Mathematics 262
Homework 10 solutions

Assignment:

- Section 8.4: 4, 5, 6, 14

§8.4, #4: If U is a unitary operator on $V = \mathbf{R}^2$, with matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (with respect to the standard basis), then since U is unitary and we're working over $\mathbf{R}$, then we have $U^{-1} = U^t$. Multiplying U by U^t and setting the result equal to I gives three equations:

$$\begin{aligned} a^2 + b^2 &= 1, \\ ac + bd &= 0, \\ c^2 + d^2 &= 1. \end{aligned}$$

Because of the first equation, I can let $a = \cos\theta$ and $b = \sin\theta$ for some angle θ between 0 and 2π. Similarly, I can let $c = \sin\varphi$ and $d = \cos\varphi$ for some angle φ. The second equation then says that

$$\cos\theta \sin\varphi + \sin\theta \cos\varphi = 0.$$

Using a trig identity, this becomes $\sin(\theta+\varphi) = 0$. Therefore, φ is either $2\pi - \theta$ or $\pi - \theta$. If the first of these holds, then the matrix for U is $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. If the second holds, then the matrix is $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$.

(a) $U_\theta U_\varphi$ is rotation by φ followed by rotation by θ, which is the same as $U_{\theta+\varphi}$, rotation by the angle $\theta + \varphi$. (One can also see this by multiplying out the matrices and using trig identities, but it's not as pleasant.)

(b) Well, $U^* = U^t$ since we're working over $\mathbf{R}$, and $U^t = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$. Since $\sin(-\theta) = -\sin\theta$ and $\cos(-\theta) = \cos\theta$, this matrix is equal to $U_{-\theta}$. Alternatively, by part (a), $U_\theta U_{-\theta} = U_0 = I$, so $U_{-\theta} = U_\theta^{-1} = U_\theta^*$. (Geometrically, if you rotate by θ and then by $-\theta$, you have performed the identity operation.)

(c) The change-of-basis matrix is U_φ, so I want to compute $U_\varphi^{-1} U_\theta U_\varphi$. I could compute this by multiplying out the matrices. I could also compute this using the previous parts:

$$U_\varphi^{-1} U_\theta U_\varphi = U_{-\varphi} U_\theta U_\varphi = U_{-\varphi+\theta+\varphi} = U_\theta.$$

(I could also try to argue that if a linear operator is a rotation, it should treat the rotated basis $\{\alpha_1, \alpha_2\}$ just like standard orthonormal basis: α_1 should get sent to $(\cos\theta)\alpha_1 + (\sin\theta)\alpha_2$, etc. So the matrix for U_θ should be the same in this new basis.)

§8.4, #5: Following the suggestion in the book, I let $\alpha = (1,1,1)$ and $\beta = (1,1,-2)$, and I use Gram-Schmidt to turn this into an orthonormal basis: $\alpha_1 = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$ and $\alpha_2 = (\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{2}{\sqrt{6}})$. I find a vector orthogonal to both α and β by using the cross product, and I divide by its norm to get $\alpha_3 = (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$.

Now, U sends α_3 to itself, and it rotates the subspace spanned by α_1 and α_2 by the angle θ, so with respect to the orthonormal basis $\{\alpha_1, \alpha_2, \alpha_3\}$, U is represented by the matrix

$$A = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let P be the change-of-basis matrix:

$$P = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \end{bmatrix}.$$

This is supposed to be unitary (i.e., orthogonal, since we're working over $\mathbf{R}$), and you should check this (e.g., check that the columns are an orthonormal set, or that $P^{-1} = P^t = P^*$). With respect to the standard basis, then, U is represented by the matrix

$$PAP^{-1} = PAP^t = \begin{bmatrix} \frac{1}{2}\cos\theta + \frac{1}{2} & \frac{1}{2}\cos\theta - \frac{1}{2} & \frac{1}{\sqrt{2}}\sin\theta \\ \frac{1}{2}\cos\theta - \frac{1}{2} & \frac{1}{2}\cos\theta + \frac{1}{2} & \frac{1}{\sqrt{2}}\sin\theta \\ -\frac{1}{\sqrt{2}}\sin\theta & -\frac{1}{\sqrt{2}}\sin\theta & \cos\theta \end{bmatrix}.$$

(Since I'm changing from the basis $\{\alpha_1, \alpha_2, \alpha_3\}$ to the standard basis, I want PAP^{-1}. If I were changing from the standard basis to $\{\alpha_1, \alpha_2, \alpha_3\}$, I would want $P^{-1}AP$.) You could check this by seeing if $U\alpha_3 = \alpha_3$, and seeing if $U\alpha_i$ looks reasonable for $i = 1, 2$. (For instance, it's not so hard to see that $U\alpha_1$ is a linear combination of α_1 and α_2, so at least it's in the right general area.)

§8.4, #6: Choose orthonormal bases for W and $W^\perp$. Together, they form an orthonormal basis for V. With respect to this basis, U is represented by the matrix $A = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$. (The I in the upper left is the identity matrix on W. The $-I$ in the lower right is negative the identity matrix on $W^\perp$.)

(a) From the matrix description, $A = A^t = A^*$, so U is self-adjoint. Also, $AA^* = AA^t = A^2 = I$, so U is unitary (and also orthogonal).

(b) I guess I should find orthonormal bases for W and $W^\perp$ and do a change of basis: W is spanned by $(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$. $W^\perp$ is spanned by any two vectors orthogonal to this one, say $(1, 0, -1)$ and $(0, 1, 0)$. These happen to be orthogonal (if they weren't, I'd use Gram-Schmidt), so an orthonormal basis for $W^\perp$ is $\{(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}), (0, 1, 0)\}$. Let P be the change-of-basis matrix

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix}.$$

This is a unitary matrix (because the columns are an orthonormal basis), so $P^{-1} = P^* = P^t$. With respect to this basis, U is represented by the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

So with respect to the standard basis, U is represented by

$$PAP^{-1} = PAP^t = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

(Again, since I'm changing from some other basis to the standard basis, I look at PAP^{-1}, not $P^{-1}AP$.) Does this answer make sense? Well, ε_2 is in $W^\perp$, so $U\varepsilon_2 = -\varepsilon_2$. To compute $U\varepsilon_1$, I write $\varepsilon_1 = \frac{1}{2}(1, 0, 1) + \frac{1}{2}(1, 0, -1)$; the first of these vectors is in W, the second in $W^\perp$. Thus $U\varepsilon_1 = \frac{1}{2}(1, 0, 1) - \frac{1}{2}(1, 0, -1) = (0, 0, 1) = \varepsilon_3$. Similarly, $U\varepsilon_3 = \varepsilon_1$. So this is indeed the correct matrix.

§8.4, #14:

(a) Let's see. If $T(0) = 0$, then

$$||T\alpha|| = ||T\alpha - 0|| = ||T\alpha - T0|| = ||\alpha - 0|| = ||\alpha||,$$

so if I knew that T were linear, I could conclude that T is unitary. So I have to prove linearity.

First, I'll show that T preserves inner products. Since $(\alpha|\beta) = \frac{1}{4}(||\alpha+\beta||^2 - ||\alpha-\beta||^2)$, and since $||T\alpha - T\beta|| = ||\alpha - \beta||$, it's good enough to show that $||T\alpha + T\beta|| = ||\alpha + \beta||$. Well,

$$||T\alpha + T\beta|| = ||T\alpha - (-T\beta)||.$$

If I knew that $-T\beta = T(-\beta)$, then this would equal

$$||T\alpha - T(-\beta)|| = ||\alpha - (-\beta)|| = ||\alpha + \beta||.$$

So it's good enough to show that $T(-\beta) = -T\beta$ for each vector β. Find a vector γ perpendicular to β. Then $||T\beta|| = ||\beta||$, $||T\gamma|| = ||\gamma||$, and $||T\gamma - T\beta|| = ||\gamma - \beta||$. Since the vectors β, γ, and $\gamma - \beta$ form a right triangle, and since the vectors $T\beta$, $T\gamma$, and $T\gamma - T\beta$ have the same lengths, then they must also form a right triangle. In particular, $T\beta$ is perpendicular to $T\gamma$. Similarly, $T(-\beta)$ is perpendicular to $T\gamma$. Two vectors in $\mathbf{R}^2$ which are perpendicular to the same vector must be parallel, so $T\beta$ and $T(-\beta)$ are parallel. They also have the same length, so $T(-\beta) = \pm T\beta$. Lastly, $||T\beta - T(-\beta)|| = ||\beta - (-\beta)|| = 2||\beta||$; therefore $T(-\beta)$ must be $-T\beta$. Therefore T preserves inner products.

Let $\varepsilon_1 = (1,0)$ and $\varepsilon_2 = (0,1)$. Since T preserves norms and inner products, then $T\varepsilon_1$ and $T\varepsilon_2$ are an orthonormal basis for $\mathbf{R}^2$. I claim that T is linear. Suppose $\alpha = (x_1, x_2) = x_1\varepsilon_1 + x_2\varepsilon_2$. Then $(T\alpha|T\varepsilon_i) = (\alpha|\varepsilon_i) = x_i$, so $T\alpha = x_1 T\varepsilon_1 + x_2 T\varepsilon_2$. In other words,

$$T(x_1\varepsilon_1 + x_2\varepsilon_2) = x_1 T\varepsilon_1 + x_2 T\varepsilon_2.$$

This implies that T is linear. If you want to be explicit: if $\beta = y_1\varepsilon_1 + y_2\varepsilon_2$, then

$$\begin{aligned} T(c\alpha + \beta) &= T((cx_1 + y_1)\varepsilon_1 + (cx_2 + y_2)\varepsilon_2) \\ &= (cx_1 + y_1)T\varepsilon_1 + (cx_2 + y_2)T\varepsilon_2 \\ &= (cx_1 T\varepsilon_1 + cx_2 T\varepsilon_2) + (y_1 T\varepsilon_1 + y_2 T\varepsilon_2) \\ &= cT\alpha + T\beta. \end{aligned}$$

(b) Given a rigid motion f of $\mathbf{R}^2$, find a point γ so that $f(\gamma) = 0$ (justification below), and define a function $U : \mathbf{R}^2 \to \mathbf{R}^2$ by $U(\alpha) = f(\alpha + \gamma)$. Then $U(0) = f(\gamma) = 0$ and U is a rigid motion:

$$||U\alpha - U\beta|| = ||f(\alpha + \gamma) - f(\beta + \gamma)|| = ||(\alpha + \gamma) - (\beta + \gamma)|| = ||\alpha - \beta||.$$

By part (a), then, U is a unitary (linear) operator. Furthermore, for all vectors α, $f(\alpha + \gamma) = U(\alpha)$. Letting $\beta = \alpha - \gamma$, we get

$$f(\beta) = U(\beta - \gamma) = U(T_{-\gamma}(\beta)) = U \circ T_{-\gamma}(\beta),$$

where $T_{-\gamma}$ is translation by $-\gamma$: $T_{-\gamma}(\beta) = \beta - \gamma$.

I still owe you the justification that there is a point γ so that $f(\gamma) = 0$. The idea is that if f is a rigid motion, then it preserves all distances, angles, and other geometric information. So if f takes 0 to some point α, then f takes the circle of radius $||\alpha||$ centered at 0 to the circle of radius $||\alpha||$ centered at α. In particular, since 0 is on this second circle, then f takes some point on the first circle to 0. One can do this more carefully (just from the definition of rigid motion, without hand-waving about "preserving other geometric information"), but I don't feel like writing down the details.

(c) This follows from problem 4. Every unitary operator on $\mathbf{R}^2$ is either a rotation or a reflection followed by a rotation, and every rigid motion is a translation followed by one of these.