June 4, 2013

1. (a) We have

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

and hence

$$\|x+y\|^2 + \|x-y\|^2 = 2\langle x,x\rangle + 2\langle y,y\rangle = 2\|x\|^2 + 2\|y\|^2.$$

(b) ($\Leftarrow$) After rearrangement we could assume that $y = \alpha x$ for some $\alpha \in F$. Then

$$\begin{aligned}|\langle x,y\rangle| = |\langle x,\alpha x\rangle| &= |\alpha| \cdot \|x\|^2 \\ &= \|x\| \cdot \|\alpha x\| = \|x\| \cdot \|y\|.\end{aligned}$$

($\Rightarrow$) If $x = 0$, then $x = 0 \cdot y$ and we are done. Now assume that $x \neq 0$. Write $y = \alpha x + \varepsilon$ for some $\alpha \in F$ and $\varepsilon \in \operatorname{Span}\{x\}^\perp$. Then

$$|\langle x,y\rangle|^2 = |\alpha|^2 \cdot \|x\|^4$$

$$\|x\|^2\|y\|^2 = \|x\|^2\|\alpha x + \varepsilon\|^2 = \|x\|^2 \left(|\alpha|^2\|x\|^2 + \|\varepsilon\|^2\right).$$

Thus we have $\|\varepsilon\|^2 = 0$, which implies that $\varepsilon = 0$ and hence y is a multiple of x.

2. (a) Suppose $v \in W_1$. Then for all $w \in W_1^\perp$ we have $\langle v, w\rangle = 0$. Thus $W_1 \subset (W_1^\perp)^\perp$. Since

$$\begin{aligned}\dim(W_1^\perp)^\perp &= \dim V - \dim W_1^\perp \\ &= \dim V - (\dim V - \dim W_1) \\ &= \dim W_1.\end{aligned}$$

Therefore $W_1 = (W_1^\perp)^\perp$.

(b) Since $W_1 + W_2$ contains W_1 and W_2, we see that $(W_1 + W_2)^\perp$ is contained in both $W_1^\perp$ and $W_2^\perp$.

On the other hand, suppose $v \in W_1^\perp \cap W_2^\perp$. For any $w \in W_1 + W_2$, one can write $w = w_1 + w_2$ for some $w_i \in W_i$. Then we have

$$\langle v, w\rangle = \langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle = 0 + 0 = 0.$$

Thus $W_1^\perp \cap W_2^\perp \subset (W_1 + W_2)^\perp$.

(c) Plugging the two subspaces $W_1^\perp$ and $W_2^\perp$ into (b) and using (a), one obtains

$$(W_1^\perp + W_2^\perp)^\perp = (W_1^\perp)^\perp \cap (W_2^\perp)^\perp = W_1 \cap W_2.$$

Taking perp in the above equation and applying (a) again, we get

$$(W_1 \cap W_2)^\perp = \left((W_1^\perp + W_2^\perp)^\perp\right)^\perp = W_1^\perp + W_2^\perp.$$

3. Suppose W is T-invariant. Let $v \in W^\perp$. For all $w \in W$, we have

$$\langle T^*(v), w\rangle = \langle v, T(w)\rangle = 0.$$

Thus $W^\perp$ is T^*-invariant.

On the other hand, if $W^\perp$ is T^*-invariant, then $(W^\perp)^\perp$ is $(T^*)^*$-invariant. Since $(W^\perp)^\perp = W$ and $(T^*)^* = T$, the statement follows.

4. We apply the Gram-Schmidt process to the basis

$$v_1 = (1,2,2), e_1 = (1,0,0), e_2 = (0,1,0)$$

of $\mathbb{R}^3$.

Using e_1 to modify v to obtain a vector in $\mathrm{Span}\{v\}^{\perp}$, one gets the vector

$$v_2 = -9e_1 + v_1 = (-8,2,2).$$

Now suppose

$$v_3 := v_1 + (\alpha-1)e_1 + (\beta-2)e_2 = (\alpha,\beta,2) \in \mathrm{Span}\{v_1,e_1\}^{\perp}.$$

Then

$$\begin{aligned} \alpha + 2\beta + 4 &= \langle v_3, v_1\rangle = 0 \\ \alpha &= \langle v_3, e_1\rangle = 0. \end{aligned}$$

Thus we obtain the orthogonal basis v_1, v_2 and

$$v_3 = (0,-2,2).$$

Normalizing the vectors, we obtain the orthogonal matrix

$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{-4}{\sqrt{18}} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{18}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

5. (a) The direction ($\Rightarrow$) is clear since two similar matrices have the same characteristic polynomial.

   ($\Leftarrow$) A normal complex matrix is unitarily similar to a diagonal matrix whose diagonal entries are the eigenvalues. Thus A and B are unitarily similar to a certain diagonal matrix and hence they are unitarily similar.

   (b) In fact there are unitary matrices X, Y and a diagonal matrix D with non-negative diagonal entries $\sigma_1, \cdots, \sigma_n$ such that $A = XDY$. Then

$$\begin{aligned} A^*A &= Y^*D^*X^*XDY = Y^*D^2Y \\ AA^* &= XDYY^*D^*X^* = XD^2X^*. \end{aligned}$$

   Therefore A^*A and AA^* are both unitarily similar to D^2 and consequently they are unitarily similar.

6. In $M_n(\mathbb{R})$, there exists an orthogonal matrix X and a diagonal matrix D with diagonal $\lambda_1, \cdots, \lambda_n$ such that $A = XDX^t$. Since $I = X^tA^rX = D^r$, we see that the real numbers λ_i satisfy $\lambda_i^r = 1$. Thus $\lambda_i = \pm 1$ and we have $A^2 = XD^2X^t = I$.

7. (a) We have

$$\begin{aligned} 0 = H(v+w, v+w) &= H(v,v) + H(v,w) + H(w,v) + H(w,w) \\ &= H(v,w) + H(w,v). \end{aligned}$$

   (b) We find the desired basis by induction on $\dim V$.

   First if $\dim V = 1$, then any basis $\{u\}$ would satisfy $H(u,u) = 0$. In this case, $a = 0, b = 1$.

Now suppose $\dim V \geq 2$. If in case $H(x,y) = 0$ for all $x, y \in V$, then any basis $\{u_1, \cdots, u_n\}$ of V satisfies the requirement (with $a = 0, b = n$).
So suppose that $\dim V \geq 2$ and there exist $v_1, w_1 \in V$ such that $H(v_1, w_1) = c \neq 0$. First notice that $v_1, w_1 \neq 0$. Secondly after replacing w_1 by $\frac{1}{c} \cdot w_1$, we can assume that $H(v_1, w_1) = 1$. Thirdly any of v_1 and w_1 cannot be a multiple of the other by (a). Thus $W := \text{Span}\{v_1, w_1\}$ is a two-dimensional subspace of V.
Define

$$W^\perp := \{v \in V \,|\, H(v,w) = 0 \ \forall w \in W\};$$

it is clearly a subspace of V. We now show that $V = W \oplus W^\perp$. Fix $v \in V$. If we have $v = \xi v_1 + \eta w_1 + \varepsilon$ for some $\xi, \eta \in F$ and $\varepsilon \in W^\perp$, then

$$\begin{aligned} H(v_1, v) &= H(v_1, \xi v_1 + \eta w_1 + \varepsilon) = \eta \\ H(w_1, v) &= H(w_1, \xi v_1 + \eta w_1 + \varepsilon) = -\xi. \end{aligned}$$

That is, ξ, η and ε are uniquely determined by v if they exist. This shows that $W \cap W^\perp = \{0\}$. On the other hand, if we let $(\xi, \eta) = (-H(w_1, v), H(v_1, v))$ and $\varepsilon = v - \xi v_1 - \eta w_1$, then a direct computation shows that $H(v_1, \varepsilon) = H(w_1, \varepsilon) = 0$. This shows that $V = W + W^\perp$.
At this point, we just restrict the bilinear form H to the subspace $W^\perp$. Since $\dim W^\perp$ is strictly smaller than $\dim V$, the induction hypothesis implies that there exists a basis $\beta' = \{v_2, w_2, \cdots, u_1, \cdots\}$ of $W^\perp$ such that the matrix $\psi_{\beta'}(H_{W^\perp})$ is of the correct shape. Let $\beta = \{v_1, w_1\} \cup \beta'$. Then β is a basis of V and it satisfies the requirement.

8. (a) Let $\{e_1, \cdots, e_n\}$ be the standard basis of $\mathbb{R}^n$. We use induction on n to find $v_r = e_r +$ (a linear combination of $e_1, \cdots, e_{r-1}$) that satisfies the requirement.
If $n = 1$, we take v_1 to be the standard basis. Then $v_1^t A v_1 = \Delta_1$.
Assume that $n \geq 2$. Then the matrix A_{n-1} is a symmetric matrix whose upper left sub-matrices have determinants $\neq 0$. Thus by induction, there exist column vectors $v_1', \cdots, v_{n-1}'$ in $\mathbb{R}^{n-1}$ of the shape $v_r' = e_r +$ (a linear combination of $e_1, \cdots, e_{r-1}$) satisfying $(v_r')^t A_{n-1} v_r' = \Delta_r / \Delta_{r-1}$, etc (with $\Delta_0 = 1$).
For $r = 1, \cdots, n-1$, let $v_r \in \mathbb{R}^n$ obtained by adding 0 in the n-th entry to v_r'. Let

$$\begin{aligned} v_n &= e_n + \sum_{i=1}^{n-1} \frac{-a_{ni}\Delta_{i-1}}{\Delta_i} v_i \\ &= e_n + (\text{a linear combination of } e_1, \cdots, e_{n-1}). \end{aligned}$$

Then by a direct computation, we have $v_n^t A v_r = 0$ for $1 \leq r < n$. To compute $\alpha := v_n^t A v_n$, let $P = (v_1, \cdots, v_n) \in M_n(\mathbb{R})$. We have

$$P = \begin{pmatrix} 1 & & * & \\ & 1 & & \\ & & \ddots & \\ & 0 & & 1 \end{pmatrix}$$

is upper-triangular with 1 in the diagonal, and

$$P^t A P = (v_i^t A v_j) = \begin{pmatrix} \Delta_1 & & & 0 & \\ & \frac{\Delta_2}{\Delta_1} & & & \\ & & \ddots & & \\ & & & \frac{\Delta_{n-1}}{\Delta_{n-2}} & \\ & 0 & & & \alpha \end{pmatrix}.$$

Taking determinants on both sides, we obtain $\alpha = \det(A)/\Delta_{n-1}$, which is what we want.

[This can be regarded as Gram-Schmidt process for symmetric bilinear forms. Notice that here we do not require the field F to be $\mathbb{R}$.]

(b) ($\Rightarrow$) We use induction on n.

If $n = 1$, then $\Delta_1 = e_1^t A e_1 > 0$ and we are done.

Suppose $n \geq 2$. As in (a), we can consider the smaller matrix A_{n-1}, which is then positive definite because

$$v^t A_{n-1} v = (v^t, 0) A \begin{pmatrix} v \\ 0 \end{pmatrix} > 0$$

for any non-zero $v \in \mathbb{R}^{n-1}$. By induction, we obtain $\Delta_r > 0$ for $1 \leq r < n$. For $\Delta_n = \det A$, notice that A is diagonalizable (since A is symmetric) and all its eigenvalues are positive (since A is positive definite). Thus $\det A =$ product of eigenvalues > 0.

($\Leftarrow$) **Method I.** By (a), there exists a basis $\{v_1, \cdots, v_n\}$ such that

$$v_i^t A v_j \begin{cases} > 0 & i = j \\ = 0 & i \neq j. \end{cases}$$

Thus for any non-zero $v \in V$, we have $v = \sum_i \alpha_i v_i$ with $\alpha_j \neq 0$ for some j and $v^t A v = \sum_i \alpha_i^2 (v_i^t A v_i) > 0$.

Method II. One can also argue by induction on n as follows. Again we look at the symmetric A_{n-1} of smaller size. Since $\Delta_r > 0$ for $1 \leq r < n$, A_{n-1} is positive definite by induction. Thus there exist an orthogonal $Q \in M_{n-1}(\mathbb{R})$ and a diagonal D with diagonal entries $\sigma_1, \cdots, \sigma_{n-1} > 0$ such that $Q^t A_{n-1} Q = D$. We then have

$$B := \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix}^t A \begin{pmatrix} Q & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} & & b_1 \\ & D & \vdots \\ b_1 & \cdots & b_n \end{pmatrix} \quad \text{for some} \quad \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n.$$

Notice that A and B are orthogonally similar, and hence A is positive definite if and only if B is. Hence we reduce to consider B. Also by direct computation, we have

$$\Delta_n = \det A = \det B = \sigma_1 \cdots \sigma_{n-1} \underbrace{\left(b_n - \frac{b_1^2}{\sigma_1} - \cdots - \frac{b_{n-1}^2}{\sigma_{n-1}} \right)}_{\sigma_n},$$

which is positive by assumption. Thus $\sigma_n > 0$. Now take any non-zero column vector $x = \sum x_i e_i \in \mathbb{R}^n$, we have

$$\begin{aligned} x^t B x &= \sum_{i=1}^{n-1} \left(\sigma_i x_i^2 + 2 b_i x_i x_n \right) + b_n x_n^2 \\ &= \sum_{i=1}^{n-1} \sigma_i \left(x_i + \frac{b_i}{\sigma_i} x_n \right)^2 + \underbrace{\left(b_n - \sum_{i=1}^{n-1} \frac{b_i^2}{\sigma_i} \right)}_{\sigma_n} x_n^2 > 0. \end{aligned}$$

Therefore B, and hence A, are positive definite.

(c) Consider $A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$. Then $\Delta_1, \Delta_2 \geq 0$ but $e_2^t A e_2 = -1 < 0$.