3}{8} \end{cases}$$

Thus $P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$.



Legendre polynomials

Problem. Find a quadratic polynomial that is the best least squares fit to the function $f(x) = |x|$ on the interval $[-1, 1]$.

The best least squares fit is a polynomial $p(x)$ that minimizes the distance relative to the integral norm

$$\|f - p\| = \left(\int_{-1}^{1} |f(x) - p(x)|^2 \, dx \right)^{1/2}$$

over all polynomials of degree 2.

The norm $\|f - p\|$ is minimal if p is the orthogonal projection of the function f on the subspace $\mathcal{P}_3$ of polynomials of degree at most 2.

The Legendre polynomials P_0, P_1, P_2 form an orthogonal basis for $\mathcal{P}_3$. Therefore

$$p(x) = \frac{\langle f, P_0 \rangle}{\langle P_0, P_0 \rangle} P_0(x) + \frac{\langle f, P_1 \rangle}{\langle P_1, P_1 \rangle} P_1(x) + \frac{\langle f, P_2 \rangle}{\langle P_2, P_2 \rangle} P_2(x).$$

$$\langle f, P_0 \rangle = \int_{-1}^{1} |x| \, dx = 2 \int_0^1 x \, dx = 1,$$

$$\langle f, P_1 \rangle = \int_{-1}^{1} |x| \, x \, dx = 0,$$

$$\langle f, P_2 \rangle = \int_{-1}^{1} |x| \, \frac{3x^2 - 1}{2} \, dx = \int_0^1 x(3x^2 - 1) \, dx = \frac{1}{4},$$

$$\langle P_0, P_0 \rangle = \int_{-1}^{1} dx = 2, \qquad \langle P_2, P_2 \rangle = \int_{-1}^{1} \left(\frac{3x^2 - 1}{2} \right)^2 dx = \frac{2}{5}.$$

In general, $\langle P_n, P_n \rangle = \dfrac{2}{2n+1}$.

Problem. Find a quadratic polynomial that is the best least squares fit to the function $f(x) = |x|$ on the interval $[-1, 1]$.

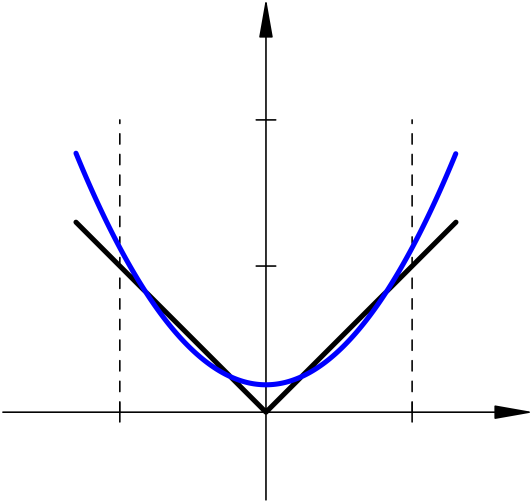
Solution: $p(x) = \frac{1}{2}P_0(x) + \frac{5}{8}P_2(x)$

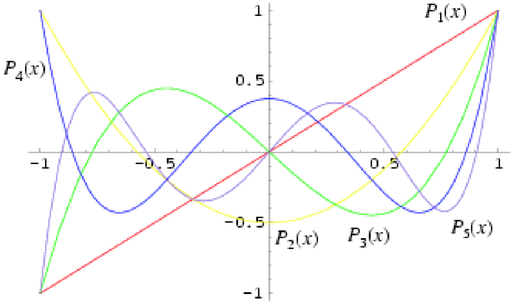
$$= \frac{1}{2} + \frac{5}{16}(3x^2 - 1) = \frac{3}{16}(5x^2 + 1).$$

Recurrent formula for the Legendre polynomials:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x).$$

For example, $4P_4(x) = 7xP_3(x) - 3P_2(x)$.





Legendre polynomials

Definition. **Chebyshev polynomials** $T_0, T_1, T_2, \ldots$ are orthogonal polynomials relative to the inner product

$$\langle p, q \rangle = \int_{-1}^{1} \frac{p(x)q(x)}{\sqrt{1-x^2}}\, dx,$$

with the standardization $T_n(1) = 1$.

Remark. "T" is like in "Tschebyscheff".

Change of variable in the integral: $x = \cos\phi$.

$$\langle p, q \rangle = -\int_0^{\pi} \frac{p(\cos\phi)\, q(\cos\phi)}{\sqrt{1-\cos^2\phi}} \cos'\phi\, d\phi$$

$$= \int_0^{\pi} p(\cos\phi)\, q(\cos\phi)\, d\phi.$$

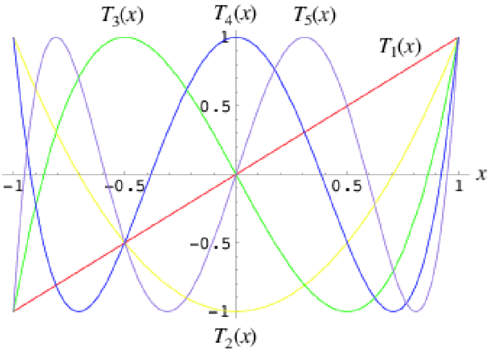
Theorem. $T_n(\cos\phi) = \cos n\phi$.

$$\langle T_n, T_m \rangle = \int_0^\pi T_n(\cos\phi)\, T_m(\cos\phi)\, d\phi$$
$$= \int_0^\pi \cos(n\phi)\cos(m\phi)\, d\phi = 0 \quad \text{if} \quad n \neq m.$$

Recurrent formula: $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$.

$T_0(x) = 1, \quad T_1(x) = x,$
$T_2(x) = 2x^2 - 1,$
$T_3(x) = 4x^3 - 3x,$
$T_4(x) = 8x^4 - 8x^2 + 1, \ \dots$

That is, $\cos 2\phi = 2\cos^2\phi - 1,$
$\cos 3\phi = 4\cos^3\phi - 3\cos\phi,$
$\cos 4\phi = 8\cos^4\phi - 8\cos^2\phi + 1, \ \dots$



Chebyshev polynomials

Topics for the final exam: Part I

Elementary linear algebra (Leon 1.1–1.5, 2.1–2.2)

- Systems of linear equations: elementary operations, Gaussian elimination, back substitution.

- Matrix of coefficients and augmented matrix. Elementary row operations, row echelon form and reduced row echelon form.

- Matrix algebra. Inverse matrix.

- Determinants: explicit formulas for 2×2 and 3×3 matrices, row and column expansions, elementary row and column operations.

Topics for the final exam: Part II

Abstract linear algebra (Leon 3.1–3.6, 4.1–4.3)

- Vector spaces (vectors, matrices, polynomials, functional spaces).
- Subspaces. Nullspace, column space, and row space of a matrix.
- Span, spanning set. Linear independence.
- Bases and dimension.
- Rank and nullity of a matrix.
- Coordinates relative to a basis.
- Change of basis, transition matrix.

- Linear transformations.
- Matrix transformations.
- Matrix of a linear mapping.
- Similarity of matrices.

Topics for the final exam: Parts III–IV

Advanced linear algebra (Leon 5.1–5.7, 6.1–6.3)

- Euclidean structure in $\mathbb{R}^n$ (length, angle, dot product)
- Inner products and norms
- Orthogonal complement
- Least squares problems
- The Gram-Schmidt orthogonalization process
- Orthogonal polynomials

- Eigenvalues, eigenvectors, eigenspaces
- Characteristic polynomial
- Bases of eigenvectors, diagonalization
- Matrix exponentials
- Complex eigenvalues and eigenvectors
- Orthogonal matrices
- Rotations in space

Problem. Consider a linear operator $L : \mathbb{R}^3 \to \mathbb{R}^3$ defined by $L(\mathbf{v}) = \mathbf{v}_0 \times \mathbf{v}$, where $\mathbf{v}_0 = (3/5, 0, -4/5)$.

(a) Find the matrix B of the operator L.
(b) Find the range and kernel of L.
(c) Find the eigenvalues of L.
(d) Find the matrix of the operator L^{2011} (L applied 2011 times).

$$L(\mathbf{v}) = \mathbf{v}_0 \times \mathbf{v}, \quad \mathbf{v}_0 = (3/5, 0, -4/5).$$

Let $\mathbf{v} = (x, y, z) = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$. Then

$$L(\mathbf{v}) = \mathbf{v}_0 \times \mathbf{v} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 3/5 & 0 & -4/5 \\ x & y & z \end{vmatrix}$$

$$= \tfrac{4}{5}y\mathbf{e}_1 - \left(\tfrac{4}{5}x + \tfrac{3}{5}z\right)\mathbf{e}_2 + \tfrac{3}{5}y\mathbf{e}_3.$$

In particular, $L(\mathbf{e}_1) = -\frac{4}{5}\mathbf{e}_2$, $L(\mathbf{e}_2) = \frac{4}{5}\mathbf{e}_1 + \frac{3}{5}\mathbf{e}_3$, $L(\mathbf{e}_3) = -\frac{3}{5}\mathbf{e}_2$.

Therefore $B = \begin{pmatrix} 0 & 4/5 & 0 \\ -4/5 & 0 & -3/5 \\ 0 & 3/5 & 0 \end{pmatrix}.$

$$B = \begin{pmatrix} 0 & 4/5 & 0 \\ -4/5 & 0 & -3/5 \\ 0 & 3/5 & 0 \end{pmatrix}.$$

The range of the operator L is spanned by columns of the matrix B. It follows that $\mathrm{Range}(L)$ is the plane spanned by $\mathbf{v}_1 = (0, 1, 0)$ and $\mathbf{v}_2 = (4, 0, 3)$.

The kernel of L is the nullspace of the matrix B, i.e., the solution set for the equation $B\mathbf{x} = \mathbf{0}$.

$$\begin{pmatrix} 0 & 4/5 & 0 \\ -4/5 & 0 & -3/5 \\ 0 & 3/5 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 3/4 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\implies x + \tfrac{3}{4}z = y = 0 \implies \mathbf{x} = t(-3/4, 0, 1).$$

Alternatively, the kernel of L is the set of vectors $\mathbf{v} \in \mathbb{R}^3$ such that $L(\mathbf{v}) = \mathbf{v}_0 \times \mathbf{v} = \mathbf{0}$.
It follows that this is the line spanned by $\mathbf{v}_0 = (3/5, 0, -4/5)$.

Characteristic polynomial of the matrix B:

$$\det(B - \lambda I) = \begin{vmatrix} -\lambda & 4/5 & 0 \\ -4/5 & -\lambda & -3/5 \\ 0 & 3/5 & -\lambda \end{vmatrix}$$

$$= -\lambda^3 - (3/5)^2\lambda - (4/5)^2\lambda = -\lambda^3 - \lambda = -\lambda(\lambda^2 + 1).$$

The eigenvalues are 0, i, and $-i$.

The matrix of the operator L^{2011} is B^{2011}.

Since the matrix B has eigenvalues 0, i, and $-i$, it is diagonalizable in $\mathbb{C}^3$. Namely, $B = UDU^{-1}$, where U is an invertible matrix with complex entries and

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Then $B^{2011} = UD^{2011}U^{-1}$. We have that $D^{2011} =$ $= \mathrm{diag}\big(0, i^{2011}, (-i)^{2011}\big) = \mathrm{diag}(0, -i, i) = -D$.
Hence

$$B^{2011} = U(-D)U^{-1} = -B = \begin{pmatrix} 0 & -4/5 & 0 \\ 4/5 & 0 & 3/5 \\ 0 & -3/5 & 0 \end{pmatrix}.$$