1 Some examples

1. The functions below are examples of inner products:

   (a). $V = \mathbb{C}([0,1]) = \{f : [0,1] \to \mathbb{C} \text{ continuous}\}$.
   $< f, g >= \int_0^1 f\overline{g}$

   (b). $V = M_n(\mathbb{C}), < A, B >= \mathrm{tr}(AB^*)$, where $B^* = \overline{B^t}$.
   Proof: The conditions of the inner products can be established as below:
   • $< A+B, C >= \mathrm{tr}((A+B)C^*) = \mathrm{tr}(AC^* + BC^*) = \mathrm{tr}(AC^*) + \mathrm{tr}(BC^*) =< A+C, B+C >$.
   • for any constant c, $< cA, B >= \mathrm{tr}(c(AB^*)) = c\mathrm{tr}(AB^*) = c < A, B >$.
   • $< A, B >= \mathrm{tr}(AB^*) = \mathrm{tr}(A\overline{B^t}) = \sum(A\overline{B^t})_{ii} = \sum A_{ij}\overline{B^t_{ji}} = \sum A_{ij}\overline{B_{ij}}$, $\forall 1 \leq i, j \leq n$. Similarly, $< B, A >= \sum B_{ij}\overline{A_{ij}}$. Now for $a, b \in \mathbb{C}$ we have $\overline{a} + \overline{b} = \overline{a+b}$, $\overline{ab} = \overline{a}\overline{b}$ and $\overline{\overline{a}} = a$. Therefore $\overline{a\overline{b}} = \overline{a}\overline{\overline{b}} = \overline{a}b$. This gives $A_{ij}\overline{B_{ij}} = \overline{B_{ij}\overline{A_{ij}}}$ and therefore $< A, B >= A_{ij}\overline{B_{ij}} = \overline{B_{ij}\overline{A_{ij}}} = \overline{< B, A >}$.
   • From above, $< A, A >= \sum A_{ij}\overline{A_{ij}} = \sum ||A_{ij}||^2$. This is obviously nonnegative, and it is zero if and only if all $||A_{ij}||$'s are zero, meaning that A_{ij} must be itself a zero (i.e. a zero vector).

2. In assignment 1 problem 1, we have seen that the pairing isn't an inner product because there exists nonzero vector $\boldsymbol{x}$ satisfying $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$. We now show that the pairing $\boldsymbol{x}^t \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \overline{\boldsymbol{y}}$ satisfies all other properties.

   Notice that, if $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ then

   $$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \boldsymbol{x}^t A \overline{\boldsymbol{y}} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \begin{pmatrix} \overline{y_1} \\ \overline{y_2} \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} \overline{y_1} + i\overline{y_2} \\ -i\overline{y_1} + \overline{y_2} \end{pmatrix} = \begin{pmatrix} x_1(\overline{y_1} + i\overline{y_2}) + x_2(-i\overline{y_1} + \overline{y_2}) \end{pmatrix}.$$

   We establish the following:
   • $\langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z} \rangle = \boldsymbol{x} + \boldsymbol{y}^t A \overline{\boldsymbol{z}} = (\boldsymbol{x}^t + \boldsymbol{y}^t) A \overline{\boldsymbol{z}} = \boldsymbol{x}^t A \overline{\boldsymbol{z}} + \boldsymbol{y}^t A \overline{\boldsymbol{z}} = \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \boldsymbol{y}, \boldsymbol{z} \rangle$.
   • For any constant c, $\langle \boldsymbol{cx}, \boldsymbol{y} \rangle = (c\boldsymbol{x}^t) A \overline{\boldsymbol{y}} = c(\boldsymbol{x}^t A \overline{\boldsymbol{y}}) = c \langle \boldsymbol{x}, \boldsymbol{y} \rangle$.
   • Before proving $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}$, we need the following properties about complex numbers: for any complex numbers a and b, we have $\overline{a} + \overline{b} = \overline{a+b}$; for any complex numbers a and b, $\overline{a} \cdot \overline{b} = \overline{ab}$. Therefore,

   $$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \begin{pmatrix} x_1(\overline{y_1} + i\overline{y_2}) + x_2(-i\overline{y_1} + \overline{y_2}) \end{pmatrix}, \langle \boldsymbol{y}, \boldsymbol{x} \rangle = \begin{pmatrix} y_1(\overline{x_1} + i\overline{x_2}) + y_2(-i\overline{x_1} + \overline{x_2}) \end{pmatrix}.$$

   We have $x_1\overline{y_1} = \overline{\overline{x_1 y_1}} = \overline{\overline{x_1} y_1}$, and similarly $x_2\overline{y_2} = \overline{\overline{x_2 y_2}} = \overline{\overline{x_2} y_2}$. In addition, $i(x_1\overline{y_2} - x_2\overline{y_1}) = i(\overline{\overline{x_1}y_2 - \overline{x_2}y_1}) = -i(\overline{\overline{x_2}y_1 - \overline{x_1}y_2}) = \overline{i}(\overline{\overline{x_2}y_1 - \overline{x_1}y_2}) = \overline{i(\overline{x_2}y_1 - \overline{x_1}y_2)}$. Therefore,

   $$x_1(\overline{y_1}+i\overline{y_2})+x_2(-i\overline{y_1}+\overline{y_2}) = x_1\overline{y_1}+x_2\overline{y_2}+i(x_1\overline{y_2}-x_2\overline{y_1}) = \overline{\overline{x_1}y_1}+\overline{\overline{x_2}y_2}+\overline{i(\overline{x_2}y_1 - \overline{x_1}y_2)} = \overline{\overline{x_1}y_1 + \overline{x_2}y_2 + i(\overline{x_2}y_1 - \overline{x_1}y_2)}$$

   $= \overline{y_1(\overline{x_1} + i\overline{x_2}) + y_2(-i\overline{x_1} + \overline{x_2}}$, establishing the claim.
   • Now $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \begin{pmatrix} x_1(\overline{x_1} + i\overline{x_2}) + x_2(-i\overline{x_1} + \overline{x_2}) \end{pmatrix} = (x_1\overline{x_1} + x_2\overline{x_2} + i(x_1\overline{x_2} - x_2\overline{x_1}) = |x_1|^2 + |x_2|^2 + ix_1\overline{x_2} + (-i)\overline{x_1}x_2 = |x_1|^2 + |x_2|^2 + ix_1\overline{x_2} + \overline{i}\overline{x_1}\overline{\overline{x_2}} = |x_1|^2 + |x_2|^2 + ix_1\overline{x_2} + \overline{ix_1\overline{x_2}} = |x_1|^2 + |x_2|^2 + 2Re(ix_1\overline{x_2})$, because $a + \overline{a} = 2Re(a)$. Now, $|2Re(ix_1\overline{x_2})| \leq |2(ix_1\overline{x_2})| \leq 2|x_1x_2|$ so $-2|x_1x_2| \leq |2Re(ix_1\overline{x_2})| \leq 2|x_1x_2|$, so $|x_1|^2 + |x_2|^2 + 2Re(ix_1\overline{x_2}) \geq |x_1|^2 + |x_2|^2 - 2|x_1||x_2| = (|x_1| - |x_2|)^2$, so the pairing is always nonnegative. Notice, however, it could happen that this quantity is indeed 0 even with both x_1, x_2 nonzero.

2 Proofs of identities

1. Given basis $\{\vec{w_1}, \cdots, \vec{w_n}\}$ of an inner product space, prove that the the set of vectors $\{\vec{v_1}, \cdots, \vec{v_n}\}$ defined as $\vec{v_1} = \vec{w_1}$ and

   $$\vec{v_k} = \vec{w_k} - \sum_{i=1}^{k-1} \frac{\langle \vec{w_k}, \vec{v_i} \rangle}{||\vec{v_i}||^2} \vec{v_i}, \qquad \forall\, k \in [2, n]$$

   is an orthogonal basis.

Proof: (Credits to textbook and prof). First, we prove that $\langle\vec{i},\vec{j}=0,\ \forall i\neq j$. We also proeed by inducting on n. Base case where $n=1$ is trivial. Suppose the claim holds for $n=1,2,\cdots k-1$ for some k, we have: for any $j<k$,

$$\langle\vec{v_k},\vec{v_j}\rangle=\langle\vec{w_k}-\sum_{i=1}^{k-1}\frac{\langle\vec{w_k},\vec{v_i}\rangle}{||\vec{v_i}||^2}\vec{v_i},\vec{v_j}\rangle=\langle\vec{w_k},\vec{v_j}\rangle-\langle\sum_{i=1}^{k-1}\frac{\langle\vec{w_k},\vec{v_i}\rangle}{||\vec{v_i}||^2}\vec{v_i},\vec{v_j}\rangle=\langle\vec{w_k},\vec{v_j}\rangle-\sum_{i=1}^{k-1}\frac{\langle\vec{w_k},\vec{v_i}\rangle}{||\vec{v_i}||^2}\langle\vec{v_i},\vec{v_j}\rangle$$

$$=\langle\vec{w_k},\vec{v_j}\rangle-\frac{\langle\vec{w_k},\vec{v_j}\rangle}{||\vec{v_j}||^2}\langle\vec{v_j},\vec{v_j}\rangle=\langle\vec{w_k},\vec{v_j}\rangle-\langle\vec{w_k},\vec{v_j}\rangle=0,$$

justifying the claim. (By induction hypothesis, $\langle\vec{i},\vec{j}\rangle=0$ for any $i<j<k$.

Next, notice that none of the vectors $\vec{v_i}$ can be zero; each of the vectors $\vec{v_k}$ can be written as the linear combination of $\vec{w_1},\cdots,\vec{w_k}$, with the coefficient of $\vec{w_k}$ being 1. Since $\vec{w_1},\cdots,\vec{w_k}$ are linearly independent, the claim follows.

Finally, in class we have seen that a set of nonzero orthogonal vectors must be linearly independent. Since the set of vectors $\{\vec{v_1},\cdots,\vec{v_n}\}$ has n elements and are linearly independent, this set is also a basis. The conclusion follows.

2. Given a finite dimensional inner-product space V and let W be its subspace with orthonormal basis $\{\vec{w_1},\cdots,\vec{w_k}\}$. Then for each $\vec{v}\in V$ there exists a unique $\vec{w}\in W$ and $\vec{w'}\in W^\perp$ satisfying $\vec{w}+\vec{w'}=\vec{v}$, given by the following formula:

$$\vec{w}=\sum_{i=1}^{k}\frac{\langle\vec{v},\vec{w_i}\rangle}{||\vec{w_i}||^2}\vec{w_i},\vec{w'}=\vec{v}-\vec{w}.$$

Proof: since a subspace (or a vector space, in general) is closed under addition, $\vec{w}$ described above is in W. To show that $\vec{w'}\in W^\perp$, we notice the following for all $j\in[1,n]$:

$$\langle\vec{w'},\vec{w_j}\rangle=\langle\vec{v}-\sum_{i=1}^{k}\frac{\langle\vec{v},\vec{w_i}\rangle}{||\vec{w_i}||^2}\vec{w_i},\vec{w_j}\rangle=\langle\vec{v},\vec{w_j}\rangle-\langle\sum_{i=1}^{k}\frac{\langle\vec{v},\vec{w_i}\rangle}{||\vec{w_i}||^2}\vec{w_i},\vec{w_j}\rangle=\langle\vec{v},\vec{w_j}\rangle-\sum_{i=1}^{k}\frac{\langle\vec{v},\vec{w_i}\rangle}{||\vec{w_i}||^2}\langle\vec{w_i},\vec{w_j}\rangle=\langle\vec{v},\vec{w_j}\rangle-\langle\vec{v},\vec{w_j}\rangle=0,$$

because $\langle\vec{w_i},\vec{w_j}\rangle$ vanishes whenever $i\neq j$, and $\frac{\langle\vec{v},\vec{w_j}\rangle}{||\vec{w_j}||^2}\langle\vec{w_j},\vec{w_j}\rangle=\langle\vec{v},\vec{w_j}\rangle$.

To show that the numbers $\vec{w}$ and $\vec{w'}$ are unique, suppose that there exists $\vec{w_1},\vec{w_2}\in W$ and $\vec{w_1'},\vec{w_2'}\in W^\perp$ satisfying $\vec{w_1}+\vec{w_1'}=\vec{w_2}+\vec{w_2'}$. Now, $\vec{w_1}-\vec{w_2}\in W$ and $\vec{w_1'}-\vec{w_2'}=-(\vec{w_1}-\vec{w_2})\in W^\perp$, which means the vector $\vec{w_1}-\vec{w_2}$ is in both W and $W^\perp$ (the product of any vector in W and any scalar constant is also in W). Notice, however, that this means $||\vec{w_1}-\vec{w_2}||=0$ by the definition of W and $W^\perp$, so $\vec{w_1}-\vec{w_2}=0$ or $\vec{w_1}=\vec{w_2}$, showing that such pair of numbers must be unique.

3. Let V be a finite dimensional transformation. Then for each transformation $T:V\to V$ there is a unique tranformation T^* satisfying $\langle T(\vec{x}),\vec{y}\rangle=\langle\vec{x},T^*(\vec{y})\rangle$ for all $\vec{x},\vec{y}\in V$.

Proof: Let n be the dimension of V, and denote $\{\vec{v}_1,\cdots,\vec{v}_n\}$ by an orthonormal basis of V. We use the fact that each linear transformation is uniquely determined by the values of $T(\vec{v}_1),\cdots,T(\vec{v})n)$. That is, for each n-tuples of vectors $\{\vec{w}_1.\cdots,\vec{w}_n\}$ there is a unique linear transformation T such that $T(\vec{v}_i)=\vec{w}_i$. Suppose that numbers $a_{ij}\in\mathbb{C},1\leq i,j\leq n$ are such that $T(\vec{v}_i)=\sum_{i=1}^{n}a_{ij}\vec{v}_j$, we have, for each i,k, $\langle T(\vec{v}_i),\vec{v}_k\rangle=\langle\sum_{i=1}^{n}a_{ij}\vec{v}_j,\vec{v}_k\rangle=a_{ik}$. Suppose that there is a linear transformation T^* such that $\langle T(\vec{x}),\vec{y}\rangle=\langle\vec{x},T^*(\vec{y})\rangle$ for all $\vec{x},\vec{y}\in V$. Let b_{ij} be numbers such that $T^*(\vec{v}_i)=\sum_{i=1}^{n}b_{ij}\vec{v}_j$ then we have $a_{ik}=\langle T(\vec{v}_i),\vec{v}_k\rangle=\langle(\vec{v}_i),T^*(\vec{v}_k)\rangle=\overline{\langle T^*(\vec{v}_k),\vec{v}_i\rangle}=\overline{\langle\sum_{i=1}^{n}b_{kj}\vec{v}_j,\vec{v}_i\rangle}=\overline{a_{ki}}$, therefore we must have $T^*(\vec{v}_i)=\sum_{i=1}^{n}b_{ij}\vec{v}_j=T^*(\vec{v}_i)=\sum_{i=1}^{n}\overline{b_{ji}}\vec{v}_j$. This uniquely defines T^*.

Conversely, let T^* be as defined, given T. From above we already have the relation $\langle T(\vec{v}_i), \vec{v}_k\rangle = \langle (\vec{v}_i), T^*(\vec{v}_k)\rangle$ for each pair of orthonormal basis. Let $\vec{x} = \sum_{i=1}^{n} x_i\vec{v}_i$ and $\vec{y} = \sum_{i=1}^{n} y_i\vec{v}_i$ then we have

$$\langle T(\vec{x}), \vec{y}\rangle = \langle T(\sum_{i=1}^{n} x_i\vec{v}_i), \sum_{i=1}^{n} y_i\vec{v}_i\rangle = \langle \sum_{i=1}^{n} x_i T(\vec{v}_i), \sum_{i=1}^{n} y_i\vec{v}_i\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\langle T(\vec{v}_i), \vec{v}_j\rangle$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} x_i\overline{y_j}\langle \vec{v}_i, T^*(\vec{v}_j)\rangle = \langle \sum_{i=1}^{n} x_i\vec{v}_i, \sum_{j=1}^{n} y_j T^*(\vec{v}_j)\rangle = \langle \sum_{i=1}^{n} x_i\vec{v}_i, T^*(\sum_{j=1}^{n} y_j\vec{v}_j)\rangle = \langle \vec{x}, T^*(\vec{y})\rangle$$

4. Let $A = [T]_\beta$ for some orthonormal basis β is a finite dimensional space V. Then $[T]_\beta^* = [T^*]_\beta$.

   Proof: Let our orthonormal basis be $\{\vec{v}_1, \cdots, \vec{v}_n\}$. This proof relies on the following fact: $A_{ij} = \langle T(\vec{v}_j), \vec{v}_i\rangle$. This is because for each j, $[T\vec{v}_j]_\beta = [T]_\beta[\vec{v}_j]_\beta = \mathrm{Col}_j(A)$ so $\langle T(\vec{v}_j), \vec{v}_i\rangle = \sum_{i=1}^{n} A_{kj}\langle \vec{v}_k, \vec{v}_i\rangle = A_{ij}$, as desired. Thus for each i, j we have $([T]_\beta^*)_{ij} = (A^*)_{ij} = \overline{A_{ij}^t} = \overline{A_{ji}} = \overline{\langle T(\vec{v}_i), \vec{v}_j\rangle} = \overline{\langle \vec{v}_i, T^*(\vec{v}_j)\rangle} = \langle T^*(\vec{v}_j), \vec{v}_i\rangle = ([T*]_\beta)_{ij}$.