Linear Algebra

Lecture Notes

Rostyslav Hryniv

Ukrainian Catholic University
Business Analytics and Computer Science Programmes

4th term
Spring 2019

Lecture 9. Orthogonalization and QR

Outline

1. Orthonormal bases and orthogonal matrices
   - Orthonormal bases in $\mathbb{R}^n$
   - Orthogonal matrices
2. Gram–Schmidt orthogonalization and QR-decomposition
   - Gram–Schmidt orthogonalization
   - QR-decomposition
   - Applications of QR
3. Fourier transform
   - Hilbert spaces

Orthogonal systems

Definition

A set $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ of nonzero vectors in $\mathbb{R}^n$ is orthogonal if

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0 \text{ for } i \neq j.$$

An orthogonal set S s.t. $\|\mathbf{u}\| = 1$ for each $\mathbf{u} \in S$ is orthonormal.

Example

- The standard basis system of $\mathbb{R}^n$ is an orthonormal set
- The set $(1, 0, 1)$, $(1, 1, -1)$ and $(1, -2, -1)$ is orthogonal

Remark

Every *orthogonal* set S can be made *orthonormal* if we replace each $\mathbf{u}$ in S by $\mathbf{u}/\|\mathbf{u}\|$.

Orthogonal sets are linearly independent

Lemma

Every orthogonal set is linearly independent.

Proof.

Assume that a set $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ is orthogonal and $c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k = \mathbf{0}$. Then

$$\begin{aligned} 0 = \langle \mathbf{u}_m, \mathbf{0} \rangle &= \langle \mathbf{u}_m, c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k \rangle \\ &= c_1\langle \mathbf{u}_m, \mathbf{u}_1 \rangle + \cdots + c_k\langle \mathbf{u}_m, \mathbf{u}_k \rangle = c_m\|\mathbf{u}_m\|^2 \end{aligned}$$

yielding $c_m = 0$. □

Every orthogonal set $S \subset \mathbb{R}^n$ is contained in an orthogonal basis of $\mathbb{R}^n$:

- If $\mathrm{ls}(S) \neq \mathbb{R}^n$, take any nonzero vector $\mathbf{u} \in \big(\mathrm{ls}(S)\big)^\perp$
- Denote $S' := S \cup \{\mathbf{u}\}$; then S' is orthogonal
- Continue (finitely many times) to get $\mathrm{ls}(S') = \mathbb{R}^n$

Coordinate representation in orthogonal bases

Theorem

Assume that $S = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ *is an orthonormal basis (ONB) of* $\mathbb{R}^n$*. Then, for every* $\mathbf{u} \in \mathbb{R}^n$*,*

$$\boxed{\mathbf{u} = \langle \mathbf{u}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{u}, \mathbf{u}_n \rangle \mathbf{u}_n}$$

Proof.

The vector $\mathbf{v} := \mathbf{u} - \langle \mathbf{u}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{u}, \mathbf{u}_n \rangle \mathbf{u}_n$ is orthogonal to every $\mathbf{u}_k$, thus to $\mathrm{ls}\{\mathbf{u}_1, \ldots, \mathbf{u}_n\} = \mathbb{R}^n \implies \mathbf{v} = \mathbf{0}$. □

Corollary

Assume that $S = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ *is an orthogonal basis of* $\mathbb{R}^n$*. Then, for every* $\mathbf{u} \in \mathbb{R}^n$*,*

$$\boxed{\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 + \cdots + \frac{\langle \mathbf{u}, \mathbf{u}_n \rangle}{\|\mathbf{u}_n\|^2} \mathbf{u}_n}$$

$\langle \mathbf{u}, \mathbf{u}_k \rangle / \|\mathbf{u}_k\|^2$ *is the* k^{th} *coordinate of* $\mathbf{u}$ *in the basis* $\mathbf{u}_1, \ldots, \mathbf{u}_n$
$(\langle \mathbf{u}, \mathbf{u}_k \rangle / \|\mathbf{u}_k\|^2) \mathbf{u}_k$ *is the component* $=$ *projection of* $\mathbf{u}$ *onto* $\mathbf{u}_k$*.*

Orthogonal columns and least squares solutions

Properties of matrices with orthogonal columns:

- If an $m \times n$ matrix $A = Q$ has orthonormal columns ($\mathbf{q}_i^T \mathbf{q}_j = \delta_{ij}$), then
$$\boxed{Q^T Q = I_n}$$
- the least squares solution of $Q\mathbf{x} = \mathbf{b}$ is $\hat{\mathbf{x}} = (Q^T Q)^{-1} Q^T \mathbf{b} = Q^T \mathbf{b}$
- the projection $\mathbf{p}$, i.e., the best approximation in $\mathcal{C}(Q)$ is $QQ^T \mathbf{b}$
- the orthogonal projection operator onto $\mathcal{C}(Q) \subset \mathbb{R}^m$ is
$$\boxed{P = QQ^T}$$
- $\mathbf{p} = QQ^T \mathbf{b}$ is the basis decomposition of $\mathbf{p}$ in the ONB $\mathbf{q}_k$ of $\mathcal{C}(Q)$:
$$\mathbf{p} = \mathbf{q}_1(\mathbf{q}_1^T \mathbf{b}) + \mathbf{q}_2(\mathbf{q}_2^T \mathbf{b}) + \cdots + \mathbf{q}_n(\mathbf{q}_n^T \mathbf{b})$$
- When columns of Q are only orthogonal, then the above becomes
$$\mathbf{p} = \frac{\mathbf{q}_1^T \mathbf{b}}{\|\mathbf{q}_1\|^2}\mathbf{q}_1 + \frac{\mathbf{q}_2^T \mathbf{b}}{\|\mathbf{q}_2\|^2}\mathbf{q}_2 + \cdots + \frac{\mathbf{q}_n^T \mathbf{b}}{\|\mathbf{q}_n\|^2}\mathbf{q}_n$$

Orthogonal projectors

Explicit formula for the orthogonal projector P_W onto a subspace W

- Assume W is a subspace of $\mathbb{R}^m$ of dimension n
- Choose an orthogonal basis $\{\mathbf{q}_1, \dots, \mathbf{q}_n\}$ of W
- Denote by Q the $m \times n$ matrix with columns $\mathbf{q}_j$
- Then

$$\boxed{P_W = Q(Q^TQ)^{-1}Q^T = \frac{\mathbf{q}_1\mathbf{q}_1^T}{\|\mathbf{q}_1\|^2} + \frac{\mathbf{q}_2\mathbf{q}_2^T}{\|\mathbf{q}_2\|^2} + \cdots + \frac{\mathbf{q}_n\mathbf{q}_n^T}{\|\mathbf{q}_n\|^2}}$$

and the projection $P_W\mathbf{x}$ is equal to

$$\boxed{P_W\mathbf{x} = \frac{\mathbf{q}_1^T\mathbf{x}}{\|\mathbf{q}_1\|^2}\mathbf{q}_1 + \frac{\mathbf{q}_2^T\mathbf{x}}{\|\mathbf{q}_2\|^2}\mathbf{q}_2 + \cdots + \frac{\mathbf{q}_n^T\mathbf{x}}{\|\mathbf{q}_n\|^2}\mathbf{q}_n}$$

Example (Orthogonal projection)

Let $\mathbf{x} = (x_1, x_2, x_3)^\top$, $\mathbf{q}_1 = (1, 1, 0)^\top$, and $\mathbf{q}_2 = (1, -1, 0)^\top$. Find an orthogonal projection of the vector $\mathbf{x}$ onto the plane $W = \text{ls}\{\mathbf{q}_1, \mathbf{q}_2\}$.

Solution

Since $\mathbf{q}_1 \perp \mathbf{q}_2$ and $\|\mathbf{q}_1\| = \|\mathbf{q}_2\| = \sqrt{2}$, the orthogonal projector onto W is equal to

$$P_W = \frac{\mathbf{q}_1\mathbf{q}_1^\top}{2} + \frac{\mathbf{q}_2\mathbf{q}_2^\top}{2} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \text{diag}\{1, 1, 0\}$$

i.e., P_W is the projector onto the xy-plane. Clearly, $P_W\mathbf{x} = (x_1, x_2, 0)^\top$. Alternatively, we find the orthogonal projection via

$$P_W\mathbf{x} = \frac{\mathbf{q}_1^\top\mathbf{x}}{2}\mathbf{q}_1 + \frac{\mathbf{q}_2^\top\mathbf{x}}{2}\mathbf{q}_2 = \frac{x_1 + x_2}{2}\mathbf{q}_1 + \frac{x_1 - x_2}{2}\mathbf{q}_2 = (x_1, x_2, 0)^\top$$

Orthogonal matrices

- When Q has orthonormal columns and is square, then
$$Q^{-1} = Q^T$$
- the least squares solution is then exact: $\hat{\mathbf{x}} = Q^T\mathbf{b} \implies Q\hat{\mathbf{x}} = \mathbf{b}$

Definition

An $n \times n$ matrix U is called orthogonal if $U^{-1} = U^T$, i.e., if
$$U^T U = UU^T = I_n$$

Criterion for orthogonality

$$\boxed{U \text{ is orthogonal} \iff \text{its columns form an ONB of } \mathbb{R}^n}$$

If $\mathbf{u}_j$ is the j^{th} column of U, then $U^\top U = I_n$ amounts to $\mathbf{u}_k^\top \mathbf{u}_j = \delta_{jk}$.

Another derivation of $\mathbf{u} = \langle \mathbf{u}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \cdots + \langle \mathbf{u}, \mathbf{u}_n \rangle \mathbf{u}_n$

- $\mathbf{u} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$ is just the matrix equality $\mathbf{u} = U\mathbf{c}$
- if $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ forms an ONB of $\mathbb{R}^n$, then $\mathbf{c} = U^{-1}\mathbf{u} = U^T\mathbf{u}$
- then $\mathbf{u} = U\mathbf{c} = UU^T\mathbf{u}$ reads componentwise
$$\mathbf{u} = \mathbf{u}_1(\mathbf{u}_1^T\mathbf{u}) + \mathbf{u}_2(\mathbf{u}_2^T\mathbf{u}) + \cdots + \mathbf{u}_n(\mathbf{u}_n^T\mathbf{u})$$

Properties of orthogonal matrices

An orthogonal matrix U does not change scalar products

Reason: $\langle U\mathbf{u}, U\mathbf{v}\rangle = \langle U^T U\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$

Orthogonality criterion for square matrices

$$\boxed{\text{A matrix } U \text{ is orthogonal} \iff U \text{ does not change length of vectors}}$$

$\Longrightarrow$ from the above

$\Longleftarrow$ if $\forall \mathbf{x} \quad \|U\mathbf{x}\| = \|\mathbf{x}\|$, then $\forall \mathbf{u}, \mathbf{v}$: $\langle U^T U\mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle \implies U^T U = I$

Corollary: An orthogonal matrix preserves the angle between vectors

Reason: $\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u}\|\|\mathbf{v}\| \cos\phi$

Orthogonal matrices and bases

$$\boxed{\text{A matrix } U \text{ is orthogonal} \iff U \text{ sends every ONB into ONB}}$$

$\Longleftarrow$ because then U is square and does not change length of vectors

Gram–Schmidt orthogonalization

Remark

Having a matrix with orthogonal columns or an orthogonal set is very useful

Problem: Construct an orthogonal system $\mathbf{w}_1, \mathbf{w}_2, \dots$ given a system of linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \dots$

- set $\mathbf{w}_1 = \mathbf{v}_1$; $\mathbf{v}_2$ need not be orthogonal to $\mathbf{w}_1$
- subtract from $\mathbf{v}_2$ its projection onto $\mathbf{w}_1$:

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{w}_1^\top \mathbf{v}_2}{\mathbf{w}_1^\top \mathbf{w}_1}\mathbf{w}_1 \implies \mathbf{w}_2^T \perp \mathbf{w}_1 = 0$$

- subtract from $\mathbf{v}_3$ its projection onto the plane $\mathrm{ls}\{\mathbf{w}_1, \mathbf{w}_2\}$:

$$\mathbf{w}_3 := \mathbf{v}_3 - \frac{\mathbf{v}_3^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{v}_3^T \mathbf{w}_2}{\mathbf{w}_2^T \mathbf{w}_2}\mathbf{w}_2$$

- and so on for $\mathbf{w}_4, \mathbf{w}_5, \dots$

Gram–Schmidt orthogonalization

Example (Gram–Schmidt orthogonalization)

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \; \mathbf{v}_2 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \; \mathbf{v}_3 = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix}$$

Solution:

$$\mathbf{w}_1 = \mathbf{v}_1; \qquad \mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2^\top \mathbf{w}_1}{\mathbf{w}_1^\top \mathbf{w}_1}\mathbf{w}_1 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\mathbf{v}_3^T \mathbf{w}_1}{\mathbf{w}_1^T \mathbf{w}_1}\mathbf{w}_1 - \frac{\mathbf{v}_3^T \mathbf{w}_2}{\mathbf{w}_2^T \mathbf{w}_2}\mathbf{w}_2 = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix} - 3\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

QR factorization = matrix form of Gram–Schmidt

- Assume A has linearly independent columns $\mathbf{a}_1, \dots, \mathbf{a}_n$
- perform Gram–Schmidt orthogonalization to get an orthogonal set $\mathbf{w}_1, \dots, \mathbf{w}_n$
- then normalize to get $\mathbf{q}_1, \dots, \mathbf{q}_n$
- at each step, $\mathbf{q}_k$ is a linear combination of $\mathbf{a}_k$, $\mathbf{q}_1, \dots, \mathbf{q}_{k-1}$
- thus $\mathbf{a}_k$ is in the span of $\mathbf{q}_1, \dots, \mathbf{q}_k$
- $\mathbf{a}_k = P_1\mathbf{a}_k + \cdots + P_k\mathbf{a}_k = \mathbf{q}_1\mathbf{q}_1^\top\mathbf{a}_k + \cdots + \mathbf{q}_k\mathbf{q}_k^\top\mathbf{a}_k$
- in matrix form, this becomes a *QR* factorization:

$$\underbrace{\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n\end{pmatrix}}_{A} = \underbrace{\begin{pmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \dots & \mathbf{q}_n\end{pmatrix}}_{Q} \underbrace{\begin{pmatrix}\mathbf{q}_1^\top\mathbf{a}_1 & \mathbf{q}_1^\top\mathbf{a}_2 & \dots & \mathbf{q}_1^\top\mathbf{a}_n \\ & \mathbf{q}_2^\top\mathbf{a}_2 & \dots & \mathbf{q}_2^\top\mathbf{a}_n \\ & & \dots & \dots \\ & & & \mathbf{q}_n^\top\mathbf{a}_n\end{pmatrix}}_{R}$$

- then $R = Q^\top A$: indeed, $Q^\top A = Q^\top(QR) = (Q^\top Q)R = R$

QR factorization, or QR decomposition

QR factorization of an $m \times n$ matrix $A \equiv$ its representation as $A = QR$, where

- an $m \times n$ matrix Q has orthonormal columns;
- an $n \times n$ matrix R is upper-triangular.

A has linearly independent columns $\iff$ R is non-singular

$$A^\top A = (QR)^\top QR = R^\top (Q^\top Q) R = R^\top I_n R = R^\top R$$

Another reason: $\operatorname{rank} A = n \iff \operatorname{rank}(QR) = n \overset{?}{\iff} \operatorname{rank} R = n$

Uniqueness of QR-factorization of A with linearly independent columns:

- $Q_1 R_1 = Q_2 R_2 \iff Q_2^\top Q_1 = R_2 R_1^{-1}$ upper triangular
- however, $Q_2^\top Q_1$ is lower triangular since linear spans of the first k columns of Q_1 and Q_2 are the same for $k = 1, \ldots, n$
- thus $Q_2^\top Q_1$ is a diagonal matrix $D \implies R_2 = DR_1$ and $Q_1 = Q_2 D$
- columns of Q_j are of length 1 $\implies D = \operatorname{diag}\{\pm 1, \pm 1, \ldots, \pm 1\}$

Lifehack: normalize at the very end!

- In the above algorithm, $\mathbf{q}_k = \mathbf{w}_k/\|\mathbf{w}_k\|$, where $\mathbf{w}_1, \dots, \mathbf{w}_n$ are obtained by GS algorithm from columns of A ($\mathbf{w}_k$ not normalized!)
- denote by W the matrix with columns $\mathbf{w}_1, \dots, \mathbf{w}_n$ and by D the diagonal matrix $\operatorname{diag}\{\|\mathbf{w}_1\|, \dots, \|\mathbf{w}_n\|\}$
- then $Q = WD^{-1}$ and $R = Q^\top A = D^{-1}W^\top A$
- therefore, must divide by $\|\mathbf{w}_k\|$ the k^{th} column of W and the k^{th} row of $W^\top A$ to get Q and R respectively

Example

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & 0 & -3 \\ 0 & -2 & 3 \end{pmatrix}, \quad W = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -2 & 1 \end{pmatrix}, \quad D = \operatorname{diag}\{\sqrt{2}, \sqrt{6}, \sqrt{3}\}$$

$$Q \leftarrow \quad \frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{6}} \quad \frac{1}{\sqrt{3}}$$

$$DR = W^T A = \begin{pmatrix} 2 & 2 & 6 \\ 0 & 6 & -6 \\ 0 & 0 & 3 \end{pmatrix} \begin{array}{l} :\sqrt{2} \\ :\sqrt{6} \\ :\sqrt{3} \end{array} \quad \mapsto \quad R$$

Full QR factorization and applications

- In the above form, Q is $m \times n$ and R is $n \times n$
- often called the reduced QR factorization
- Q is not orthogonal (as it is not square)
- add $m - n$ columns to get an orthogonal $\tilde{Q}$ and add $m - n$ zero rows to R; then $A = \tilde{Q}\tilde{R}$ is the full QR factorization

Application of QR to least squares:

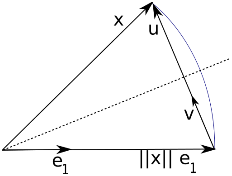
- $A^\top A = (QR)^\top (QR) = R^\top R$;
- since $A^\top A$ is invertible, such is also $R^\top R$ and thus R
- therefore, $R^\top R\,\hat{\mathbf{x}} = R^\top Q^T \mathbf{b} \implies R\hat{\mathbf{x}} = Q^\top \mathbf{b}$
- as R is upper-triangular, this is very fast!

Advantages of QR-decomposition:

Orthogonal columns of Q make algorithm stable
(norms do not increase or decrease)

Householder's reflection algorithm of QR

- another methods to find Q and R involve Householder's reflections $Q = I - 2\mathbf{v}\mathbf{v}^T$ with $\|v\| = 1$
- Householder's reflection Q can be chosen so that $Q\mathbf{x} = \|\mathbf{x}\|\mathbf{e}_1$: set $\mathbf{u} = \mathbf{x} - \|\mathbf{x}\|\mathbf{e}_1$ and $\mathbf{v} = \mathbf{u}/\|\mathbf{u}\|$



- the impact on A: take $\mathbf{x}$ to be the first column of A; then

$$QA = \begin{pmatrix} \|\mathbf{x}\| & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \vdots & \vdots \\ 0 & * & \dots & * \end{pmatrix}$$

- Then consider $(n-1) \times (n-1)$ submatrix and continue

Givens rotations and QR

The idea in dimension 2

Rotate a vector $(x, y)^\top$ of length r to make it collinear to $\mathbf{e}_1 = (1, 0)^\top$:

$$G(\theta)\mathbf{x} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \mathbf{x} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix}$$

We easily find that $\cos\theta = x/r = x/\sqrt{x^2+y^2}$ and $\sin\theta = -y/\sqrt{x^2+y^2}$

Algrithm:

Apply Givens rotations in coordinates 1 and n, then 1 and $n-1$ etc to make all entries below the $(1, 1)$-entry zero

Pros:

- easily parallelized
- fast for sparse matrices

Applications

- *QR* eigenvalue algorithm
- Fourier transform and Fast Fourier transform
- ...

Hilbert space $L_2(0, 2\pi)$

- Consider the space of functions over $[0, 2\pi]$ that are square integrable:
$$\|f\|^2 := \int_0^{2\pi} |f(t)|^2 \, dt$$
- this is the inner product space called the Hilbert space $L_2(0, 2\pi)$
- scalar product:
$$\langle f, g \rangle := \int_0^{2\pi} f(t)\overline{g(t)} \, dt$$
- the set of functions 1, $\cos nx$, $\sin nx$, $n = 1, 2 \ldots,$ forms an orthogonal set in $L_2(0, \pi)$:

$$\sin nx \cos mx = \tfrac{1}{2}\sin(n+m)x + \tfrac{1}{2}\sin(n-m)x \implies$$
$$\langle \sin nx, \cos mx \rangle = 0;$$
$$\langle 1, \sin nx \rangle = \langle 1, \cos mx \rangle = 0 \qquad \text{etc}$$

Fourier series

- $\{1/\sqrt{2\pi}, \sin nx/\sqrt{\pi}, \cos nx/\sqrt{\pi}\}$ forms an orthonormal basis of $L_2(0, 2\pi)$
- every function $f \in L_2(0, 2\pi)$ is equal (in some sense) to its Fourier series

$$f(x) \sim \tfrac{1}{2}a_0 + \sum_{k>0}(a_k \cos kx + b_k \sin kx)$$

with

$$a_k := \frac{1}{\pi}\int_0^{2\pi} f(t)\cos kt\, dt, \qquad b_k := \frac{1}{\pi}\int_0^{2\pi} f(t)\sin kt\, dt,$$

being the Fourier coefficients of f

Complex trigonometric Fourier series

- This time we consider complex-valued functions over $(0, 2\pi)$
- can associate any such function with a function on the unit circle
- Another basis for $L_2(0, 2\pi)$ is $\{e^{inx}\}$, $n \in \mathbb{Z}$, with $i = \sqrt{-1}$
- This is an orthogonal basis:

$$\langle e^{inx}, e^{imx} \rangle = \int_0^{2\pi} e^{inx}\overline{e^{imx}}\,dx = \int_0^{2\pi} e^{i(n-m)x}\,dx = 2\pi\delta_{n,m}$$

- Fourier series of a function f:

$$f \sim \sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

  where

$$c_k := \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-ikt}\,dt$$

  are the Fourier coefficients

The Fourier transform

Definition

The mapping $\mathscr{F} : L_2(0,2\pi) \to \ell_2(\mathbb{Z})$ given by

$$\mathscr{F}(f) = (c_k)$$

is the Fourier transform

- ℓ_2 is a vector space of all complex-valued sequences $\mathbf{c} = (c_k)_{k=-\infty}^{\infty}$ with norm $\|\mathbf{c}\| = \sqrt{\sum |c_k|^2}$
- The Pythagorean thm = Parseval thm:

$$\|f\|^2 = \sum |c_k|^2$$

- Plancherel thm: $f \sim \sum c_k e^{ikx}$, $g \sim \sum d_k e^{ikx} \implies$

$$\langle f, g \rangle = \sum c_k \overline{d_k}$$

Example (Fourier transform of $f = \cos 2x$)

$\cos 2x = \frac{1}{2} e^{2ix} + \frac{1}{2} e^{-2ix}$ is the Fourier series of f;

$\mathscr{F}(f) = (\ldots, 0, 0, \ldots, 0, \frac{1}{2}, 0, \frac{1}{2}, 0, \ldots, 0, 0, \ldots)$