## Solutions to Homework Assignment 12

Assigned Fri 04/29. Due never.

1. Let $a, b, c \in \mathbb{R}$, and define $C \subseteq \mathbb{R}^{2}$ by

$$
C=\left\{\left(x_{1}, x_{2}\right) \mid a x_{1}^{2}+2 b x_{1} x_{2}+c x_{2}^{2}=1\right\}
$$

The point of this problem is to decide if $C$ is a ellipse, parabola or hyperbola. I'm assuming you must have seen this condition (probably without proof) in high school coordinate geometry. The advantage of the approach here is that once you understand it, you can use the trick in higher dimensions. Plus it uses a bunch of material from the last week, so it serves as a nice review.
(a) Show that there exists $\theta \in[-\pi, \pi), \lambda_{1}, \lambda_{2} \in \mathbb{R}$ such that

$$
a x_{1}^{2}+2 b x_{1} x_{2}+c x_{2}^{2}=\lambda_{1} y_{1}^{2}+\lambda_{2} y_{2}^{2},
$$

where

$$
y_{1}=x_{1} \cos \theta-x_{2} \sin \theta, \quad \text { and } \quad y_{2}=x_{1} \sin \theta+x_{2} \cos \theta
$$

Solution. Let $A=\left(\begin{array}{cc}a & b \\ b & c\end{array}\right)$. Then $a x_{1}^{2}+2 b x_{1} x_{2}+c x_{2}^{2}=\langle A x, x\rangle$. Since $A$ is real and symmetric, the spectral theorem implies there exists an orthogonal matrix $P$ such that $P A P *=D$, where $D=\binom{\lambda_{1}}{\lambda_{2}}$, for some $\lambda_{1}, \lambda_{2} \in \mathbb{R}$. Replacing $P$ with $-P$ if necessary, we can assume $\operatorname{det}(P)=1$.
Thus, by your previous homework, $P=\left(\begin{array}{c}\cos \theta-\sin \theta \\ \sin \theta \\ \cos \theta\end{array}\right)$ for some $\theta \in$ $[-\pi, \pi)$. Now define $y=P x$. That is, let

$$
y=\binom{y_{1}}{y_{2}}=\left(\begin{array}{cc}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{array}\right)\binom{x_{1}}{x_{2}}=\binom{x_{1} \cos \theta-x_{2} \sin \theta}{x_{1} \sin \theta+x_{2} \cos \theta} .
$$

Then

$$
a x_{1}^{2}+2 b x_{1} x_{2}+c x_{2}^{2}=\langle A x, x\rangle=\left\langle P^{*} D P x, x\right\rangle=\langle D y, y\rangle=\lambda_{1} y_{1}^{2}+\lambda_{2} y_{2}^{2}
$$

(b) Find a condition on $a, b, c$ which determines if $C$ is an ellipse, parabola or hyperbola.

Solution. Observe that the coordinates $\left(y_{1}, y_{2}\right)$ in the previous subpart are a rotation of the standard coordinates $\left(x_{1}, x_{2}\right)$. Thus the shape of $C$ is the same under either coordinate system.
In the coordinates given by $y$, we immediately see the shape of $C$. First note that for $C$ to be non-empty, we must have either $\lambda_{1}>0$, or $\lambda_{2}>0$. Without loss, assume $\lambda_{1}>0$. Then we immediately see that if $\lambda_{2}>0, C$ is an ellipse. If $\lambda_{2}=0, C$ is a parabola, and if $\lambda_{2}<0$ then $C$ is a hyperbola. Since $\lambda_{1}, \lambda_{2}$ are roots of $\lambda^{2}-(a+c) \lambda+\left(a c-b^{2}\right)$, one can easily check:

1. If $a c-b^{2}>0$, then $C$ is an ellipse.
2. If $a c-b^{2}=0$, then $C$ is a parabola.
3. If $a c-b^{2}<0$, then $C$ is a hyperbola.
4. Let $V$ be a finite dimensional inner product space. Show that $T \in \mathcal{L}(V, V)$ is positive if and only if there exists $S \in L(V, V)$ such that $T=S^{*} S$.

Solution. We've seen in class that $S^{*} S$ is positive. For the converse, note that by definition positive operators are self adjoint. Thus, by the spectral theorem, there exists an orthonormal basis $\left\{v_{1}, \ldots, v_{n}\right\}$, and $\lambda_{i} \in \mathbb{R}$ such that $T v_{i}=\lambda v_{i}$. Note that since $T$ is positive, $\lambda_{i} \geqslant 0$ for all $i$.

Now let $S \in \mathcal{L}(V, V)$ be the unique linear operator such that $S v_{i}=\sqrt{\lambda_{i}} v_{i}$. Since $\left\{v_{1}, \ldots, v_{n}\right\}$ is orthonormal you can check that $S$ is self adjoint. Once you know that, observe $S^{*} S v_{i}=\lambda_{i} v_{i}=T v_{i}$. Since defining a linear transformation on a basis determines it uniquely, we must have $S^{*} S=T$.
3. We say $T \in \mathcal{L}(V, V)$ is unitary if $T^{*}=T^{-1}$. (Note, from a similar question on your previous homework, this is equivalent to saying $\langle T u, T v\rangle=\langle u, v\rangle$. .)
(a) Suppose $T$ is unitary, and $B$ is any orthonormal basis of $V$. Show that $\overline{\mathcal{M}}_{B}(T)=\mathcal{M}_{B}(T)^{-1}$. [NoTE: This result is not true if $B$ is not orthonormal.]

Solution. This follows quickly from a more general observation. Namely $\mathcal{M}_{B}\left(T^{*}\right)=\overline{\mathcal{M}}_{B}(T)$, as long as $B$ is orthonormal (regardless of weather $T$ is unitary or not).
To see this, let $B=\left\{u_{1}, \ldots, u_{n}\right\}, \mathcal{M}_{B}(T)=\left(a_{i, j}\right)$, and $\mathcal{M}_{B}\left(T^{*}\right)=\left(b_{i, j}\right.$. Then $\left\langle T u_{i}, u_{j}\right\rangle=a_{i, j}$ by definition of $\mathcal{M}_{B}(T)$, and because $B$ is orthonormal.

On the other hand, $\left\langle T u_{i}, u_{j}\right\rangle=\left\langle u_{i}, T^{*} u_{j}\right\rangle=\bar{b}_{j, i}$ by definition of $\mathcal{M}_{B}\left(T^{*}\right)$, and because $B$ is orthonormal. Consequently $a_{i, j}=\bar{b}_{j, i}$, proving $\mathcal{M}_{B}\left(T^{*}\right)={\overline{\mathcal{M}_{B}(T)}}^{\mathrm{t}}$.

Now if $T^{*}=T^{-1}$, the claim immediately follows.
(b) Conversely, suppose there exists an orthonormal basis $B$ of $V$ such that ${\overline{\mathcal{M}_{B}(T)}}^{\mathrm{t}}=\mathcal{M}_{B}(T)^{-1}$, then show that $T$ is unitary.

Solution. Very similar to reversing the previous subpart.

Assume subsequently that $U, V$ be are two finite dimensional inner product spaces.
4. If $T \in \mathcal{L}(U, V)$ is injective, show that it's singular values are nonzero.

Solution. Recall that the singular values of $T$ are exactly the square root of eigenvalues of $T^{*} T$. However, if $T$ is injective, we've seen that $T^{*} T$ is invertible, so 0 is not an eigenvalue of $T^{*} T$ ! Consequently, 0 is not a singular value of $T$.
5. Say $\left\{u_{1}, \ldots, u_{m}\right\},\left\{v_{1}, \ldots, v_{n}\right\}$ are orthonormal basis of $U$ and $V$ respectively, and $\sigma_{1} \geqslant \cdots \geqslant \sigma_{m}$ are the singular values of $T$. Let $r=\operatorname{rank}(T)$ be the largest number such that $\sigma_{r} \neq 0$. Let $T^{+} \in \mathcal{L}(V, U)$ be the unique linear transformation such that $T v_{i}=\frac{1}{\sigma_{i}} u_{i}$ for $i \leqslant r$, and $T v_{i}=0$ for $i \geqslant r$. This is called the pseudo-inverse of $T$. [Of course, $T \in \mathcal{L}(U, V)$ need not be invertible. If $T$ is invertible, then $T^{+}=T^{-1}$, as you see from (b) below. If $T$ is not invertible, then $T^{+}$is the next best thing to $T^{-1}$ (as you also see from (b) below).]
(a) Show that $T^{+} T=I$ is the orthogonal projection of $U$ onto $\operatorname{ker}(T)^{\perp}$.

Solution. Note first that $\operatorname{ker}(T)=\operatorname{span}\left\{u_{r+1}, \ldots, u_{m}\right\}$. This is because if $T\left(\sum_{i=1}^{m} a_{i} u_{i}\right)=0$, then $\sum_{i=1}^{m} \sigma_{i} a_{i} v_{i}=0$, which implies $\sigma_{i} a_{i}=0$ for all $i \leqslant m$, showing $a_{i}=0$ for all $i \leqslant r$. This will immediately show $\operatorname{ker}(T)=$ $\operatorname{span}\left\{u_{r+1}, \ldots, u_{m}\right\}$. From this, it follows that $\operatorname{ker}(T)^{\perp}=\operatorname{span}\left\{u_{1}, \ldots, u_{r}\right\}$.
Let $P \in \mathcal{L}\left(U, \operatorname{ker}(T)^{\perp}\right)$ be the orthogonal projection of $P$ onto $\operatorname{ker}(T)^{\perp}$. Observe that for $i \leqslant r, T^{+} T\left(u_{i}\right)=T^{+}\left(\sigma_{i} v_{i}\right)=u_{i}=P u_{i}$. For $i>r$, then $T^{+} T\left(u_{i}\right)=T^{+}(0)=0=P u_{i}$. Consequently, $T^{+} T$ and $P$ are equal on basis vectors, and hence, by linearity must be equal.
(b) Show that $T T^{+}$is the orthogonal projection of $V$ onto $\operatorname{im}(T)$.

Solution. Similar to the previous part.
(c) If $T$ is injective, show $T^{+}=\left(T^{*} T\right)^{-1} T^{*}$.

Solution. Let $\delta_{i, j}=1$ if $i=j$, and 0 if $i \neq j$. Then

$$
\sigma_{i} \delta_{i, j}=\left\langle T u_{i}, v_{j}\right\rangle=\left\langle u_{i}, T^{*} v_{j}\right\rangle
$$

which forces $T^{*} v_{i}=\sigma_{i} u_{i}$.
Now let $\lambda_{i}=\sigma_{i}^{2}$ be the eigenvalues of $\left(T^{*} T\right)$. If $T$ is injective, $T^{*} T$ is invertible, and $\lambda_{i} \neq 0$ for all $i$. Thus $\left(T^{*} T\right)^{-1}\left(u_{i}\right)=\frac{1}{\lambda_{i}} u_{i}$. And so $\left(T^{*} T\right)^{-1} T v_{i}=\frac{1}{\lambda_{i}} \sigma_{i} u_{i}=\frac{1}{\sigma_{i}} u_{i}=T^{+} u_{i}$. Thus $\left(T^{*} T\right)^{-1} T$ and $T^{+}$are equal on basis vectors, and hence, by linearity, must be equal.
(d) Suppose now $A \in \operatorname{Mat}(n, m, \mathbb{C})$ and $T \in \mathcal{L}\left(C^{n}, C^{m}\right)$ is given by multiplication with $A$. Let $A=P \Sigma Q^{*}$ be the singular value decomposition of $A$. Show that $T^{+}$is multiplication by the matrix $Q \Sigma^{+} P^{*}$, where $\Sigma^{+}$ is obtained by taking the reciprocal of all the non-zero entries in $\Sigma$, and transposing the result.

Solution. This is exactly the matrix definition of $T^{+}$. Computing what $T^{+}$ does to columns of $Q$ will immediately give the result.

