2. The Dot Product, Norm, Angle and Orthogonal Projections in $\mathbb{R}^n$

2.1 Definition: For vectors $x, y \in \mathbb{R}^n$ we define the **dot product** of x and y to be

$$x \cdot y = y^T x = \sum_{i=1}^{n} x_i y_i \,.$$

2.2 Theorem: (*Properties of the Dot Product*) *For all $x, y, z \in \mathbb{R}^n$ and all $t \in \mathbb{R}$ we have*
(1) (Bilinearity) $(x+y) \cdot z = x \cdot z + y \cdot z$, $(tx) \cdot y = t(x \cdot y)$
$x \cdot (y+z) = x \cdot y + x \cdot z$, $x \cdot (ty) = t(x \cdot y)$,
(2) (Symmetry) $x \cdot y = y \cdot x$, *and*
(3) (Positive Definiteness) $x \cdot x \geq 0$ *with* $x \cdot x = 0$ *if and only if* $x = 0$.

Proof: The proof is left as an exercise.

2.3 Definition: For a vector $x \in \mathbb{R}^n$, we define the **length** (or **norm**) of x to be

$$\|x\| = \sqrt{x \cdot x} = \Big(\sum_{i=1}^{n} {x_i}^2 \Big)^{1/2}.$$

We say that x is a **unit vector** when $\|x\| = 1$.

2.4 Theorem: (*Properties of Length*) *Let $x, y \in \mathbb{R}^n$ and let $t \in \mathbb{R}$. Then*
(1) (Positive Definiteness) $\|x\| \geq 0$ *with* $\|x\| = 0$ *if and only if* $x = 0$,
(2) (Scaling) $\|tx\| = |t| \|x\|$,
(3) $\|x \pm y\|^2 = \|x\|^2 \pm 2(x \cdot y) + \|y\|^2$.
(4) (The Polarization Identities) $x \cdot y = \frac{1}{2}\big(\|x+y\|^2 - \|x\|^2 - \|y\|^2\big) = \frac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2\big)$,
(5) (The Cauchy-Schwarz Inequality) $|x \cdot y| \leq \|x\| \, \|y\|$ *with* $|x \cdot y| = \|x\| \, \|y\|$ *if and only if the set* $\{x, y\}$ *is linearly dependent, and*
(6) (The Triangle Inequality) $\|x + y\| \leq \|x\| + \|y\|$.

Proof: We leave the proofs of Parts (1), (2) and (3) as an exercise, and we note that (4) follows immediately from (3). To prove part (5), suppose first that $\{x, y\}$ is linearly dependent. Then one of x and y is a multiple of the other, say $y = tx$ with $t \in \mathbb{R}$. Then

$$|x \cdot y| = |x \cdot (tx)| = |t(x \cdot x)| = |t| \, \|x\|^2 = \|x\| \, \|tx\| = \|x\| \, \|y\|.$$

Suppose next that $\{x, y\}$ is linearly independent. Then for all $t \in \mathbb{R}$ we have $x + ty \neq 0$ and so

$$0 \neq \|x + ty\|^2 = (x + ty) \cdot (x + ty) = \|x\|^2 + 2t(x \cdot y) + t^2 \|y\|^2.$$

Since the quadratic on the right is non-zero for all $t \in \mathbb{R}$, it follows that the discriminant of the quadratic must be negative, that is

$$4(x \cdot y)^2 - 4\|x\|^2 \|y\|^2 < 0.$$

Thus $(x \cdot y)^2 < \|x\|^2 \|y\|^2$ and hence $|x \cdot y| < \|x\| \, \|y\|$. This proves part (5).

Using part (5) note that

$$\begin{aligned} \|x + y\|^2 = \|x\|^2 + 2(x \cdot y) + \|y\|^2 &\leq \|x + y\|^2 + 2|x \cdot y| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \, \|y\| + \|y\|^2 = \big(\|x\| + \|y\|\big)^2 \end{aligned}$$

and so $\|x + y\| \leq \|x\| + \|y\|$, which proves Part (6).

2.5 Definition: For points $a, b \in \mathbb{R}^n$, we define the **distance** between a and b to be

$$\mathrm{dist}(a,b) = \|b-a\|.$$

2.6 Theorem: *(Properties of Distance) Let $a, b, c \in \mathbb{R}^n$. Then*
(1) (Positive Definiteness) $\mathrm{dist}(a,b) \geq 0$ with $\mathrm{dist}(a,b) = 0$ if and only if $a = b$,
(2) (Symmetry) $\mathrm{dist}(a,b) = \mathrm{dist}(b,a)$, and
(3) (The Triangle Inequality) $\mathrm{dist}(a,c) \leq \mathrm{dist}(a,b) + \mathrm{dist}(b,c)$.

Proof: The proof is left as an exercise.

2.7 Definition: For nonzero vectors $0 \neq x, y \in \mathbb{R}^n$, we define the **angle** between x and y to be

$$\theta(x,y) = \cos^{-1}\left(\frac{x \bullet y}{\|x\|\,\|y\|}\right) \in [0, \pi].$$

Note that $\theta(x,y) = \frac{\pi}{2}$ if and only if $x \bullet y = 0$. For vectors $x, y \in \mathbb{R}^n$, we say that x and y are **orthogonal** when $x \bullet y = 0$.

2.8 Theorem: *(Properties of Angle) Let $0 \neq x, y \in \mathbb{R}^n$. Then*

(1) $\theta(x,y) \in [0,\pi]$ with $\begin{cases} \theta(x,y) = 0 \text{ if and only if } y = tx \text{ for some } t > 0, \text{ and} \\ \theta(x,y) = \pi \text{ if and only if } y = tx \text{ for some } t < 0, \end{cases}$
(2) (Symmetry) $\theta(x,y) = \theta(y,x)$,
(3) (Scaling) $\theta(tx,y) = \theta(x,ty) = \begin{cases} \theta(x,y) & \text{if } 0 < t \in \mathbb{R}, \\ \pi - \theta(x,y) & \text{if } 0 > t \in \mathbb{R}, \end{cases}$
(4) (The Law of Cosines) $\|y-x\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\,\|y\| \cos\theta(x,y)$,
(5) (Pythagoras' Theorem) $\theta(x,y) = \frac{\pi}{2}$ if and only if $\|y-x\|^2 = \|x\|^2 + \|y\|^2$, and
(6) (Trigonometric Ratios) if $(y-x) \bullet x = 0$ then $\cos\theta(x,y) = \frac{\|x\|}{\|y\|}$ and $\sin\theta(x,y) = \frac{\|y-x\|}{\|y\|}$.

Proof: The Law of Cosines follows from the identity $\|y-x\|^2 = \|y\|^2 - 2(y \bullet x) + \|x\|^2$ and the definition of $\theta(x,y)$. Pythagoras' Theorem is a special case of the Law of Cosines. We Prove Part (6). Let $0 \neq x, y \in \mathbb{R}^n$ and write $\theta = \theta(x,y)$. Suppose that $(y-x) \bullet x = 0$. Then we have $y \bullet x - x \bullet x = 0$ so that $x \bullet y = \|x\|^2$, and so we have

$$\cos\theta = \frac{x \bullet y}{\|x\|\,\|y\|} = \frac{\|x\|^2}{\|x\|\,\|y\|} = \frac{\|x\|}{\|y\|}\,.$$

By Pythagoras' Theorem we have $\|x\|^2 + \|y-x\|^2 = \|y\|^2$ so that $\|y\|^2 - \|x\|^2 = \|y-x\|^2$, and so

$$\sin^2\theta = 1 - \cos^2\theta = 1 - \frac{\|x\|^2}{\|y\|^2} = \frac{\|y\|^2 - \|x\|^2}{\|y\|^2} = \frac{\|y-x\|^2}{\|y\|^2}\,.$$

Since $\theta \in [0,\pi]$ we have $\sin\theta \geq 0$, and so taking the square root on both sides gives

$$\sin\theta = \frac{\|y-x\|}{\|y\|}\,.$$

2.9 Definition: For points $a, b, c \in \mathbb{R}^n$ with $a \neq b$ and $b \neq c$ we define

$$\angle abc = \theta(a-b, c-b).$$

2.10 Definition: Let $S \subseteq \mathbb{R}^n$ be a subset. We define the **orthogonal complement** of S in $\mathbb{R}^n$ to be
$$S^\perp = \big\{x \in \mathbb{R}^n \big| x \cdot u = 0 \text{ for all } u \in S\big\}.$$

2.11 Theorem: (*Properties of the Orthogonal Complement*) *Let $S \subseteq \mathbb{R}$ be a subset, let $U \subseteq \mathbb{R}^n$ be a subspace, and let $A \in M_{k\times n}(\mathbb{R})$. Then*
(1) $S^\perp$ is a subspace of $\mathbb{R}^n$,
(2) If $U = \text{Span}\,(S)$ then $U^\perp = S^\perp$,
(3) $(\text{Row}A)^\perp = \text{Null}A$.
(4) $\dim(U) + \dim(U^\perp) = n$
(5) $U \oplus U^\perp = \mathbb{R}^n$,
(6) $(U^\perp)^\perp = U$,
(7) $(\text{Null}A)^\perp = \text{Row}A$.

Proof: Note that $0 \in S^\perp$ since $0 \cdot u = 0$ for all $u \in S$. If $x, y \subset S^\perp$ so that $x \cdot u = 0$ and $y \cdot u = 0$ for all $u \in S$ then we have $(x + y) \cdot u = x \cdot u + y \cdot u = 0$ for all $u \in S$ and so $x + y \in S^\perp$. If $x \in S^\perp$ so that $x \cdot u = 0$ for all $u \in S$ and $t \in \mathbb{R}$ then we have $(tx) \cdot u = t(x \cdot u) = 0$ for all $u \in S$ and so $tu \in S^\perp$. This shows that $S^\perp$ is a subspace of $\mathbb{R}^n$, proving Part (1).

To prove Part (2), let $T = \big\{x \in \mathbb{R}^n \big| x \cdot u = 0 \text{ for all } u \in S\big\}$. It is clear that $U^\perp \subseteq T$. Let $x \in T$. Let $u \in U = \text{Span}\,(S)$, say $u = \sum\limits_{i=1}^n t_i u_i$ with each $t_i \in \mathbb{R}$ and each $u_i \in S$. Then $x \cdot u = x \cdot \sum\limits_{i=1}^n t_i u_i = \sum\limits_{i=1}^n t_i (x \cdot u_i) = 0$. Thus $x \in U^\perp$ and so we have $T \subseteq U^\perp$.

To prove Part (3), let $A = \begin{pmatrix} {v_1}^T \\ \vdots \\ {v_k}^T \end{pmatrix}$ where each $v_i \in \mathbb{R}^n$. Note that $Ax = \begin{pmatrix} x \cdot v_1 \\ \vdots \\ x \cdot v_k \end{pmatrix}$ so we have $x \in \text{Null}A \iff x \cdot v_i = 0 \text{ for all } i \iff x \in \text{Span}\,\{v_1, v_2, \cdots, v_k\}^\perp = (\text{Row}A)^\perp$ by part (2).

Part (4) follows from Part (3) since if we choose A so that $\text{Row}A = U$ then we have $\dim(U) + \dim(U^\perp) = \dim \text{Row}A + \dim(\text{Row}A)^\perp = \dim \text{Row}A + \dim \text{Null}A = n$.

To prove Part (5), in light of Part (4), it suffices to show that $U \cap U^\perp = \{0\}$. Let $x \in U \cap U^\perp$. Since $x \in U^\perp$ we have $x \cdot u = 0$ for all $u \in U$. In particular, since $x \in U$ we have $x \cdot x = 0$, and hence $x = 0$. Thus $U \cap U^\perp = \{0\}$ and so $U \oplus U^\perp = \mathbb{R}^n$.

To prove Part (6), let $x \in U$. By the definition of $U^\perp$ we have $x \cdot v = 0$ for all $v \in U^\perp$. By the definition of $(U^\perp)^\perp$ we see that $x \in (U^\perp)^\perp$. Thus $U \subseteq (U^\perp)^\perp$. By part (4) we know that $\dim U + \dim U^\perp = n$ and also that $\dim U^\perp + \dim(U^\perp)^\perp = n$. It follows that $\dim U = n - \dim U^\perp = \dim(U^\perp)^\perp$. Since $U \subseteq (U^\perp)^\perp$ and $\dim U = \dim(U^\perp)^\perp$ we have $U = (U^\perp)^\perp$, as required.

By Parts (3) and (6) we have $(\text{Null}A)^\perp = \big((\text{Row}A)^\perp\big)^\perp = \text{Row}A$, proving Part (7).

2.12 Definition: For a subspace $U \subseteq \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, we define the **orthogonal projection** of x onto U, denoted by $\text{Proj}_U(x)$, as follows. Since $\mathbb{R}^n = U \oplus U^\perp$, we can choose unique vectors $u, v \in \mathbb{R}^n$ with $u \in U$, $v \in U^\perp$ and $x = u + v$. We then define

$$\text{Proj}_U(x) = u.$$

Note that since $U = (U^\perp)^\perp$, for u and v as above we have $\text{Proj}_{U^\perp}(x) = v$. When $y \in \mathbb{R}^n$ and $U = \text{Span}\{y\}$, we also write $\text{Proj}_y(x) = \text{Proj}_U(x)$ and $\text{Proj}_{y^\perp}(x) = \text{Proj}_{U^\perp}(x)$.

2.13 Theorem: *Let $U \subseteq \mathbb{R}^n$ be a subspace and let $x \in \mathbb{R}^n$. Then $\text{Proj}_U(x)$ is the unique point in U which is nearest to x.*

Proof: Let $u, v \in \mathbb{R}^n$ with $u \in U$, $v \in V$ and $u + v = x$ so that $\text{Proj}_U(x) = u$. Let $w \in U$ with $w \neq u$. Since $v \in U^\perp$ and $u, w \in U$ we have $v \cdot u = v \cdot w = 0$ and so $v \cdot (w - u) = v \cdot w - v \cdot u = 0$. Thus we have

$$\begin{aligned} \|x - w\|^2 &= \|u + v - w\|^2 = \|v - (w - u)\|^2 = \big(v - (w - u)\big) \cdot \big(v - (w - u)\big) \\ &= \|v\|^2 - 2\,v \cdot (w - u) + \|w - u\|^2 = \|v\|^2 + \|w - u\|^2 = \|x - u\|^2 + \|w - u\|^2 . \end{aligned}$$

Since $w \neq u$ we have $\|w - u\| > 0$ and so $\|x - w\|^2 > \|x - u\|^2$. Thus $\|x - w\| > \|x - u\|$, that is $\text{dist}(x, w) > \text{dist}(x, u)$, so u is the vector in U nearest to x, as required.

2.14 Theorem: *For any matrix $A \in M_{n \times l}(\mathbb{R})$ we have $\text{Null}(A^TA) = \text{Null}(A)$ and $\text{Col}(A^TA) = \text{Col}(A^T)$ so that $\text{nullity}(A^TA) = \text{nullity}(A)$ and $\text{rank}(A^TA) = \text{rank}(A)$.*

Proof: If $x \in \text{Null}(A)$ then $Ax = 0$ so $A^TAx = 0$ hence $x \in \text{Null}(A^TA)$. This shows that $\text{Null}(A) \subseteq \text{Null}(A^TA)$. If $x \in \text{Null}(A^TA)$ then $A^TAx = 0$ so that $x^TA^TAx = 0$, that is $\|Ax\|^2 = (Ax)^T(Ax) = x^TA^TAx = 0$, and so $Ax = 0$, that is $x \in \text{Null}(A)$. This shows that $\text{Null}(A^TA) \subseteq \text{Null}(A)$. Thus we have $\text{Null}(A^TA) = \text{Null}(A)$. It then follows that

$$\text{Col}(A^T) = \text{Row}(A) = \text{Null}(A)^\perp = \text{Null}(A^TA)^\perp = \text{Row}(A^TA) = \text{Col}\big((A^TA)^T\big) = \text{Col}(A^TA).$$

2.15 Theorem: *(Orthogonal Projection Formula) Let $A \in M_{n \times l}(\mathbb{R})$, let $U = \text{Col}(A)$ and let $x \in \mathbb{R}^n$. Then*

(1) the matrix equation $A^TA\,t = A^Tx$ has a solution $t \in \mathbb{R}^l$, and for any solution t we have

$$\text{Proj}_U(x) = At,$$

(2) if $\text{rank}(A) = l$ then A^TA is invertible and

$$\text{Proj}_U(x) = A(A^TA)^{-1}A^Tx.$$

Proof: Note that $U^\perp = (\text{Col}A)^\perp = \text{Row}(A^T)^\perp = \text{Null}(A^T)$. Let $u, v \in \mathbb{R}^n$ with $u \in U$, $v \in U^\perp$ and $u + v = x$ so that $\text{Proj}_U(x) = u$. Since $u \in U = \text{Col}A$ we can choose $t \in \mathbb{R}^l$ so that $u = At$. Then we have $x = u + v = At + v$. Multiply by A^T to get $A^Tx = A^TA\,t + A^Tv$. Since $v \in U^\perp = \text{Null}(A^T)$ we have $A^Tv = 0$ so $A^TA\,t = A^Tx$. Thus the matrix equation $A^TA\,t = A^Tx$ does have a solution $t \in \mathbb{R}^l$.

Now let $t \in \mathbb{R}^l$ be any solution to $A^TA\,t = A^tx$. Let $u = At$ and $v = x - u$. Note that $x = u + v$, $u = At \in \text{Col}(A) = U$, and $A^Tv = A^T(x - u) = A^T(x - At) = A^Tx - A^TA\,t = 0$ so that $v \in \text{Null}(A^T) = U^\perp$. Thus $\text{Proj}_U(x) = u = At$, proving part (1).

Now suppose that $\text{rank}(A) = l$. Since $A^TA \in M_{l \times l}(\mathbb{R})$ with $\text{rank}(A^TA) = \text{rank}(A) = l$, the matrix A^TA is invertible. Since A^TA is invertible, the unique solution $t \in \mathbb{R}^l$ to the matrix equation $A^TA\,t = A^Tx$ is the vector $t = (A^TA)^{-1}A^Tx$, and so from Part (1) we have $\text{Proj}_U(x) = At = A(A^TA)^{-1}A^Tx$, proving Part (2).

2.16 Definition: For a subset $\mathcal{A} \subseteq \mathbb{R}^n$, we say that $\mathcal{A}$ is **orthogonal** when $x \cdot y = 0$ for all $x, y \in \mathcal{A}$ with $x \neq y$. We say that $\mathcal{A}$ is **orthonormal** when $\mathcal{A}$ is orthogonal and $\|x\| = 1$ for every $x \in \mathcal{A}$.

2.17 Note: Let $u_1, \cdots, u_l \in \mathbb{R}^n$, let $\mathcal{A} = \{u_1, \cdots, u_l\}$ and let $A = (u_1, \cdots, u_l) \in M_{n\times l}(\mathbb{R})$. Then

$$A^T A = \begin{pmatrix} {u_1}^T \\ \vdots \\ {u_l}^T \end{pmatrix} (u_1, \cdots, u_l) = \begin{pmatrix} u_1 \cdot u_1 & u_1 \cdot u_2 & \cdots & u_1 \cdot u_l \\ u_2 \cdot u_1 & u_2 \cdot u_2 & \cdots & u_2 \cdot u_l \\ \vdots & \vdots & & \vdots \\ u_l \cdot u_1 & u_l \cdot u_2 & \cdots & u_l \cdot u_l \end{pmatrix}.$$

It follows that $\mathcal{A}$ is orthogonal if and only if $A^T A$ is diagonal, in which case we have $A^T A = \mathrm{diag}(\|u_1\|^2, \|u_2\|^2, \cdots, \|u_l\|^2)$, and $\mathcal{A}$ is orthonormal if and only if $A^T A = I$.

2.18 Note: Recall that when $\mathcal{A} = \{u_1, u_2, \cdots, u_l\}$ is a basis for a vector space U over a field F, a vector $x \in U$ can be written uniquely as a linear combination $x = \sum\limits_{i=1}^{l} t_i u_i$ with each $t_i \in F$, and then we define the coordonate vector of x with respect to $\mathcal{A}$ to be

$$[x]_{\mathcal{A}} = t = (t_1, t_2, \cdots, t_l)^T \in F^l.$$

2.19 Theorem: *Let $u_1, u_2, \cdots, u_l \in \mathbb{R}^n$, let $\mathcal{A} = \{u_1, u_2, \cdots, u_l\}$, let $U = \mathrm{Span}\,\mathcal{A}$, and let $x \in \mathbb{R}^n$ Then*

(1) if $\mathcal{A}$ is orthogonal with each $u_i \neq 0$ then $\mathcal{A}$ is a basis for U and

$$[x]_{\mathcal{A}} = \left(\frac{x \cdot u_1}{\|u_1\|^2}, \frac{x \cdot u_2}{\|u_2\|^2}, \cdots, \frac{x \cdot u_l}{\|u_l\|^2} \right)^T, \text{ and}$$

(2) if $\mathcal{A}$ is orthonormal then $\mathcal{A}$ is a basis for U and

$$[x]_{\mathcal{A}} = \Big(x \cdot u_1, x \cdot u_2, \cdots, x \cdot u_l \Big)^T.$$

Proof: Suppose $\mathcal{A}$ is orthogonal with each $u_i \neq 0$. Let $A = (u_1, u_2, \cdots, u_l) \in M_{n\times l}(\mathbb{R})$ so that $U = \mathrm{Col}(A)$. Since $\mathcal{A}$ is orthogonal we have $A^T A = \mathrm{diag}(\|u_1\|^2, \cdots, \|u_l\|^2)$. Since each $u_i \neq 0$ we see that $A^T A$ is invertible. Since $\mathrm{rank}(A) = \mathrm{rank}(A^T A) = l$, the columns of A are linearly independent, so $\mathcal{A}$ is a basis for U. Write x as a linear combination $x = \sum\limits_{i=1}^{l} t_i u_i = At$ with $t \in \mathbb{R}^l$. Then we have $A^T x = A^T A\, t$ and so

$$\begin{aligned} [x]_{\mathcal{A}} = t = (A^T A)^{-1} A^T x &= \mathrm{diag}\Big(\|u_1\|^2, \cdots, \|u_l\|^2 \Big)^{-1} \begin{pmatrix} {u_1}^T \\ \vdots \\ {u_l}^T \end{pmatrix} x \\ &= \mathrm{diag}\Big(\tfrac{1}{\|u_1\|^2}, \cdots, \tfrac{1}{\|u_l\|^2} \Big) \begin{pmatrix} x \cdot u_1 \\ \vdots \\ x \cdot u_l \end{pmatrix} = \begin{pmatrix} \dfrac{x \cdot u_1}{\|u_1\|^2} \\ \vdots \\ \dfrac{x \cdot u_l}{\|u_l\|^2} \end{pmatrix} \end{aligned}$$

This proves part (1), and part (2) follows immediately from part (1).

2.20 Theorem: *Let $u_1, u_2, \cdots, u_l \in \mathbb{R}^n$, let $\mathcal{A} = \{u_1, u_2, \cdots, u_l\}$, let $U = \operatorname{Span} \mathcal{A}$, and let $x \in \mathbb{R}^n$. Then*

(1) if $\mathcal{A}$ is orthogonal with each $u_i \neq 0$ then we have

$$\operatorname{Proj}_U(x) = \sum_{i=1}^{l} \frac{x \cdot u_i}{\|u_i\|^2}\, u_i,$$

(2) if $\mathcal{A}$ is orthonormal then

$$\operatorname{Proj}_U(x) = \sum_{i=1}^{l} (x \cdot u_i)\, u_i.$$

Proof: Suppose that $\mathcal{A}$ is orthogonal with each $u_i \neq 0$. Let $A = (u_1, u_2, \cdots, u_l) \in M_{n \times l}(\mathbb{R})$ so that $U = \operatorname{Col}(A)$ and we have $A^T A = \operatorname{diag}(\|u_1\|^2, \cdots, \|u_l\|^2)$, which is invertible. Then

$$\operatorname{Proj}_U(x) = A\,(A^T A)^{-1} A^T x = \Big(u_1, \cdots, u_l\Big) \operatorname{diag}\left(\frac{1}{\|u_1\|^2}, \cdots, \frac{1}{\|u_l\|^2}\right) \begin{pmatrix} {u_1}^T \\ \vdots \\ {u_l}^T \end{pmatrix} x$$

$$= \left(\frac{u_1}{\|u_1\|^2}, \cdots, \frac{u_l}{\|u_l\|^2}\right) \begin{pmatrix} x \cdot u_1 \\ \vdots \\ x \cdot u_l \end{pmatrix} = \frac{x \cdot u_1}{\|u_1\|^2}\, u_1 + \cdots + \frac{x \cdot u_l}{\|u_l\|^2}\, u_l.$$

This proves part (1), and part (2) follows immediately from part (1).