

- Let \vec{x} and \vec{y} be vectors in an inner product space and let $\|\cdot\|$ be derived from the inner product. Prove that $\vec{x} \perp \vec{y}$ if and only if $\|\vec{y}\| \leq \|\vec{y} + t\vec{x}\|$ for all scalars $t \in \mathbb{R}$.

Solution: Consider the equivalent statement $\vec{x} \perp \vec{y}$ if and only if $\|\vec{y}\| \leq \|\vec{y} + t\vec{x}\|$ for all scalars $t \in \mathbb{R}$. Expand

$$\|\vec{y} + t\vec{x}\|^2 = \langle \vec{y} + t\vec{x}, \vec{y} + t\vec{x} \rangle = \|\vec{y}\|^2 + 2t\langle \vec{y}, \vec{x} \rangle + t^2\|\vec{x}\|^2.$$

If $\vec{x} \perp \vec{y}$ then the implication is obvious.

Now consider the minimum value of $\|\vec{y} + t\vec{x}\|^2$ as a function of t , which occurs at the critical point where the derivative with respect to t is zero. The critical point occurs at the solution of

$$2\langle \vec{y}, \vec{x} \rangle + 2t\|\vec{x}\|^2 = 0$$

i.e.

$$t = -\frac{\langle \vec{y}, \vec{x} \rangle}{\|\vec{x}\|^2}.$$

The minimum value is obtained by plugging this value of t back in, which yields

$$\min \|\vec{y} + t\vec{x}\|^2 = \|\vec{y}\|^2 - \frac{(\langle \vec{y}, \vec{x} \rangle)^2}{\|\vec{x}\|^2}.$$

The only way that this minimum value can be equal to $\|\vec{y}\|^2$ is if $\vec{x} \perp \vec{y}$.

- Prove that $\langle \mathbf{A}, \mathbf{B} \rangle = \text{Tr}(\mathbf{A}\mathbf{B}^T)$ is an inner product on the space of real $m \times n$ matrices. (Note that this implies that $(\langle \mathbf{A}, \mathbf{A} \rangle)^{1/2}$ defines a norm on real $m \times n$ matrices; this norm is called the Frobenius norm. It is not an operator norm. It is not the matrix 2-norm.)

Solution: Using the definition of the trace and the definition of matrix multiplication we have

$$\text{Tr}(\mathbf{A}\mathbf{B}^T) = \sum_i \sum_k a_{ik} b_{ik}.$$

I.e. we multiply the elements of the matrices and sum them up. This is the same as the dot product, so it satisfies the same properties. If you want to check each property by hand:

- Positivity: $\text{Tr}(\mathbf{A}\mathbf{A}^T) = \sum_i \sum_k a_{ik}^2$. This is non-negative and the only way to get 0 is for all the entries of \mathbf{A} to be zero.
- Symmetry: Taking the transpose of a matrix does not change the diagonal elements, so it does not change the trace of the matrix. Therefore

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{Tr}(\mathbf{A}\mathbf{B}^T) = \text{Tr}((\mathbf{A}\mathbf{B}^T)^T) = \text{Tr}(\mathbf{B}\mathbf{A}^T) = \langle \mathbf{B}, \mathbf{A} \rangle.$$

- Bilinearity:

$$\begin{aligned} \langle \alpha\mathbf{A} + \beta\mathbf{B}, \mathbf{C} \rangle &= \text{Tr}((\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{C}^T) = \text{Tr}((\mathbf{A}\mathbf{B}^T)^T) = \text{Tr}(\alpha\mathbf{A}\mathbf{C}^T + \beta\mathbf{B}\mathbf{C}^T) \\ &= \alpha\text{Tr}(\mathbf{A}\mathbf{C}^T) + \beta\text{Tr}(\mathbf{B}\mathbf{C}^T) = \alpha\langle \mathbf{A}, \mathbf{C} \rangle + \beta\langle \mathbf{B}, \mathbf{C} \rangle \end{aligned}$$

3. Suppose that you have a norm $\|\cdot\|$ on a real vector space but you don't know if it comes from an inner product. Prove that if the norm does come from an inner product, then the inner product must be

$$\langle \vec{x}, \vec{y} \rangle = \frac{1}{4} (\|\vec{x} + \vec{y}\|^2 - \|\vec{x} - \vec{y}\|^2).$$

This is called a polarization identity. If the norm does not come from an inner product, which of the three properties of an inner product will it fail to satisfy?

Solution:

$$\begin{aligned} \|\vec{x} + \vec{y}\|^2 &= \langle \vec{x} + \vec{y}, \vec{x} + \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, \vec{y} \rangle + \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle = \|\vec{x}\|^2 + 2\langle \vec{x}, \vec{y} \rangle + \|\vec{y}\|^2 \\ \|\vec{x} - \vec{y}\|^2 &= \langle \vec{x} - \vec{y}, \vec{x} - \vec{y} \rangle = \langle \vec{x}, \vec{x} \rangle - \langle \vec{x}, \vec{y} \rangle - \langle \vec{y}, \vec{x} \rangle + \langle \vec{y}, \vec{y} \rangle = \|\vec{x}\|^2 - 2\langle \vec{x}, \vec{y} \rangle + \|\vec{y}\|^2 \end{aligned}$$

Top row minus the bottom row, divided by 4 yields the identity.

The expression above is symmetric and non-negative, so if it fails to be an inner product it must fail the bilinearity condition.

4. Consider the vector space of polynomials of degree $\leq n$ and the inner product $\langle p, q \rangle = \int_0^1 p(x)q(x)dx$. Find an expression for the Gram matrix associated with the monomial basis for this space. This matrix is called the Hilbert matrix. Explain why it is positive definite.

Solution: Let the monomial basis be $\vec{v}_i = x^{i-1}$. The entries of the Gram matrix are the inner products of these vectors:

$$(\mathbf{H})_{ij} = \langle \vec{v}_i, \vec{v}_j \rangle = \int_0^1 x^{i+j-2} dx = \frac{1}{i+j-1}.$$

Every Gram matrix is symmetric and at least non-negative definite. It is positive definite if and only if the vectors used to construct the Gram matrix are linearly independent. Here we used the monomial *basis* to construct the Gram matrix, and the vectors in a basis are always linearly independent, so the Gram matrix must be positive definite.

5. Every square matrix \mathbf{A} can be uniquely decomposed into a sum of a symmetric and an antisymmetric matrix:

$$\mathbf{A} = \frac{\mathbf{A} + \mathbf{A}^T}{2} + \frac{\mathbf{A} - \mathbf{A}^T}{2}.$$

Show that if \mathbf{J} is real and antisymmetric, then $\vec{x}^T \mathbf{J} \vec{x} = 0$ for every $\vec{x} \in \mathbb{R}^n$.

Solution: Note that $\vec{x}^T \mathbf{J} \vec{x}$ is a scalar, so it is equal to its own transpose:

$$\vec{x}^T \mathbf{J} \vec{x} = (\vec{x}^T \mathbf{J} \vec{x})^T = \vec{x}^T \mathbf{J}^T \vec{x} = -\vec{x}^T \mathbf{J} \vec{x}.$$

The only number equal to its negative is zero.

6. Riesz representation theorem:

- (a) Consider the linear functional $L[x\vec{e}_1 + y\vec{e}_2 + z\vec{e}_3] = y$ and the inner product defined by

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^T \mathbf{K} \vec{y} \text{ where } \mathbf{K} = \begin{bmatrix} 2 & 1 \\ 1 & 2 & 1 \\ & 1 & 2 \end{bmatrix}.$$

Find the vector $\vec{z} \in \mathbb{R}^3$ such that $L[\vec{x}] = \langle \vec{x}, \vec{z} \rangle$ for all \vec{x} . (Use `numpy` or Mathematica if you don't want to solve by hand.)

Solution: We can write

$$L[\vec{x}] = \vec{x}^T \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

We want to find a \vec{z} that solves

$$\vec{x}^T \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \vec{x}^T \mathbf{K} \vec{z}$$

for every \vec{x} , i.e.

$$\mathbf{K} \vec{z} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

The solution is

$$\vec{z} = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{pmatrix}.$$

- (b) Consider the linear functional $L[p] = \int_0^1 e^{-x} p(x) dx$ defined on polynomials of degree ≤ 2 and using the inner product $\langle p, q \rangle = \int_0^1 p(x) q(x) dx$. Find the polynomial $r(x)$ of degree ≤ 2 such that $L[p] = \langle p, r \rangle$ for every polynomial p of degree ≤ 2 . (Use Mathematica or similar to speed up the analysis.)

Solution: Use the monomial basis. Any p in the space can be written $p(x) = a + bx + cx^2$. The polynomial r that we're looking for can also be written in the monomial basis: $r(x) = u + vx + wx^2$. The condition is that

$$\int_0^1 e^{-x} (a + bx + cx^2) dx = \int_0^1 (u + vx + wx^2) (a + bx + cx^2) dx \quad \forall a, b, c \in \mathbb{R}.$$

As in part (a) we can get a system of equations by considering one equation for a , one for b , and one for c . This yields the linear system

$$\begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix} \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1 - \frac{1}{e} \\ 1 - \frac{2}{e} \\ 2 - \frac{5}{e} \end{pmatrix}.$$

The solution is

$$\begin{pmatrix} \frac{3(11e-29)}{e} \\ -\frac{12(17e-46)}{e} \\ \frac{30(7e-19)}{e} \end{pmatrix} \text{ i.e. } r(x) = \frac{3(11e-29)}{e} + \left(-\frac{12(17e-46)}{e}\right)x + \left(\frac{30(7e-19)}{e}\right)x^2$$

If you plot $r(x)$ and e^{-x} on $x \in [0, 1]$ you will see that $r(x)$ closely approximates e^{-x} over this range.

7. Consider the linear function $L[f](x) = d[xf'(x)]/dx$ acting on polynomials of degree ≤ 2 , and the inner product $\langle p, q \rangle = \int_0^1 xp(x)q(x)dx$. What is the adjoint of L with respect to this inner product? Hint: Express your solution as a map from monomial coefficients to monomial coefficients, and use Mathematica or numpy to aid the computation.

Solution: You may be tempted to use integration by parts here, but it will not work well because it doesn't know that we're dealing with polynomials of degree ≤ 2 .

We want a linear function L^\dagger on polynomials such that

$$\langle p, L[q] \rangle = \int_0^1 xp(x) \frac{d[xq'(x)]}{dx} dx = \langle L^\dagger[p], q \rangle = \int_0^1 xL^\dagger[p](x)q(x)dx$$

for every p and q in the space.

We could find a matrix representation of L as well as of the inner product, and then use matrix algebra to find the adjoint. We will instead pursue an equivalent approach.

As above, let $\vec{v}_i = x^{i-1}$ denote the monomial basis vectors. Let

$$L^\dagger[\vec{v}_j] = \sum_i a_{ij} \vec{v}_i.$$

We want the above condition to hold for every p and q , so we can impose that it holds for every combination $p = \vec{v}_i$, $q = \vec{v}_j$, which leads to

$$\langle \vec{v}_i, L[\vec{v}_j] \rangle = \int_0^1 xx^{i-1}(j-1)^2x^{j-2}dx = \langle L^\dagger[\vec{v}_i], \vec{v}_j \rangle = \int_0^1 x(\sum_m a_{mi}x^{m-1})x^{j-1}dx$$

i.e.

$$\frac{(j-1)^2}{i+j-1} = \sum_m \frac{a_{mi}}{m+j} \text{ for } i = 1, 2, 3 \text{ and } j = 1, 2, 3.$$

This is a system of 9 linear equations in 9 unknowns.

The solution is

$$L^\dagger[a + bx + cx^2] = x^2(440a + 360b + 300c) - x(510a + 420b + 351c) + (120a + 100b + 84c)$$