Fall 2002 ARE211

Mid-Term Answer Key

1. Let $d$ be the discrete metric on $S$ and let $\rho$ be the usual metric on $T$. Pick $\bar{s} \in S$ and suppose that $s_n$ converges to $\bar{s}$. Since $d$ is the discrete metric, there exists $N \in \mathbb{N}$ such that for $n > N$, $s_n = \bar{s}$. Hence for $n > N$, $f(s_n) = f(\bar{s})$. Hence for all $\varepsilon > 0$ and all $n > N$, $f(s_n) \in B(f(\bar{s}), \varepsilon)$. Hence $(f(s_n))$ converges to $f(\bar{s})$.

2. “if:” Assume that $f$ is a locally constant function. We’ll show that $f \in F^0$. Pick $\bar{s}$ arbitrarily from $S$ and an arbitrary sequence $(s_n)$ in $S$ that converges to $\bar{s}$. Since $f$ is a locally constant function, there exists $\varepsilon > 0$ such that for $s' \in B(\bar{s}, \varepsilon)$, $f(s') = f(\bar{s})$. Moreover, since $(s_n)$ converges to $\bar{s}$, there exists $N$ such that for $n > N$, $s_n \in B(\bar{s}, \varepsilon)$ and hence $f(s_n) = f(\bar{s})$. Hence for all $\varepsilon > 0$ and all $n > N$, $f(s_n) \in B(f(\bar{s}), \varepsilon)$. Hence $(f(s_n))$ converges to $f(\bar{s})$.

“only if:” Now assume that $f$ is not a locally constant function. In this case, there exists $\bar{s} \in S$ such that for all $\varepsilon > 0$, there exists $s' \in B(\bar{s}, \varepsilon)$ such that $f(s') \neq f(\bar{s})$. Pick a sequence $s_n$ such that for all $n$, $s_n \in B(\bar{s}, \varepsilon)$ and $f(s_n) \neq f(\bar{s})$. Clearly, $s_n$ converges to $\bar{s}$, but for all $n$, $\rho(f(s_n), f(\bar{s})) = 1$. Hence $f(s_n)$ does not converge to $f(\bar{s})$.

3. (a) Let $d$ be the usual metric on $S$ and $\rho$ be the discrete metric on $T$.

(i) Suppose $f$ is constant and $f(s) = t$ for all $s$. Pick $s \in S$ and $\varepsilon > 0$ arbitrarily. For all $s' \in B(x, \varepsilon)$, $f(s') = t = f(s)$. Hence $f$ is locally constant.

(ii) Let $S = \{-1, 1\}$ and for $s \in S$, let $f(s) = s$. Clearly $f$ is locally constant but not constant.
(4) (a) Consider a sequence \((x_n)\) defined on a bounded set \(S \subseteq \mathbb{R}\) which has no convergent subsequence. Suppose that \(S\) is contained in the interval \([-b/2, b/2]\) for some \(b \in \mathbb{R}\). Proceed exactly as in the proof of the Bolzano-Weierstrass theorem (Lecture #9) to construct a Cauchy subsequence \((y_n)\) of \((x_n)\). Since \((y_n)\) is a Cauchy sequence in \(\mathbb{R}\), it must converge to some point \(y \in \mathbb{R}\). If \(y\) belonged to \(S\), then \((y_n)\) would be a convergence subsequence of \((x_n)\). Since by assumption, no such subsequence exists, we can conclude that \(y \notin S\). Since \((y_n)\) is a sequence in \(S\), then \(y_n \neq y\), for all \(n\). But by construction, for all \(\varepsilon > 0\), there exists \(n \in \mathbb{N}\) such that \(y_n \in B(y, \varepsilon)\). Hence \(y\) is an accumulation point of \(S\) which is not contained in \(S\). Hence \(S\) cannot be closed.

(b) We need to prove that if every sequence in \(A\) has a convergent subsequence then \(A\) is compact. Suppose \(A\) is not compact. We will show that there exists a sequence in \(A\) that has no convergent subsequence. There are two cases to consider:

\(A\) **is not bounded** We’ll assume that \(A \subseteq \mathbb{R}^n\) is not bounded above. In this case, for all \(N\), there exists \(n > N\) and \(a_n \in A\) such that \(||a_n|| > N\). Clearly, every subsequence of \((a_n)\) increases without bound and therefore has no convergent subsequence.

\(A\) **is not closed** In this case, there exists an accumulation point \(a\) of \(A\), such that \(a \notin A\). That is, for all \(n\), there exists \(a_n \in A\) such that \(a_n \in B(a, 1/n)\). Clearly, for every subsequence \((b_n)\) of \((a_n)\), \(\lim_{n} b_n = a\). But since \(a \notin A\), \((b_n)\) is not a convergent subsequence.

(5)

(6)
To show that $P$ is a vector space, we need to show that for any $f_1, f_2 \in P$, and any $\alpha_1, \alpha_2 \in \mathbb{R}$, $g = \alpha_1 f_1 + \alpha_2 f_2 \in P$. For $i = 1, 2$, $f_i \in P_k(i)$, for some $k(i) \in \mathbb{N} \cup \{0\}$. Let $K = \max\{k(1), k(2)\}$, and note that for $i = 1, 2$, there exists $\{a_{0,i}, a_{1,i}, \ldots, a_{K,i}\}$, with $a_{k,i} = 0$, for $k > k(i)$ such that $f_i = \sum_{k=0}^{K} a_{k,i} x^k$. Define $\{b_0, b_1, \ldots, b_K\}$ by, for $\kappa = 0, 1, \ldots, K$, $b_{\kappa} = \alpha_1 a_{\kappa,1} + \alpha_2 a_{\kappa,2}$ and observe that $g(x) = \sum_{k=0}^{K} b_{k} x^k$. Now define $n$ as follows:

$$n = \begin{cases} 0 & \text{if } b_k = 0, k \in \{1, \ldots, K\} \\ \max\{k \in \{1, \ldots, K\} : b_k \neq 0\} & \text{otherwise} \end{cases}$$

and note that if $n > 0$, $g(x) = \sum_{k=0}^{n} b_k x^k$, with $b_n \neq 0$. Therefore, $g \in P_n$. On the other hand, if $n = 0$, $g(x) = 0$, for all $x \in R$. Hence $g \in P_0$. Conclude that $g \in \cup_{n=0}^{\infty} P_n = P$, proving that $P$ is a vector space.

Let $f_1(x) = x$ and $f_2(x) = 0$. Let $\alpha_1 = -1$ and $\alpha_2 = 0$. Clearly $g(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) = -x$ and so $g$ does not belong to $P$.

Pick $(p_n)$ arbitrarily such that for each $n$, $p_n \in P_n$, and $p_0$ is not the zero function. For each $n$, write $p_n(x)$ as $\sum_{k=0}^{\infty} b_{k,n} x^k$ and observe that $b_{n,n} \neq 0$, while for $k > n$, $b_{k,n} = 0$. It’s true by construction that the set of polynomials $R = \{p_0, p_1, \ldots, p_n, \ldots\}$ belongs to $P$. Therefore, to establish that $R$ is a basis for $P$ we need to show that $R$ is a minimal spanning set for $P$. We first show that $R$ spans $P$. Pick $p \in P$. Necessarily, $p \in P^n$ for some $n$. We’ll assume $n > 0$, otherwise the problem is trivial. Write $p(x)$ as $\sum_{k=0}^{\infty} a_k x^k$ and observe that $a_0 \neq 0$, while for $k > n$, $a_k = 0$. We will now construct a set of weights $(\alpha_k)_{k=0}^{\infty}$ such that $p = \sum_{k=0}^{\infty} \alpha_k p_k$. We first set $\alpha_k = 0$, for $k > n$. Now define $\alpha_n = a_n/b_{n,n}$, and note that if we ignore the first $n$ terms of $p$ (i.e., the zero’th thru $n-1’$th terms), then $p$ does indeed equal $\sum_{k=0}^{n} \alpha_k p_k$.

Now for the inductive step. For $1 \leq \kappa < n$, assume that $\alpha_r$ has been defined, for $r > \kappa$ in such a way that if we ignore the first $\kappa + 1$ terms of $p$, then $p = \sum_{k=0}^{\infty} \alpha_k p_k$. (This assumption is certainly satisfied when $\kappa = n - 1$.) We now set $\alpha_\kappa = \frac{a_\kappa - \sum_{r=\kappa+1}^{\infty} \alpha_r b_{\kappa,r}}{b_{\kappa,\kappa}}$. Note first that

$$\sum_{r=\kappa}^{n} \alpha_r b_{\kappa,r} = \sum_{r=\kappa+1}^{n} \alpha_r b_{\kappa,r} + b_{\kappa,\kappa} \frac{a_\kappa - \sum_{r=\kappa+1}^{\infty} \alpha_r b_{\kappa,r}}{b_{\kappa,\kappa}} = a_\kappa$$

while, since $b_{\kappa,r} = 0$, for $r < \kappa$ and $\alpha_r = 0$, for $r > n$

$$\sum_{r=0}^{\kappa-1} \alpha_r b_{\kappa,r} + \sum_{r=n+1}^{\infty} \alpha_r b_{\kappa,r} = 0$$

Conclude that

$$\sum_{r=0}^{n} \alpha_r b_{\kappa,r} = a_\kappa$$

We have thus shown that if we ignore the first $\kappa$ terms of $p$, then $p = \sum_{k=0}^{\infty} \alpha_k p_k$. This completes the inductive argument that any $p \in P$ can be written as a linear combination of the vectors in $R$.

To conclude the argument we need to show that $(p_n)$ is a minimal spanning set. This proves to be quite challenging if you allow for the possibility of linear combinations of an infinite number of vectors. I’ll explain more about this below. But since the difficulty didn’t seem to bother any of you, we won’t worry about it too much.
Suppose that for arbitrary \( n \), \( p_n \) were omitted from \( R \). We will show that \( p_n \) cannot be written as a finite linear combination of the other elements of \( P \), i.e., a linear combination in which only a finite number of elements of \( P \) are assigned non-zero weight. To see this, let \( q = \sum_{k=1}^{K} \alpha_k p_{m_k} \) be any finite linear combination with for each \( k \), \( \alpha_k \neq 0 \) and \( p_{m_k} \in P^{m_k} \). There are two cases to consider, depending on the value of \( \max\{m_k\} \). Since \( \max\{m_k\} \neq n \), the two cases are:

(i) \( m = \max\{m_k\} > n \), i.e., the highest order polynomial in the (finite) set is of higher order than \( p_n \) the polynomial that was omitted. In this case, the weight on \( x^m \) is \( \alpha_m \) times the weight that \( p_m \) assigns to \( x^m \) (since all other terms in the combination are lower order polynomials). Hence \( q \) is an \( m \)-order polynomial and cannot be equal to \( p_n \).

(ii) \( m = \max\{m_k\} < n \), i.e., the highest order polynomial in the set is of lower order than \( p_n \).

In this case, for all of the the polynomials that are assigned positive weight, the weight assigned to \( x^n \) is zero. Hence the weight assigned by \( q \) to \( x^n \) must also be zero, i.e., \( q \) is an \( m \)-order polynomial and cannot be equal to \( p_n \).

Notice the point where this argument would break down if we hadn’t restricted our attention to finite linear combinations: in this case, we wouldn’t have been able to identify a highest order polynomial in the set of polynomials that is assigned a positive weight.
(9) (a) Following Simon-Blume, to solve for the eigenvalues, you solve the quadratic equation: 
\[ \lambda^2 - (3 + x)\lambda + (3x - 1) = 0. \]
Now apply the quadratic formula, i.e. to solve \( ax^2 + bx + c \) 
you set 
\[ x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \]
In our case this gives the following values for \( \lambda \):
\[
\begin{align*}
\lambda_1 &= \left(3 + x\right) - \sqrt{\left(13 - 6x + x^2\right)} / 2; \\
\lambda_2 &= \left(3 + x\right) + \sqrt{\left(13 - 6x + x^2\right)} / 2; \quad (9.1)
\end{align*}
\]
Now given \( \lambda_i \) we solve the equation \( Av_i = \lambda_i v_i \) to obtain the corresponding eigenvector, i.e., That is, \( v_1^i \left(3 - \lambda_i\right) = -v_2^i \). Since \( v^i \) has norm one, we also know that \( v_2^i = \sqrt{1 - (v_1^i)^2} \). Combining these two expressions yields
\[
\begin{align*}
v_1^i &= \frac{1}{\sqrt{\left(3 - \lambda_i\right)^2 + 1}} = \frac{1}{\sqrt{10 - 6\lambda_i + \left(\lambda_i\right)^2}} \\
v_2^i &= -v_1^i \left(3 - \lambda_i\right)
\end{align*}
\]
Note that the \( v^i \)'s do depend on \( x \), but only through the \( \lambda_i \)'s.

(b) In order to have
(i) \( v^i A v > 0 \) for most but not all \( v \)'s, you need one negative \( \lambda \) and a larger positive \( \lambda \).
(ii) \( v^i A v < 0 \) for most but not all \( v \)'s, you need one positive \( \lambda \) and a larger negative \( \lambda \).
(iii) \( v^i A v > 0 \) for all \( v \)'s, you need two positive \( \lambda \)'s.

To accomplish this, go back to expression (9.1) and pick appropriate values for \( x \). For example
(i) set \( x = -1 \) so that

\[
\lambda_1 = \left( 3 - \sqrt{20} \right) / 2 < 0;
\]

\[
\lambda_2 = \left( 3 + \sqrt{20} \right) / 2 > -\lambda_1
\]

Note in the figure above that \( \mathbf{v}_2 \) is stretched by more than \( \mathbf{v}_1 \), which is flipped through 180 degrees. Vectors in the shaded area of the unit circle make an acute angle with their images under \( A \). e.g. \( \mathbf{w}_1 \) is mapped to \( A\mathbf{w}_1 \), \( \mathbf{w}_2 \) is mapped to \( A\mathbf{w}_2 \) while vectors in between \( \mathbf{w}_1 \) and \( \mathbf{w}_2 \) are mapped into the nonnegative cone defined by \( A\mathbf{w}_1 \) and \( A\mathbf{w}_2 \). Note that more than half of the unit circle is shaded, reflecting the fact that with probability greater than 0.5 (but less than one), a vector chosen at random will make an acute angle with its image.
(ii) set $x = -5$ so that

\[
\lambda_1 = \frac{-2 - \sqrt{68}}{2} < 0; \\
\lambda_2 = \frac{-2 + \sqrt{68}}{2} \in [0 - \lambda_2)
\]

In this case, note in the figure above that $v^2$ is stretched by less than $v^1$, which again is flipped through 180 degrees. Vectors in the shaded area of the unit circle again make an acute angle with their images under $A$. Again, vectors in between $w^1$ and $w^2$ are mapped into the nonnegative cone defined by $Aw^1$ and $Aw^2$. Note that in this case less than half of the unit circle is shaded, reflecting the fact that with probability greater than 0.5 (but less than one), a vector chosen at random will make an obtuse angle with its image.
(iii) set $x = 1$ so that

$$\lambda_1 = \left(4 - \sqrt{13}\right)/2 > 0;$$

$$\lambda_2 = \left(4 + \sqrt{13}\right)/2 > 0$$

I didn’t bother to draw the picture for this one.