Chapter 2

Vectors and Vector Spaces

2.1 (informal proof)

Essentially, we want to prove:

$$x_1 v_{11} + \cdots + x_{n+1} v_{1\,n+1} = 0$$
$$\vdots$$
$$x_1 v_{n1} + \cdots + x_{n+1} v_{n\,n+1} = 0$$

has non-zero solution for $(x_1 \cdots x_{n+1})$.
From alegebra, we know that: a system of n-equations but with n+1 unknowns is solvable and has non-trivial solution
Hence $(v_1, \cdots v_{n+1})$ are linearly dependent. The maximum number of linearly independent set is n.

2.2.

(1). Proof of 2.10: (informal)
Suppose W is a spanning set of $\mathcal{W}$ and each vector in W is linearly independent to each other (i.e. W is a basis for $\mathcal{W}$)

We form the spanning set V of $\mathcal{V}$ in the following manner:
$$V = W \cup (V - W)$$
Here $(V - W)$ spans $\mathcal{V} - \mathcal{W}$. (informal)
Therefore, $W \subseteq V$.
By definition (of linear independency), $\dim(W) \leq \dim(V)$.

(2) (informal)
We know that we can find spanning and linearly independent sets of $\mathcal{V} \cap \mathcal{W}$, $\mathcal{V}$ and $\mathcal{W}$ such that: $\text{span}(V \cap W) = \mathcal{V} \cap \mathcal{W}$.
$\text{span}(V) = \mathcal{V}$, $(V, W$ are vector sets$)$
$\text{span}(W) = \mathcal{W}$.
Then, since $V \cap W \subseteq V$ and $V \cap W \subseteq W$, by definition
$$\dim(\mathcal{V} \cap \mathcal{W}) \leq \min(\dim(\mathcal{V}), \dim(\mathcal{W}))$$

2.3.

a) Suppose all vectors $\mathbb{R}^2$, we choose the subset of vectors:

$$V = \{(x,0),\ x \in \mathbb{R}\} \cup \{(0,y),\ y \in \mathbb{R}\}.$$

Then V with axpy operation defined on it is not a vector space, since its not closed under axpy

b) Similar to (a), let $\mathcal{V}_1 = \{(x,0)\ x \in \mathbb{R}\}$, $\mathcal{V}_2 = \{(0,y) \cdot y \in \mathbb{R}\}$, then $\mathcal{V}_1$ and $\mathcal{V}_2$ (together with the axpy operation) are vector spaces. However, $\mathcal{V}_1 \cup \mathcal{V}_2 = \{(x,0),\ x \in \mathbb{R}\} \cup \{(0,y),\ y \in \mathbb{R}\}$ is not vector space, since it's not closed under axpy operation.

2.4

(1) Proof of 2.15: (informal)

Let $\mathcal{V}$ and $\mathcal{W}$ be in $\mathbb{R}^n$. We know that $\{e_1, \cdots, e_n\}$ spans $\mathbb{R}^n$. And we let $V \subseteq \{e_1, \cdots, e_n\}$, $W \subseteq \{e_1, \cdots, e_n\}$ be such that $\mathrm{span}(V) = \mathcal{V}$, $\mathrm{span}(W) = \mathcal{W}$.

And, we know that $\mathrm{span}(V \cup W) = \mathcal{V} \oplus \mathcal{W}$.

By relationship of set cardinality:

$$|V \cup W| = |V| + |W| - |V \cap W|$$

The dimension relationship is:

$$\dim(\mathcal{V} \oplus \mathcal{W}) = \dim(\mathcal{V}) + \dim(\mathcal{W}) - \dim(\mathcal{V} \cap \mathcal{W})$$

(2) Proof of 2.16: (informal)

Since $\mathcal{V}_i$ are essentially disjoint, we let V_i be the basis for $\mathcal{V}_i$, and get:

$$V = V_1 \cup V_2 \cdots \cup V_n \quad \text{and} \quad V_i \cap V_j = \phi,\ i \neq j$$

Applying 2.15 iteratively, we get:

$$\dim(\mathcal{V}) = \dim(\mathcal{V}_1) + \cdots + \dim(\mathcal{V}_n) = \sum_{i=1}^{n} \dim(\mathcal{V}_i)$$

2.5 Proof of 2.19: (informal)

Suppose V is a basis for $\mathcal{V}$, W is a basis for $\mathcal{W}$. then we form the basis of $\mathcal{V} \otimes \mathcal{W}$ by:

1. appending 0's to V-vectors
2. preppendinding 0's to W-vectors

Then, the new set $\left\{ \begin{pmatrix} V \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} W \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$ is a basis for $\mathcal{V} \otimes \mathcal{W}$.

By definition, $\dim(\mathcal{V} \otimes \mathcal{W}) = \dim(\mathcal{V}) + \dim(\mathcal{W})$.

2.6 $\langle x, v_i \rangle$

$= \langle \sum_j b_j V_j \cdot V_i \rangle$

$= \sum_j \langle b_j V_j, V_i \rangle = \sum_j b_j \langle V_j, V_i \rangle$

$= b_i$ (since, for $j \neq i$, $\langle V_j, V_i \rangle = 0$; for $j = i$, $\langle V_j, V_i \rangle = 1$)

2.7 Written in inner-product format:

$LHS = 2 \sum_i x_i^2 + 2 \sum_i y_i^2$

$RHS = \sum_i (x_i + y_i)^2 + \sum_i (x_i - y_i)^2$

$= \sum_i (x_i^2 + 2x_i y_i + y_i^2) + \sum_i (x_i^2 - 2x_i y_i + y_i^2)$

$= 2 \sum_i x_i^2 + 2 \sum_i y_i^2$

$= LHS$

2.8 Example:

let $x = (9, 25)$, $y = (16, 144)$ then

$$\rho(x+y) = \left(\sqrt{25} + \sqrt{169}\right)^2 = 324$$

$$\rho(x) = \left(\sqrt{9} + \sqrt{25}\right)^2 = 64, \quad \rho(y) = \left(\sqrt{16} + \sqrt{144}\right)^2 = 256$$

Hence, $\rho(x+y) > \rho(x) + \rho(y)$

This violates the triangle inequality in the definition of a norm.

So $\rho(x) = \left(\sum_i |x_i|^{\frac{1}{2}}\right)^2$ is not a norm.

2.9 Let $V_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $V_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $\left(L_1 \text{ norm: } \|x\| = \sum_i |x_i| \right)$

Then $\|V_1\|_1 = 2$, $\|V_2\|_1 = 2$, $\|V_1 + V_2\|_1 = 2$, $\|V_1 - V_2\|_1 = 2$

In the parallelogram equation: LHS $= 8$, RHS $= 4$

LHS $\neq$ RHS

So, L_1 norm is not induced by inner product.

2.10 Proof of 2.34:

For the left-hand side, let $\|x\|_\infty = \max_i |x_i|$

It's easy to see that $\max_i |x_i| \leq \left(\sum_i |x_i|^2\right)^{\frac{1}{2}}$

For the right-hand side, to prove $\|x\|_2 \leq \|x\|_1$,

it's equivalent to prove $\left(\sum_i |x_i|^2\right)^{\frac{1}{2}} \leq \sum_i |x_i|$

$$\Leftrightarrow \quad \sum_i |x_i|^2 \leq \left(\sum_i |x_i|\right)^2$$

$$\Leftrightarrow \quad \sum_i |x_i|^2 \leq \sum_i |x_i|^2 + 2 \sum_{i=1}^{n} \sum_{j=i}^{n} |x_i||x_j|$$

which is obvious. Hence the inequality is proved.

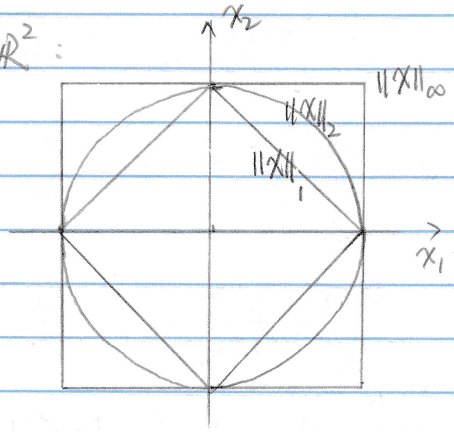
2.10 (continue)

With equality, we need one element of X to be non-zero, and the rest elements to be all zeros.

In this way, we have: ① $\max|x_i| = \left(\sum_i |x_i|^2\right)^{\frac{1}{2}}$

② $\sum_i |x_i|^2 = \sum_i |x_i|^2 + 2\sum_{i=1}^{n}\sum_{j=i}^{n} |x_i||x_j|$

Graphically, in $\mathbb{R}^2$:



① All points on the diamond have the same $\|x\|_1$

② All points on the circle have the same $\|x\|_2$

③ All points on the square have the same $\|x\|_\infty$

We can see that only at $x_1 = 0$ or $x_2 = 0$, the three norms are equal. And, the bounds are "sharp"

2.11

(a) Proof:

To prove $\langle x, y\rangle \le \|x\|_p\|y\|_q$, we consider proving: $\frac{\langle x, y\rangle}{\|x\|_p\|y\|_q} \le 1$

Notice that $p+q = pq \Leftrightarrow \frac{1}{p}+\frac{1}{q} = 1$, we relate this to convex functions and draft the following proof.

Suppose: $a_i \le |a_i| = e^{\frac{s_i}{p}}$, $b_i \le |b_i| = e^{\frac{t_i}{q}}$ for some s_i and t_i.
(a_i and b_i are some arbitrary number). Then

$$a_i b_i \le |a_i||b_i| = e^{\frac{s_i}{p}+\frac{t_i}{q}} \quad ①$$

Since e^x is convex, we have $\exp\{\lambda x_1 + (1-\lambda)x_2\} \le \lambda e^{x_1} + (1-\lambda)e^{x_2}$.
Plugging this relation into ①, we get:

$$a_i b_i \le |a_i||b_i| = \exp\left\{\frac{1}{p}s_i + \frac{1}{q}t_i\right\} \le \frac{1}{p}e^{s_i} + \frac{1}{q}e^{t_i}$$

~~Pe~~ plug in $e^{s_i} = |a_i|^p$, $e^{t_i} = |b_i|^q$, we have:

$$a_i b_i \le |a_i||b_i| \le \frac{1}{p}|a_i|^p + \frac{1}{q}|b_i|^q$$

Now, since a_i and b_i are arbitrary, we let $|a_i| = \frac{|x_i|}{\|x\|_p}$, $|b_i| = \frac{|y_i|}{\|y\|_q}$
Then:

$$a_i b_i \le |a_i||b_i| \le \frac{1}{p}\frac{|x_i|^p}{\|x\|_p^p} + \frac{1}{q}\frac{|y_i|^q}{\|y\|_q^q}$$

$$\Leftrightarrow \frac{x_i}{\|x\|_p}\frac{y_i}{\|y\|_q} \le \frac{|x_i|}{\|x\|_p}\frac{|y_i|}{\|y\|_q} \le \frac{1}{p}\frac{|x_i|^p}{\|x\|_p^p} + \frac{1}{q}\frac{|y_i|^q}{\|y\|_q^q}$$

$$\Rightarrow \sum_i \frac{x_i y_i}{\|x\|_p\|y\|_q} \le \sum_i \frac{|x_i||y_i|}{\|x\|_p\|y\|_q} \le \frac{1}{p}\frac{\sum_i|x_i|^p}{\|x\|_p^p} + \frac{1}{q}\frac{\sum_i|y_i|^q}{\|y\|_q^q}$$

remember: $\|x\|_p = \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$

$\Rightarrow \|x\|_p^p = \sum_i |x_i|^p$

$$\Rightarrow \frac{\langle x, y\rangle}{\|x\|_p\|y\|_q} \le \frac{\langle |x|, |y|\rangle}{\|x\|_p\|y\|_q} \le \frac{1}{p}\frac{\|x\|_p^p}{\|x\|_p^p} + \frac{1}{q}\frac{\|y\|_q^q}{\|y\|_q^q}$$

$$\Rightarrow \frac{\langle x, y\rangle}{\|x\|_p\|y\|_q} \le \frac{\langle |x|, |y|\rangle}{\|x\|_p\|y\|_q} \le \frac{1}{p} + \frac{1}{q} = 1$$

Hence: $\langle x, y\rangle \le \langle |x|, |y|\rangle \le \|x\|_p\|y\|_q$
The Hölder's inequality is proved.

(b). Proof:
First we make transformations to $\|x+y\|_p^p$:

$$\|x+y\|_p^p = \sum_i |x_i+y_i|^p = \sum_i |x_i+y_i|^{p-1}|x_i+y_i|$$

$$\le \sum_i |x_i+y_i|^{p-1}(|x_i|+|y_i|) \quad \text{by } |x_i+y_i| \le |x_i|+|y_i|$$

$$= \sum_i |x_i+y_i|^{p-1}|x_i| + \sum_i |x_i+y_i|^{p-1}|y_i| \quad (1)$$

Let's denote $\begin{pmatrix} |x_1+y_1|^{p-1} \\ \vdots \\ |x_n+y_n|^{p-1} \end{pmatrix}$ as vector $|x+y|^{p-1}$,

And let $q = \frac{p}{p-1}$. (Solved from $p+q = pq \Rightarrow q = \frac{p}{p-1}$)
Then we can apply Hölder's inequality to each of the right-hand side terms in (1):

$$\sum_i |x_i+y_i|^{p-1}|x_i| = \langle |x+y|^{p-1}, |x| \rangle$$

$$\leq \| |x+y|^{p-1} \|_q \, \| |x| \|_p$$

$$= \left(\sum_i |x_i+y_i|^{(p-1)q}\right)^{\frac{1}{q}} \left(\sum_i |x_i|^p\right)^{\frac{1}{p}}$$

$$\sum_i |x_i+y_i|^{p-1}|y_i| = \langle |x+y|^{p-1}, |y| \rangle$$

$$\leq \| |x+y|^{p-1} \|_q \, \| |y| \|_p$$

$$= \left(\sum_i |x_i+y_i|^{(p-1)q}\right)^{\frac{1}{q}} \left(\sum_i |y_i|^p\right)^{\frac{1}{p}}$$

Plugging them into ①:

$$\sum_i |x_i+y_i|^p \leq \left(\sum_i |x_i+y_i|^{(p-1)q}\right)^{\frac{1}{q}} \left[\left(\sum_i |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_i |y_i|^p\right)^{\frac{1}{p}}\right]$$

Also, plug in $q = \frac{p}{p-1}$, we have $(p-1)q = p$, $\frac{1}{q} = 1 - \frac{1}{p}$:

$$\sum_i |x_i+y_i|^p \leq \left(\sum_i |x_i+y_i|^p\right)^{1-\frac{1}{p}} \left[\left(\sum_i |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_i |y_i|^p\right)^{\frac{1}{p}}\right]$$

$$\Rightarrow \left(\sum_i |x_i+y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_i |y_i|^p\right)^{\frac{1}{p}}$$

$$\Rightarrow \|x+y\|_p \leq \|x\|_p + \|y\|_p$$

The Minkowski's inequality is proved.

2.12 ① By definition, if $x \neq y$, then $\Delta(x,y) = \|x-y\| \geq 0$.

② $\Delta(x,y) = \|x-y\| = \|y-x\| = \Delta(y,x)$, by properties of norm

③ $\Delta(x,y) = \|x-y\| = \|x-z+z-y\| \leq \|x-z\| + \|z-y\| = \Delta(x,z) + \Delta(z,y)$

So, $\Delta(x,y) = \|x-y\|$ is a metric, by the definition of a metric

2.13 Let's consider the right-hand side:

$$\hat{y} = \frac{\langle x, y\rangle}{\|x\|^2} x \qquad r = y - \frac{\langle x, y\rangle}{\|x\|^2} x$$

then: $\|\hat{y}\|^2 + \|r\|^2$

$$= \sum_i \left(\frac{\langle x,y\rangle}{\|x\|^2} x_i\right)^2 + \sum_i \left(y_i - \frac{\langle x,y\rangle}{\|x\|^2} x_i\right)^2$$

$$= \frac{\langle x,y\rangle^2}{\|x\|^4} \sum_i x_i^2 + \sum_i \left(y_i^2 - 2\frac{\langle x,y\rangle}{\|x\|^2} x_i y_i + \frac{\langle x,y\rangle^2}{\|x\|^4} x_i^2\right)$$

$$= \frac{\langle x,y\rangle^2}{\|x\|^4}\|x\|^2 + \sum_i y_i^2 - 2\frac{\langle x,y\rangle}{\|x\|^2}\sum_i x_i y_i + \frac{\langle x,y\rangle^2}{\|x\|^4}\sum_i x_i^2$$

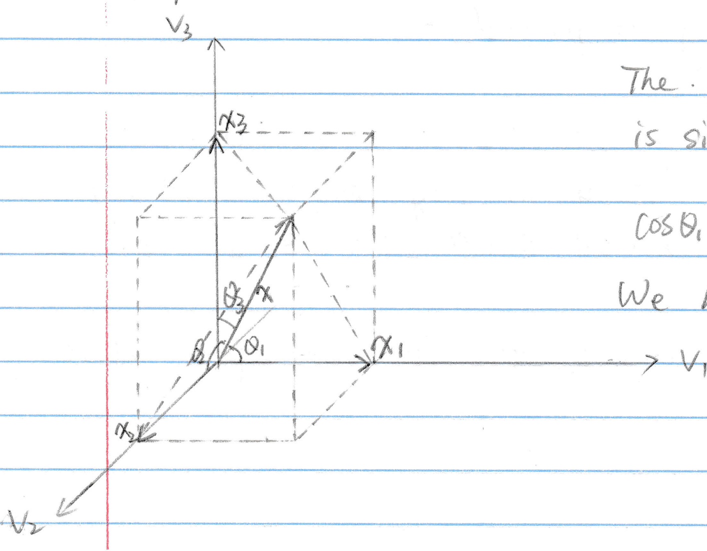
$$= \frac{\langle x,y\rangle^2}{\|x\|^2} + \|y\|^2 - 2\frac{\langle x,y\rangle^2}{\|x\|^2} + \frac{\langle x,y\rangle^2}{\|x\|^2}$$

$$= \|y\|^2$$

Hence the equation 2.53 is proved

2.14 Suppose $\mathcal{V}_1 \perp \mathcal{V}_2$, and $v \in \mathcal{V}_1$ and $v \in \mathcal{V}_2$.
Then by definition of orthogonal vector spaces, v is orthogonal to itself.
This can only be true for the zero vector.

2.15 To prove $\sum_i \cos^2\theta_i = 1$: $\quad \sum_i \cos^2\theta_i = \sum_i \frac{\langle x, e_i\rangle^2}{\|x\|^2\|e_i\|^2} = \frac{\sum_i x_i^2}{\|x\|^2} = \frac{\|x\|^2}{\|x\|^2} = 1$



The projection of the vector x onto each axis is simply the coordinate on each axis.

$$\cos\theta_1 = \frac{x_1}{\|x\|}, \quad \cos\theta_2 = \frac{x_2}{\|x\|}, \quad \cos\theta_3 = \frac{x_3}{\|x\|}$$

We know the length of x is: $\sqrt{x_1^2 + x_2^2 + x_3^2}$

Hence, $\cos\theta_1 + \cos\theta_2 + \cos\theta_3$

$$= \frac{\sqrt{x_1^2 + x_2^2 + x_3^2}}{|x|} = 1$$

2.16 In $\mathbb{R}^3$: $\cos\theta_1 = \frac{\langle \vec{1}, e_1 \rangle}{\|\vec{1}\| \|e_1\|} = \frac{1}{\sqrt{3}}$

We assume $e_1 = (1, 0, 0)$ Since e_2, e_3 are symmetric to e_1, $\cos\theta_2 = \cos\theta_3 = \cos\theta_1 = \frac{1}{\sqrt{3}}$.

In the following, we only consider one θ.

In $\mathbb{R}^{10}$: $\cos\theta = \frac{\langle \vec{1}, e_1 \rangle}{\|\vec{1}\| \|e_1\|} = \frac{1}{\sqrt{10}}$

In $\mathbb{R}^{100}$: $\cos\theta = \frac{1}{10}$

In $\mathbb{R}^{1000}$: $\cos\theta = \frac{1}{10\sqrt{10}}$

Hence, as $n \to \infty$, $\cos\theta \to 0$, implying for higher dimensions, any two vectors are almost orthogonal.

For data analysis, this means the data is always sparse in higher dimensions. And it's difficult to estimate the relationship between vectors in this case.

2.17 For an arbitrary k, using Algorithm 2.1, we obtain: (in the inner loop)

$$x_k = x_k - \langle \tilde{x}_{k-1}, x_k \rangle \tilde{x}_{k-1} - \langle \tilde{x}_{k-2}, x_k \rangle \tilde{x}_{k-2} - \cdots - \langle \tilde{x}_1, x_k \rangle \tilde{x}_1$$

For each of $\tilde{x}_{k-1} \cdots \tilde{x}_1$, they are linear combinations of the first $k-1, k-2, \cdots 1$ vectors.

If for some k, $x_k - \langle \tilde{x}_{k-1}, x_k \rangle \tilde{x}_{k-1} - \cdots - \langle \tilde{x}_1, x_k \rangle \tilde{x}_1 = 0$. That means there exists a linear combination of the first $k-1$ vectors such that $\sum_{i=1}^{k-1} c_i x_i = x_k$

This violates the statement that the initial k vectors are linearly independent.

2.18

(a) For a convex cone V that is created using the first method, we can create another cone, W, in the following manner:

For any $v_1, v_2 \in V$, and for any $a_1, a_2 \geq 0$,
we let $b_1 = \frac{a_1}{a_1+a_2}$, $b_2 = \frac{a_2}{a_1+a_2}$, then we create $w = b_1 v_1 + b_2 v_2$

Next, we scale w by (a_1+a_2). Essentially, $w' = (a_1+a_2)w$
$= (a_1+a_2)\left(\frac{a_1}{a_1+a_2} v_1 + \frac{a_2}{a_1+a_2} v_2\right) = a_1 v_1 + a_2 v_2$.

In this way, we satisfy the requirement $b_1 + b_2 = 1$, and the set $W = V$. The reverse is similar.

(b). First we prove $\frac{x_1}{x_2} \leq \frac{ax_1+by_1}{ax_2+by_2}$:

$$\frac{ax_1+by_1}{ax_2+by_2} - \frac{x_1}{x_2} = \frac{ax_1x_2 + bx_2y_1 - ax_1x_2 - bx_1y_2}{(ax_2+by_2)x_2}$$

$$= \frac{b(x_2y_1 - x_1y_2)}{(ax_2+by_2)x_2}$$

From problem statement, we know $\frac{x_1}{x_2} \leq \frac{y_1}{y_2} \Rightarrow x_1y_2 \leq x_2y_1$
(assuming $x_1, x_2, y_1, y_2 \geq 0$. However, the conclusion holds without this assumption)

Hence, $x_2y_1 - x_1y_2 \geq 0 \Rightarrow \frac{ax_1+by_1}{ax_2+by_2} - \frac{x_1}{x_2} \geq 0$.

Similarly, we have $\frac{ax_1+by_1}{ax_2+by_2} \leq \frac{y_1}{y_2}$

(c) If C_1 and C_2 are convex cones, consider $C_1 \cap C_2$:

for any two vectors $V_1, V_2 \in C_1 \cap C_2$, and for any $a, b \geq 0$:

$$aV_1 + bV_2 \in C_1 \quad \text{and} \quad aV_1 + bV_2 \in C_2.$$

Then by definition, $aV_1 + bV_2 \in C_1 \cap C_2$. Hence, $C_1 \cap C_2$ is a convex cone.

For $C_1 \cup C_2$ is not necessarily a convex cone, a counterexample is:

Obviously, C_1 and C_2 are convex cones but $C_1 \cup C_2$ is not.

2.19.

(a) Prove: consider two vectors a, b, $a \neq b$

Then $(a \times a) \times b = 0$, but $a \times (a \times b) \neq 0$.

Hence cross product is not associative.

(b) The area of the triangle is:

$$S_\Delta = \frac{1}{2} \|x\| \|y\| \sin\theta = \frac{1}{2} \|x \times Y\|$$

(by definition, $\|x \times Y\| = \sin\theta \, \|x\| \, \|Y\|$
| θ is the angle between $\vec{x}$ and $\vec{y}$

(c) First, let's calculate the LHS:

$$y \times z = (y_2 z_3 - y_3 z_2, \; y_3 z_1 - y_1 z_3, \; y_1 z_2 - y_2 z_1)$$

$$\langle x, y \times z \rangle = x_1 (y_2 z_3 - y_3 z_2) + x_2 (y_3 z_1 - y_1 z_3) + x_3 (y_1 z_2 - y_2 z_1)$$

Next, let's calculate the RHS:

$$x \times Y = (x_2 y_3 - x_3 y_2,\ x_3 y_1 - x_1 y_3,\ x_1 y_2 - x_2 y_1)$$

$$\langle x \times Y, z \rangle = (x_2 y_3 - x_3 y_2) z_1 + (x_3 y_1 - x_1 y_3) z_2 + (x_1 y_2 - x_2 y_1) z_3$$

$$= x_1 (y_2 z_3 - y_3 z_2) + x_2 (y_3 z_1 - y_1 z_3) + x_3 (y_1 z_2 - y_2 z_1)$$

$$= \langle x, y \times z \rangle$$

The triple scalar product equality is proved

(d) First, let's calculate LHS:

$$y \times z = (y_2 z_3 - y_3 z_2,\ y_3 z_1 - y_1 z_3,\ y_1 z_2 - y_2 z_1)$$

$$x \times (y \times z) = \begin{pmatrix} x_2 (y_1 z_2 - y_2 z_1) - x_3 (y_3 z_1 - y_1 z_3) \\ x_3 (y_2 z_3 - y_3 z_2) - x_1 (y_1 z_2 - y_2 z_1) \\ x_1 (y_3 z_1 - y_1 z_3) - x_2 (y_2 z_3 - y_3 z_2) \end{pmatrix}$$

Next, let's calculate RHS:

$$\langle x, z \rangle y = (x_1 z_1 + x_2 z_2 + x_3 z_3)\, y$$
$$\langle x, y \rangle z = (x_1 y_1 + x_2 y_2 + x_3 y_3)\, z$$

$$\langle x, z \rangle y - \langle x, y \rangle z = \begin{pmatrix} (x_1 z_1 + x_2 z_2 + x_3 z_3) y_1 - (x_1 y_1 + x_2 y_2 + x_3 y_3) z_1 \\ (x_1 z_1 + x_2 z_2 + x_3 z_3) y_2 - (x_1 y_1 + x_2 y_2 + x_3 y_3) z_2 \\ (x_1 z_1 + x_2 z_2 + x_3 z_3) y_3 - (x_1 y_1 + x_2 y_2 + x_3 y_3) z_3 \end{pmatrix}$$

The "triple vector product" is proved.

$$= \begin{pmatrix} x_2 (y_1 z_2 - y_2 z_1) - x_3 (y_3 z_1 - y_1 z_3) \\ x_3 (y_2 z_3 - y_3 z_2) - x_1 (y_1 z_2 - y_2 z_1) \\ x_1 (y_3 z_1 - y_1 z_3) - x_2 (y_2 z_3 - y_3 z_2) \end{pmatrix}$$

$$= x \times (y \times z)$$

(e) By definition, $x \times y = \|x\|\|y\| \sin\theta \ \vec{e}$
where e is a vector orthogonal to both x and y and has norm 1.

Hence, $\|x \times y\| = \|x\|\|y\| |\sin\theta|$

$$\Rightarrow |\sin\theta| = \frac{\|x \times y\|}{\|x\| \|y\|}$$

Since $x \times y = -y \times x$, there is a difference between $\theta(x,y)$ and $\theta(y,x)$. In the definition, we use $\sin(\theta(x,y))$.
We define $\theta(x,y)$ as the smallest angle through which x would be rotated to y. And we define $\vec{e}$ points to the "up" direction when x is rotated to y counterclock-wise.

(f) By definition, $e_1 \times e_2 = \|e_1\|\|e_2\| \sin(\theta(e_1, e_2)) \ e_3$.
And if we define $\theta(e_1, e_2)$ in the manner in (e), then:

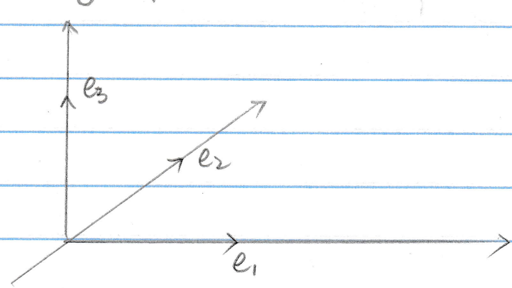
$$e_1 \times e_2 = 1 \times 1 \times \sin\frac{\pi}{2} \times e_3 = e_3.$$

Similarly: $e_2 \times e_1 = \|e_2\| \|e_1\| \sin(\theta(e_2, e_1)) \ e_3$

$$= 1 \times 1 \times \sin\frac{3\pi}{2} \times e_3 = -e_3$$

Hence, $e_3 = e_1 \times e_2 = -e_2 \times e_1$

In calculating the above quantities, especially in determinating θ, we consider the following graph:



2.20. Given that $\bar{x} = \frac{\langle x, \vec{1} \rangle}{n} \cdot \vec{1}$ and $x_c = x - \bar{x}$, we first consider the RHS:

$$RHS = \|\bar{x}\|^2 + \|x_c\|^2$$

$$= \sum_i \left(\frac{\langle x, \vec{1} \rangle}{n}\right)^2 + \sum_i \left(x_i - \frac{\langle x, \vec{1} \rangle}{n}\right)^2$$

$$= n \cdot \frac{\langle x, \vec{1} \rangle^2}{n^2} + \sum_i \left(x_i^2 - \frac{2\langle x, \vec{1} \rangle}{n} x_i + \frac{\langle x, \vec{1} \rangle^2}{n^2}\right)$$

$$= \frac{\langle x, \vec{1} \rangle^2}{n} + \sum_i x_i^2 - \frac{2\langle x, \vec{1} \rangle}{n} \sum_i x_i + n \cdot \frac{\langle x, \vec{1} \rangle^2}{n^2}$$

note: $\sum_i x_i = \langle x, \vec{1} \rangle$

$$= \frac{\langle x, \vec{1} \rangle^2}{n} + \|x\|^2 - \frac{2\langle x, \vec{1} \rangle^2}{n} + \frac{\langle x, \vec{1} \rangle^2}{n}$$

$$= \|x\|^2 = RHS$$

Hence equation 2.69 is proved.

2.21 An example: let $x = (1, 2, 3)$, $y = (3, 2, 1)$

then $\bar{x} = (2, 2, 2)$, $x_c = (-1, 0, 1)$

$\bar{y} = (2, 2, 2)$, $y_c = (1, 0, -1)$

$$\cos\theta(x, y) = \frac{\langle x, y \rangle}{\|x\| \|y\|} = \frac{5}{7}$$

$$\cos\theta(x_c, y_c) = \frac{\langle x_c, y_c \rangle}{\|x_c\| \|y_c\|} = -1$$

Hence, $\theta(x, y) \neq \theta(x_c, y_c)$

2.22 Proof:

We start from the LHS:

$$LHS = V(ax+y)$$

by definition

$$= \frac{1}{n-1}\left(\|ax+y\|^2 - \|\overline{ax+y}\|^2\right)$$

here, $\bar{x}$ and $\bar{y}$ are scalar, not vectors: $\bar{x} = \frac{1}{n}\sum_i x_i$

$$= \frac{1}{n-1}\left[\sum_i (ax_i+y_i)^2 - \sum_i (a\bar{x}+\bar{y})^2\right]$$

$$= \frac{1}{n-1}\left[\sum_i (a^2x_i^2 + 2a\, x_i y_i + y_i^2) - \sum_i (a^2\bar{x}^2 + 2a\bar{x}\bar{y} + \bar{y}^2)\right]$$

$$= \frac{1}{n-1}\left[a^2\sum_i (x_i^2 - \bar{x}^2) + \sum_i (y_i^2 - \bar{y}^2) + 2a\sum_i (x_i y_i - \bar{x}\bar{y})\right]$$

$$= \frac{a^2}{n-1}\left(\sum_i x_i^2 - \sum_i \bar{x}^2\right) + \frac{1}{n-1}\left(\sum_i y_i^2 - \sum_i \bar{y}^2\right) + \frac{2a}{n-1}\left(\sum_i x_i y_i - \sum_i \bar{x}\bar{y}\right)$$

$$= \frac{a^2}{n-1}\left(\|x\|^2 - \|\bar{x}\|^2\right) + \frac{1}{n-1}\left(\|y\|^2 - \|\bar{y}\|^2\right) + \frac{2a}{n-1}\left(\langle x, y\rangle - \langle \bar{x}, \bar{y}\rangle\right)$$

$$= a^2 V(x) + V(y) + 2a\frac{\langle x, y\rangle - \langle \bar{x}, \bar{y}\rangle}{n-1} \qquad (1)$$

The next thing we need to prove, is $\langle x, y\rangle - \langle \bar{x}, \bar{y}\rangle = \langle x_c, y_c\rangle$

We start from $\langle x_c, y_c\rangle = \langle x-\bar{x}, y-\bar{y}\rangle$

$$= \sum_i (x_i-\bar{x})(y_i-\bar{y})$$

here, $\bar{x}$ and $\bar{y}$ are scalars, not vectors.

$$= \sum_i (x_i y_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y})$$

$$= \sum_i x_i y_i - \bar{y}\sum_i x_i - \bar{x}\sum_i y_i + \sum_i \bar{x}\bar{y}$$

$$= \sum_i x_i y_i - n\bar{x}\bar{y} - n\bar{x}\bar{y} + n\bar{x}\bar{y}$$

$$= \sum_i x_i y_i - n\bar{x}\bar{y} = \sum_i x_i y_i - \sum_i \bar{x}\bar{y} = \langle x, y\rangle - \langle \bar{x}, \bar{y}\rangle$$

We plug $\langle x_c, y_c \rangle = \langle x, y \rangle - \langle \bar{x}, \bar{y} \rangle$ into ① and get:

$$LHS = V(ax+y)$$

$$= a^2 V(x) + V(y) + 2a \frac{\langle x_c, y_c \rangle}{n-1}$$

Hence the equation 2.77 is proved

2.23

(a) Proof:

First, let's consider the LHS:

$$LHS = \left(Cov(x,y)\right)^2$$

$$= \frac{\langle x-\bar{x}, y-\bar{y} \rangle^2}{(n-1)^2}$$

by Cauchy-Schwarz inequality

$$\leq \frac{1}{(n-1)^2} \|x-\bar{x}\|^2 \|y-\bar{y}\|^2$$ ①

Next, we prove that $\|x-\bar{x}\|^2 = \|x\|^2 - \|\bar{x}\|^2$

$$\|x-\bar{x}\|^2 = \sum_i (x_i - \bar{x})^2 = \sum_i (x_i^2 - 2x_i\bar{x} + \bar{x}^2)$$

$$= \sum_i x_i^2 - 2\bar{x}\sum_i x_i + \sum_i \bar{x}^2$$

$$= \|x\|^2 - 2n\bar{x}^2 + n\bar{x}^2$$

$$= \|x\|^2 - n\bar{x}^2 = \|x\|^2 - \sum_i \bar{x}^2$$

$$= \|x\|^2 - \|\bar{x}\|^2$$

Similarly, $\|y-\bar{y}\|^2 = \|y\|^2 - \|\bar{y}\|^2$

We plug them in ① and get:

$$\left(Cov(x,y)\right)^2 \le \frac{1}{(n-1)^2}\left(\|x\|^2 - \|\bar{x}\|^2\right)\left(\|y\|^2 - \|\bar{y}\|^2\right)$$

$$= \left[\frac{1}{n-1}\left(\|x\|^2 - \|\bar{x}\|^2\right)\right]\left[\frac{1}{n-1}\left(\|y\|^2 - \|\bar{y}\|^2\right)\right]$$

$$= V(x)\,V(y)$$

Hence the inequality is proved.

(b) Proof:

From (a): $\left(Cor(x,y)\right)^2 = \left(\frac{Cov(x,y)}{\sqrt{V(x)V(y)}}\right)^2$

$$= \frac{Cov^2(x,y)}{V(x)\,V(y)}$$

$$\le \frac{V(x)\,V(y)}{V(x)\,V(y)}$$

$$= 1$$

Hence $-1 \le Cor(x,y) \le 1$