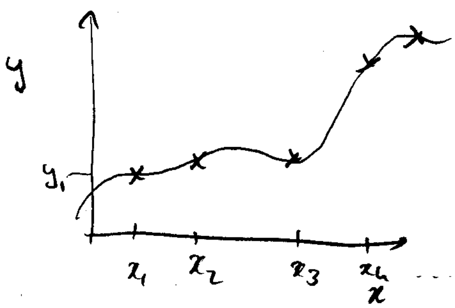
Chapter 8 Approximation theory

§ 8.1 Discrete least squares



polynomial interpolation does not always correspond well with the data. Maybe here data can be better represented by a line. $y = ax + b$ even if this line does not go through all points.

$$\underbrace{\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ x_3 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\underline{x}} \approx \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}}_{\underline{b}}$$

there are different ways of measuring closeness:

$\min_x \| Ax - b \|_\infty$ mini max problem

$\min_x \| Ax - b \|_1$ absolute deviation

$\min_x \| Ax - b \|_2^2$ least-squares

what norm we use depends on application and contrary to what your book says solving problems with $\|\cdot\|_\infty$ and $\|\cdot\|_1$ is feasible and useful (compressed sensing) but requires more knowledge of optimization and is not easy to "generalize" to infinite dimensions.
$\rightarrow$ look at least-squares only.

We assume $A = {}_{n}\Big[\quad\Big]^{m}$ $n > m$ over determined problem with more equations than unknowns.
(underdetermined problem requires an additional technicality)

we also assume A has <u>full column rank</u> i.e. $\operatorname{rank}(A) = \dim(\operatorname{Range}(A))$
$= m$

Then to get least squares solution:

$$q(x) = \|Ax - b\|_2^2 = (Ax-b)^T(Ax-b) = \text{quadratic}$$
$$= x^T A^T A x - 2 b^T A x + b^T b$$

$$\nabla q(x) = 2A^T A x - A^T b$$

$$D^2 q(x) = 2A^T A \quad \text{s.p.d because of our assumptions}$$

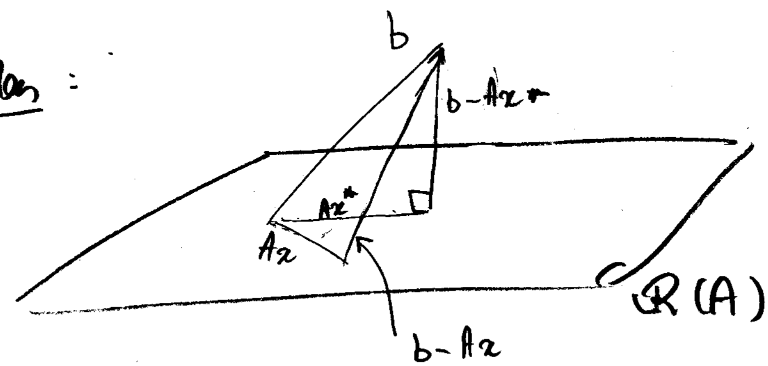
indeed: $x^T A^T A x = \|Ax\|^2 \geq 0$

and $x^T A^T A x = 0 \Rightarrow Ax = 0 \underset{\substack{\uparrow \\ A \text{ full rank}}}{\Rightarrow} x = 0$

Thus it suffices to find x_* s.t. $\nabla q(x_*) = 0$ i.e. solving:

$$\boxed{A^T A x = A^T b} = \text{normal eq}$$

Interpretation :



but x_* must have residual ⊥ to $R(A)$ can you read this from normal eq?

One can use least squares for fitting functions that are more complicated than lines! For example:

$$y = b e^{ax}$$
$$\ln y = a \ln x + \ln b$$

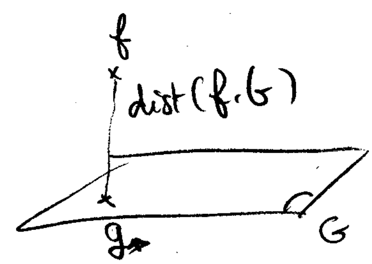
§8.2 Least Square approximation

General problem of best approximation:

Let E be a normed linear space

G be a subspace of E

$$\mathrm{dist}(f, G) = \inf_{g \in G} \| f - g \|$$



The best approx of f in G is g_* s.t. $\| f - g_* \| = \mathrm{dist}(f, G)$

One example of this problem:

$$E = C[a,b], \quad \|f\|_\infty = \max_{x \in [a,b]} |f(x)|$$
$$G = \Pi_n = \text{poly of degree} \le n$$

Question: does approx exist?

Theorem: If G is a finite-dimensional subspace in a NLS E then $\forall f \in E \ \exists g \in G$ s.t. $\|f - g\| = \mathrm{dist}(f, G)$.

We will confine ourselves to Normed Linear spaces where the norm comes from an <u>inner product</u>. This makes life easier.

Recall inner product = <u>symm</u> <u>pos</u> <u>bilinear</u> form

$$\begin{cases} (f,g) = (g,f) \\ (\alpha f+\beta g, h) = \alpha (f,h) + \beta (g,h) \\ (f,f) > 0 \quad \text{if } f \neq 0 \\ \|f\| = \sqrt{(f,f)} \end{cases}$$

examples of inner products spaces:

$\mathbb{R}^n \qquad (x,y) = \sum_{i=1}^{n} x_i y_i$

$C_w[a,b] \qquad (x,y) = \int_a^b w(x) f(x) g(x)\,dx.$

where $w(x) \geq 0$ = weight function.

Inner products give notion of angle so:

$$f \perp g \iff (f,g) = 0$$

$$f \perp G \iff \forall g \in G \ (f,g) = 0.$$

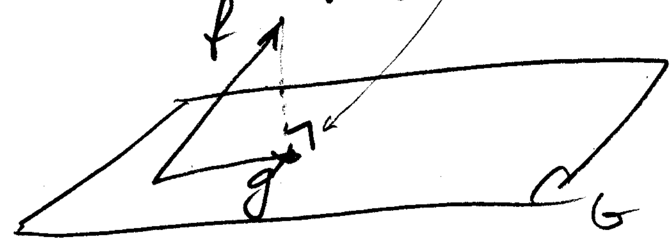
Here are some more properties of inner products that are handy and can be derived from axioms:

- $\|f+g\|^2 = \|f\|^2 + 2(f,g) + \|g\|^2$
- $\|f+g\|^2 = \|f\|^2 + \|g\|^2$, if $f \perp g$ (Pythagorean law)
- $|(f,g)| \leq \|f\|\|g\|$ Cauchy Schwarz inequality

- $\|f+g\|^2 + \|f-g\|^2 = 2\|f\|^2 + 2\|g\|^2$.

<u>Theorem</u> Let G be a subspace of an inner prod space E.

g is best approx to f in $G \iff f-g \perp G$



<u>proof</u>: $\Leftarrow$ assume $f-g \perp G$. Let $h \in G$:

$$(*) \qquad \|f-h\|^2 = \|\underbrace{f-g}_{\perp G} + \underbrace{g-h}_{\in G}\|^2 = \|f-g\|^2 + \|g-h\|^2 \geq \|f-g\|^2$$

$\Rightarrow$ Let g be best approx of f in G. Let $h \in G$ and $\lambda > 0$:

$$0 \leq \|f-g+\lambda h\|^2 - \|f-g\|^2$$

$$= \cancel{\|f-g\|^2} + 2\lambda(f-g,h) + \lambda^2\|h\|^2 - \cancel{\|f-g\|^2}$$

$$= \lambda\left(2(f-g,h) + \lambda\|h\|^2\right)$$

Letting $\lambda \to 0$: $(f-g,h) \geq 0$

Because h is arbitrary in G we must also have: $(f-g,-h) \geq 0$

$\Rightarrow (f-g,h) = 0 \Rightarrow f-g \perp G$

Notice $(*)$ implies <u>uniqueness</u>: Since if g, h are both best approx,

$$\|f-h\|^2 = \|f-g+g-h\|^2 = \|f-g\|^2 + \|g-h\|^2 = \|f-g\|^2$$

$$\Rightarrow \|g-h\| = 0 \Rightarrow g = h.$$

An example with polynomials (see also example 8.2.1)

Find the best approx to $f(x) = \sin x$ by a polynomial

$$g(x) = c_1 x + c_2 x^3 + c_3 x^5 \quad \text{on } [-1,1]$$

with norm:

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

Here $G = \text{span}\{x, x^3, x^5\}$, since x, x^3, x^5 are all lin indep.
we can enforce orthogonality of residual on each basis function.

$$(g - f, g_i) = 0 \quad , i = 1,2,3, \quad g_1(x) = x$$
$$g_2(x) = x^3$$
$$g_3(x) = x^5$$

$$(g, g_i) = (f, g_i) \quad , i = 1,2,3.$$

$$\left(\sum_{j=1}^{3} c_j g_j, g_i\right) = (f, g_i), \quad i = 1,2,3$$

$$\sum_{j=1}^{3} c_j (g_j, g_i) = (f, g_i)$$

$$G c = F \qquad c = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

$$F = \begin{pmatrix} (f, g_1) \\ (f, g_2) \\ (f, g_3) \end{pmatrix} = \begin{bmatrix} \sin(1) - \cos(1) \\ -3\sin(1) + 5\cos(1) \\ 65\sin(1) - 101\cos(1) \end{bmatrix}$$

$$G = \begin{bmatrix} (g_1, g_1) & (g_1, g_2) & (g_1, g_3) \\ (g_2, g_1) & (g_2, g_2) & (g_2, g_3) \\ (g_3, g_1) & (g_3, g_2) & (g_3, g_3) \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{5} & \frac{1}{7} \\ \frac{1}{5} & \frac{1}{7} & \frac{1}{9} \\ \frac{1}{7} & \frac{1}{9} & \frac{1}{11} \end{bmatrix}$$

in this case

Result is: $c_1 \approx -0.99998$
$c_2 \approx -0.16652$
$c_3 \approx 0.00802$

However this is not a good idea as systems tend to be ill-conditioned (i.e. large condition numbers, so very sensitive to errors)

→ Better to use orthogonal systems why?

If we have an orthonormal basis of functions of G then:

$$G = \left[(g_i, g_j) \right]_{i,j=1}^{n} = I$$ = nothing is easier to solve than this!

Gram matrix

What if we do not have an orthonormal basis?

<u>Gram Schmidt</u>: Given $\{v_1, v_2, \dots v_n\}$ basis of U then an orthonormal basis $\{u_1, u_2, \dots, u_n\}$ for U can be obtained by G.S. orthogonalization:

$$u_1 = v_1 / \|v_1\|$$
for $i = 2, \dots, n$
$$u_i = v_i - \sum_{j=1}^{i-1} (v_i, u_j) u_j$$
$$u_i = u_i / \|u_i\|$$

This can be shown by induction.

$u_1 = v_1 / \|v_1\|$ is an orthonormal base of span $\{v_1\}$

assume

$\{u_1, \dots u_i\}$ is an orthonormal basis of span $\{v_1, v_2, \dots, v_i\}$

then $u_{i+1} = v_{i+1} - \sum_{j=1}^{i} (v_{i+1}, u_j) u_j \neq 0$

(otherwise $v_{i+1} \in$ span $\{v_1, \dots v_i\}$)

By second step in loop we have $\|u_{i+1}\| = 1$
By induction hypothesis $\{u_1, \dots, u_i\}$ is orthonormal.
Now we only need to show orthogonality:

$$(u_{i+1}, u_k) = \left(v_{i+1} - \sum_{j=1}^{i} (v_{i+1}, u_j) u_j ,\ u_k\right), \quad k = 1, \dots, i$$

$$= (v_{i+1}, u_k) - \sum_{j=1}^{i} (v_{i+1}, u_j)(u_j, u_k)$$

$$= 0$$

We saw Gram-Schmidt when we were looking at Gaussian quadratures. Specifically we used Legendre polynomials and we saw that for polynomials up to degree n, the Gram-Schmidt procedure simplifies a lot. (in fact only requirement is: $(fg, h) = (f, gh)$)

$$p_0 = 1$$

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

$$p_2 = x p_1 - \frac{(x p_1, p_1)}{(p_1, p_1)} p_1 - \frac{(x p_1, p_0)}{(p_0, p_0)} p_0$$

$$p_3 = x p_2 - \frac{(x p_2, p_2)}{(p_2, p_2)} p_2 - \frac{(x p_2, p_1)}{(p_1, p_1)} p_1 - \cancel{\frac{(x p_2, p_0)}{(p_0, p_0)}} p_0 \qquad 0$$

$$\vdots$$

$$\boxed{p_n = x p_{n-1} - \frac{(x p_{n-1}, p_{n-1})}{(p_{n-1}, p_{n-1})} p_{n-1} - \frac{(x p_{n-1}, p_{n-2})}{(p_{n-2}, p_{n-2})} p_{n-2}}$$

Legendre poly are poly $\perp$ w.r.t inner product:

$$\int_{-1}^{1} f(x) g(x).$$

Another important family of orthogonal polynomials is <u>Chebyshev polynomials</u> which are $\perp$ w.r.t. inner prod:

$$(f,g) = \int_{-1}^{1} f(x)g(x)\frac{dx}{\sqrt{1-x^2}}$$

It is easier to just check orthogonality than deriving it!

$$T_n(x) = \cos(n\cos^{-1}x)$$

$$(T_n, T_m) = \int_{-1}^{1} \cos(n\cos^{-1}x)\cos(m\cos^{-1}x)\frac{dx}{\sqrt{1-x^2}}.$$

Let $x = \cos\theta$

$$\left.\begin{aligned} dx &= \sin\theta\, d\theta \\ &= \sqrt{1-x^2}\, d\theta \end{aligned}\right\} \Rightarrow (T_n, T_m) = \int_0^{\pi} \cos(n\theta)\cos(m\theta)\,d\theta$$

$$= 0 \text{ if } n \neq m$$

A nice property of polynomials that we build w/ Gram Schmidt is that p_n is the polynomial of degree n with leading term 1 (i.e. <u>monic</u>: $p_n(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots$) such that $\|p_n\|$ is minimum

$$p_n = \underset{\substack{p \in \Pi_n \\ p \text{ monic}}}{\operatorname{argmin}} \|p\|$$

Proof: Since p_n is monic of degree $\leq n$ any monic poly of degree $\leq n$ can be written as:

$$q = p_n - \sum_{i=0}^{n-1} c_i p_i$$

$\| q - 0 \|$ is smallest when $q \perp$ span $\{p_0, \dots p_{n-1}\} = \Pi_{n-1}$

but we know we can get this if $c_i = 0$, since $p_n \perp \Pi_{n-1}$.

Chebyshev Polynomials and interpolation

Chebyshev polynomials of the first kind can be defined recursively:

$$T_0(x) = 1$$
$$T_1(x) = x$$
$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x) \qquad n \geq 1$$

Here are some of them:

$$T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x$$
$$T_4(x) = 8x^4 - 8x^2 + 1$$
$$\vdots$$

Theorem: For x in $[-1, 1]$ we have: $T_n(x) = \cos(n \cos^{-1} x)$

proof: comes from cos addition formulae:

$$\cos(n+1)\theta = \cos\theta \cos n\theta - \sin\theta \sin n\theta$$
$$\cos(n-1)\theta = \cos\theta \cos n\theta + \sin\theta \sin n\theta$$

$$\Rightarrow \quad \cos(n+1)\theta = 2\cos\theta \cos n\theta - \cos(n-1)\theta$$

let $\theta = \cos^{-1} x$ $\Rightarrow$ functions $f_n(x) = \cos(n \cos^{-1} x)$ obey recurrence rel for Chebyshev poly.

Thus:

$$|T_n(x)| \leq 1 \qquad -1 \leq x \leq 1$$

$$T_n\left(\cos \frac{j\pi}{n}\right) = (-1)^j \qquad 0 \leq j \leq n \qquad n \text{ sign changes}$$

$$T_n\left(\cos\left(\frac{2j-1}{2n}\pi\right)\right) = 0 \qquad 1 \leq j \leq n \qquad n \text{ zeros}$$

Monic polynomials: polynomial where coeff of highest degree is 1.

Because of recurrence formula:

$$T_n(x) = 2^{n-1} x^n + \cdots$$

$\Rightarrow$ $2^{1-n} T_n(x)$ is a monic polynomial. (for $n \geq 1$)

Theorem on monic polynomials

If p is a monic polynomial of degree n then:

$$\|p\|_\infty = \max_{-1 \leq x \leq 1} |p(x)| \geq 2^{1-n}$$

Proof: For contradiction suppose $|p(x)| < 2^{1-n}$ for $|x| \leq 1$.

Let $q = 2^{1-n} T_n$ and $x_i = \cos(i\pi/n)$. q is monic and of degree n.

Then:

$$(-1)^i p(x_i) \leq |p(x_i)| < 2^{1-n} = (-1)^i \overbrace{q(x_i)}^{=(-1)^i}$$

$\Rightarrow$ $(-1)^i (q(x_i) - p(x_i)) > 0$ for $0 \leq i \leq n$

$\Rightarrow$ polynomial $q(x) - p(x)$ changes sign n times on $[-1, 1]$

$\Rightarrow$ it must have at least n roots on $(-1, 1)$

$\Rightarrow$ contradiction since $\underset{\text{monic}}{q(x)} - \underset{\text{monic}}{p(x)}$ is of degree $n-1$

Now recall error estimate for interpolation on $[-1,1]$ at some nodes x_i

$$f(x) - p_n(x) = f^{(n+1)}(\xi) \prod_{i=0}^{n} (x - x_i) \quad , \text{where } \xi \in [-1,1]$$

Π_n

$$\Rightarrow \quad |f(x) - p(x)| < \frac{1}{(n+1)!} \left| f^{(n+1)}(\xi) \right| \left| \prod_{i=0}^{n} (x - x_i) \right|$$

$$< \frac{1}{(n+1)!} \max_{|\xi| \leq 1} \left| f^{n+1}(\xi) \right| \left| \prod_{i=0}^{n} (x - x_i) \right|$$

$$\Rightarrow \max_{|x| \leq 1} |f(x) - p(x)| < \frac{1}{(n+1)!} \max_{|\xi| \leq 1} f^{(n+1)}(\xi) \max_{|x| \leq 1} \left| \prod_{i=0}^{n} (x - x_i) \right|$$

now we know from our results on Chebyshev poly that,

$$\max_{|x| \leq 1} \left| \underbrace{\prod_{i=0}^{n} (x - x_i)}_{\substack{\text{monic poly} \\ \in \Pi_{n+1}}} \right| \geq 2^{-n}$$

and minimum is attained when $\displaystyle\prod_{i=0}^{n} (x - x_i) = \overbrace{2^{-n} T_{n+1}(x)}^{\text{monic}}$,

i.e. when interp nodes are:

$$x_i = \cos\left(\frac{2i+1}{2n+2} \pi \right) \qquad 0 \leq i \leq n$$

Theorem on interpolation error with Chebyshev interpolation nodes

If the nodes x_i are roots of T_{n+1}, then interp error formula gives:

$$|f(x) - p(x)| \leq \frac{1}{2^{n} (n+1)!} \max_{|t| \leq 1} \left| f^{(n+1)}(t) \right|$$

Here is another example of the use of Chebyshev polynomials.

Let $P_n(x)$ be a polynomial of degree n:

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

Consider the following problem: find polynomial $P_{n-1} \in \Pi_{n-1}$ so that

$$\max_{|x|\le 1} |P_n(x) - P_{n-1}(x)| \text{ is as small as possible.}$$

Let $q(x) = \dfrac{P_n(x) - P_{n-1}(x)}{a_n}$ = monic poly $\in \Pi_n$

$$\Rightarrow \quad \max_{|x|\le 1} |q(x)| \ge \frac{1}{2^{n-1}}$$

with equality happening when $q(x) = 2^{1-n} T_n(x)$

$\Rightarrow$ we should choose:

$$\frac{P_n(x) - P_{n-1}(x)}{a_n} = 2^{1-n} T_n(x)$$

$$P_{n-1} = P_n(x) - 2^{1-n} a_n T_n(x)$$

and with this choice:

$$\max_{|x|\le 1} |P_n(x) - P_{n-1}(x)| = \frac{|a_n|}{2^{n-1}}$$

Of course this procedure can be repeated, however errors are introduced at each step, so we may want to stop when $\dfrac{|a_n|}{2^{n-1}}$ is larger than some prescribed tolerance

Trigonometric interpolation

Idea : Instead of using monomials to interpolate a function, use trigonometric functions (especially when data we want to interpolate is periodic.)

Recall Fourier series : Consider the inner product $(f,g)=\int_{-\pi}^{\pi} f(x)g(x)\,dx$

The functions $1 \; \cos x \; \cos 2x \; \cos 3x, \ldots$
$\sin x \; \sin 2x \; \sin 3x, \ldots$

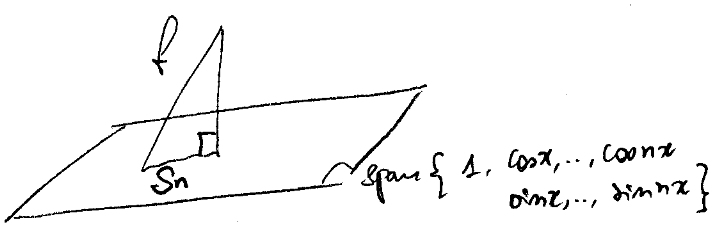
are orthogonal w.r.t. to inner product. This is because:

$$\sin nx \sin mx = \frac{1}{2}\left[\cos(n-m)x - \cos(n+m)x\right]$$

$$\cos nx \cos mx = \frac{1}{2}\left[\cos(n-m)x + \cos(n+m)x\right]$$

$$\sin nx \cos mx = \frac{1}{2}\left[\sin(n-m)x + \sin(n+m)x\right]$$

So we can find best approximation in subspace $\text{span}\{1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx\}$.



by simply taking : $S_n(x) = a_0 + \sum_{k=1}^{n} a_k \cos kx + b_k \sin kx$

$$a_0 = \frac{(f,1)}{(1,1)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\,dx$$

$$a_k = \frac{(f,\cos kx)}{(\cos kx, \cos kx)} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx$$

$$b_k = \frac{(f,\sin kx)}{(\sin kx, \sin kx)} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx$$

If $f(x)$ is continuous : $f(x) = \lim_{n\to\infty} S_n(x) = a_0 + \sum_{k=1}^{\infty} a_k \cos kx + b_k \sin kx$

The theory for Fourier series is more elegant if we use complex exponentials:

$$e^{i\theta} = \cos\theta + i\sin\theta \quad \text{Euler's formula}, \quad i^2 = -1$$

The Fourier series of $f(x)$ is given by:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx} = \sum_{k=-\infty}^{\infty} c_k (e^{ix})^k \leftarrow \text{trigonometric "polynomial".}$$

where $$c_k = \frac{(f, e^{ikx})}{(e^{ikx}, e^{ikx})} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, e^{-ikx}\, dx, \quad k \in \mathbb{Z}$$

and inner product is: $$(f, g) = \int_{-\pi}^{\pi} f(x)\, \overline{g}(x)\, dx$$

$$z = a + ib, \quad \overline{z} = a - ib \quad \text{(complex conjugate)}$$

By Euler's formula:

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x) \underbrace{(\cos kx - i\sin kx)}_{e^{-ikx}}\, dx$$

$$= \frac{1}{2}(a_k - ib_k) \qquad k \geq 1$$

$$c_{-k} = \frac{1}{2}(a_k + ib_k) \qquad c_0 = a_0$$

So the two formulations are equivalent since:

$$a_k = c_k + c_{-k}$$
$$b_k = \frac{c_{-k} - c_k}{i}$$

Now let us look at what happens in the discrete case.

Let us slightly modify inner product so that $E_k(x) \equiv e^{ikx}$ are orthonormal. i.e.

$$(E_m, E_n) = \delta_{mn} = \begin{cases} 1 & \text{if } m=n \\ 0 & \text{if } m \neq n \end{cases}$$

$$(f, g) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \overline{g(x)}\, dx$$

we added this 2π factor.

assume $m \neq n$:

$$(E_m, E_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} E_m(x) \overline{E_n(x)}\, dx$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(m-n)x}\, dx = \frac{1}{2\pi} \frac{e^{i(m-n)x}}{i(m-n)} \Bigg|_{x=-\pi}^{\pi} = 0$$

We now introduce a pseudo-inner product:

$$(f, g)_N = \frac{1}{N} \sum_{j=0}^{N-1} f\left(\frac{2\pi j}{N}\right) \overline{g\left(\frac{2\pi j}{N}\right)}$$

Let us check inner prod axioms:

- $(f, f)_N \geq 0$, however $(f, f)_N = 0 \Rightarrow f\left(\frac{2\pi j}{N}\right) = 0$ $\not\Rightarrow f = 0$.

- $(f, g)_N = \overline{(g, f)_N}$

- $(\alpha f + \beta g, h)_N = \alpha (f, h) + \beta (g, h)$

pseudo norm: $\|f\|_N = \sqrt{(f, f)_N}$

For interpolation we will need the following result:

Theorem on pseudo-inner product.

$$(E_m, E_n)_N = \begin{cases} 1 & \text{if } m-n \text{ is divisible by } N \\ 0 & \text{otherwise.} \end{cases} \qquad (N \geq 1)$$

Proof:

$$(E_m, E_n)_N = \frac{1}{N} \sum_{j=0}^{N-1} e^{\frac{i 2\pi j m}{N}} e^{-\frac{i 2\pi j n}{N}}$$

$$= \frac{1}{N} \sum_{j=0}^{N-1} \left[e^{\frac{i 2\pi}{N}(n-m)} \right]^j$$

$$= \frac{1}{N} \frac{1 - \left[e^{\frac{i 2\pi}{N}(n-m)} \right]^N}{1 - e^{\frac{i 2\pi}{N}(n-m)}} \qquad \text{if } n-m \text{ is not divisible by } N$$

$$= 0$$

when $n-m$ = divisible by N we do have: $(E_m, E_n)_N = 1$ since all terms is the sum are equal to one in this situation.

Interpolation with exponential polynomials

$$P(x) = \sum_{k=0}^{n} c_k e^{ikx} = \sum_{k=0}^{n} c_k (e^{ix})^k = \text{polynomials on unit circle.}$$

Theorem: The exponential polynomial that interpolates f at equally spaced nodes $x_j = \frac{2\pi j}{N}$, $j=0,\dots,N-1$ is:

$$P(x) = \sum_{k=0}^{N-1} c_k e^{ikx}, \text{ where } c_k = (f, E_k)_N$$

proof: We only need to verify that

$$p(x_\ell) = f(x_\ell), \quad \ell = 0, \dots N-1$$

$$p(x_\ell) = \sum_{k=0}^{N-1} c_k e^{ikx_\ell} = \sum_{k=0}^{N-1} (f, E_k)_N e^{ikx_\ell}$$

$$= \sum_{k=0}^{N-1} \frac{1}{N} \sum_{j=0}^{N-1} f(x_j) e^{-ikx_j} e^{ikx_\ell}$$

$$= \sum_{j=0}^{N-1} f(x_j) \underbrace{\frac{1}{N} \sum_{k=0}^{N-1} e^{ik(x_\ell - x_j)}}_{(E_\ell, E_j)_N = \delta_{\ell j}}$$

$$= f(x_\ell).$$