Linear System Theory

Solutions to Problem Set 1
Linear Spaces, Linear Maps, and Representations

Issue date: September 23, 2019
Due date: October 7, 2019

Solution 1. (Linear spaces [40 **points in total**]**)**

1. [18 **points in total**] We verify the axioms of a vector space.

   - Associativity of $\oplus$: trivial.
   - The identity element of $\oplus$ (that is, the zero vector) is the constant function $\theta(x) = 1$ for all $x \in S$. Indeed $[\theta \oplus f](x) = \theta(x) f(x) = 1 f(x) = f(x)$ for all $x \in S$.
   - The inverse of a function f is the function f^- defined as $f^-(x) = \frac{1}{f(x)}$ for all $x \in S$. Indeed $[f^- \oplus f](x) = f^-(x) f(x) = \frac{1}{f(x)} f(x) = 1 = \theta(x)$ for all $x \in S$.
   - Commutativity of $\oplus$: trivial.
   - Associativity of $\odot$: for all $\alpha, \beta \in \mathbb{R}$ and $f \in F$, $[\alpha \odot [\beta \odot f]](x) = [\beta \odot f](x)^\alpha = f(x)^{\alpha\beta} = [(\alpha\beta) \odot f](x)$.
   - Multiplication identity: $[1 \odot f](x) = f(x)^1 = f(x)$.
   - Distributivity, first property: $[(\alpha + \beta) \odot f](x) = f(x)^{\alpha+\beta} = f(x)^\alpha f(x)^\beta = [\alpha \odot f](x)[\beta \odot f](x) = [[\alpha \odot f] \oplus [\beta \odot f]](x)$.
   - Distributivity, second property: $[\alpha \odot [f_1 \oplus f_2]](x) = (f_1(x) f_2(x))^\alpha = f_1(x)^\alpha f_2(x)^\alpha = [\alpha \odot f_1](x)[\alpha \odot f_2](x) = [[\alpha \odot f_1] \oplus [\alpha \odot f_2]](x)$.

2. [12 **points in total**] Suppose that $\alpha \odot f_1 \oplus \beta \odot f_2 = \theta$. Then

$$
\begin{aligned}
1 &= [\alpha \odot f_1 \oplus \beta \odot f_2](a) = f_1(a)^\alpha f_2(a)^\beta = 2^\alpha 1^\beta = 2^\alpha \\
1 &= [\alpha \odot f_1 \oplus \beta \odot f_2](b) = f_1(b)^\alpha f_2(b)^\beta = 1^\alpha 3^\beta = 3^\beta
\end{aligned}
$$

   Therefore, it must hold $\alpha = \beta = 0$, and $\{f_1, f_2\}$ are linearly independent. To show that $\{f_1, f_3\}$ are linearly dependent, it suffices to recognize that $f_3(x) = f_1(x)^2$, or $f_3 = 2 \odot f_1$; hence $2 \odot f_1 \oplus (-1) \odot f_3 = \theta$ with nonzero coefficients.

3. [10 **points in total**] We verify the definition of linearity: For all $f_1, f_2, f \in F$ and $\alpha \in \mathbb{R}$,

$$
\begin{aligned}
[\varphi(f_1 \oplus f_2)](x) &= \sqrt{[f_1 \oplus f_2](x)} \\
&= \sqrt{f_1(x) f_2(x)} \\
&= \sqrt{f_1(x)} \sqrt{f_2(x)} \\
&= [\varphi(f_1) \oplus \varphi(f_2)](x)
\end{aligned}
$$

and

$$
\begin{aligned}
[\varphi(\alpha \odot f)](x) &= \sqrt{[\alpha \odot f](x)} \\
&= \sqrt{f(x)^{\alpha}} \\
&= \left(\sqrt{f(x)}\right)^{\alpha} \\
&= [\alpha \odot \varphi(f)](x).
\end{aligned}
$$

Solution 2. (Range and null space [40 **points in total**]**)**

1. [10 **points in total**] Note that by definition $\mathit{dim}(\textsc{Range}(\mathcal{A}))$ is nonnegative. We also have $\textsc{Range}(\mathcal{A}) \subseteq F^m$, and therefore $\mathit{dim}(\textsc{Range}(\mathcal{A})) \leq m$.

   Let $\{b_1, \ldots, b_n\}$ be a basis of F^n. Then, $\textsc{Range}(\mathcal{A}) \subseteq \operatorname{span}\{\mathcal{A}(b_1), \ldots, \mathcal{A}(b_n)\}$. This follows since for any $y \in \textsc{Range}(\mathcal{A})$, $\exists x \in F^n$ such that $\mathcal{A}(x) = y$, and x can be written as $x = \sum_{i=1}^{n} \xi_i b_i$, and thus $y = \mathcal{A}(x) = \sum_{i=1}^{n} \xi_i \mathcal{A}(b_i) \in \operatorname{span}\{\mathcal{A}(b_1), \ldots, \mathcal{A}(b_n)\}$. We obtain

   $$\mathit{dim}(\textsc{Range}(\mathcal{A})) \leq \mathit{dim}(\operatorname{span}\{\mathcal{A}(b_1), \ldots, \mathcal{A}(b_n)\}) \leq n.$$

   Finally, we have:

   $$0 \leq \mathit{dim}(\textsc{Range}(\mathcal{A})) \leq \min\{m, n\}.$$

2. [15 **points in total**]

   Our proof approach is as follows. We first show that:

   $$\mathit{dim}(\textsc{Null}(\mathcal{B} \circ \mathcal{A})) \leq \mathit{dim}(\textsc{Null}(\mathcal{A})) + \mathit{dim}(\textsc{Null}(\mathcal{B})),$$

   Then, we use the rank-nullity theorem to obtain the statement given in the problem.

   Clearly, we have $\textsc{Null}(\mathcal{A}) \subseteq \textsc{Null}(\mathcal{B} \circ \mathcal{A})$. This follow since if $x \in F^n$ is such that $\mathcal{A}(x) = 0$, then $\mathcal{B}(\mathcal{A}(x)) = 0$. Next, consider $\tilde{\mathcal{A}}$ as a restriction of $\mathcal{A}$:

   $$\tilde{\mathcal{A}} : \textsc{Null}(\mathcal{B} \circ \mathcal{A}) \to F^m.$$

   For the map above, we have $\textsc{Null}(\tilde{\mathcal{A}}) = \textsc{Null}(\mathcal{A})$. Moreover, since $\textsc{Range}(\tilde{\mathcal{A}})$ is contained in $\textsc{Null}(\mathcal{B})$ by construction, we have $\mathit{dim}(\textsc{Range}(\tilde{\mathcal{A}})) \leq \mathit{dim}(\textsc{Null}(\mathcal{B}))$. The rank-nullity theorem applied to $\tilde{\mathcal{A}}$ gives

   $$\mathit{dim}(\textsc{Null}(\mathcal{B}\circ\mathcal{A})) = \mathit{dim}(\textsc{Null}(\tilde{\mathcal{A}})) + \mathit{dim}(\textsc{Range}(\tilde{\mathcal{A}})) \leq \mathit{dim}(\textsc{Null}(\mathcal{A})) + \mathit{dim}(\textsc{Null}(\mathcal{B})).$$

   Subtracting n from both sides and multiplying by -1, we get

   $$n - \mathit{dim}(\textsc{Null}(\mathcal{B} \circ \mathcal{A})) \geq n - \mathit{dim}(\textsc{Null}(\mathcal{A})) - \mathit{dim}(\textsc{Null}(\mathcal{B})).$$

   Finally, applying the rank-nullity theorem again, we obtain the desired result

   $$\mathit{dim}(\textsc{Range}(\mathcal{B} \circ \mathcal{A})) \geq \mathit{dim}(\textsc{Range}(\mathcal{A})) + \mathit{dim}(\textsc{Range}(\mathcal{B})) - m.$$

Alternative Solution:[15 **points in total**]

Consider $\tilde{\mathcal{B}}$ as a restriction of $\mathcal{B}$:

$$\tilde{\mathcal{B}} : \textsc{Range}(\mathcal{A}) \to F^p.$$

It follows that $\textsc{Null}(\tilde{\mathcal{B}}) \subseteq \textsc{Null}(\mathcal{B})$. Moreover, $\textsc{Range}(\tilde{\mathcal{B}}) = \textsc{Range}(\mathcal{B} \circ \mathcal{A})$. Applying the rank-nullity theorem to the map $\tilde{\mathcal{B}}$ and subsequently to the map $\mathcal{B}$, we obtain

$$\begin{aligned} dim(\textsc{Range}(\mathcal{A})) &= dim(\textsc{Null}(\tilde{\mathcal{B}})) + dim(\textsc{Range}(\tilde{\mathcal{B}})) \\ &\leq dim(\textsc{Null}(\mathcal{B})) + dim(\textsc{Range}(\mathcal{B} \circ \mathcal{A})) \\ &= m - dim(\textsc{Range}(\mathcal{B}) + dim(\textsc{Range}(\mathcal{B} \circ \mathcal{A})) \end{aligned}$$

We conclude that

$$dim(\textsc{Range}(\mathcal{B} \circ \mathcal{A})) \geq dim(\textsc{Range}(\mathcal{A})) + dim(\textsc{Range}(\mathcal{B})) - m.$$

3. [15 **points in total**]

   We have $\textsc{Range}(\mathcal{B}\circ\mathcal{A}) \subseteq \textsc{Range}(\mathcal{B})$. This follows since for any $y \in \textsc{Range}(\mathcal{B}\circ\mathcal{A})$, $\exists x \in F^n$ such that $\mathcal{B}(\mathcal{A}(x)) = y$. Letting $z = \mathcal{A}(x) \in F^m$ implies that $\mathcal{B}(z) = y$ and $y \in \textsc{Range}(\mathcal{B})$. Thus,

   $$dim(\textsc{Range}(\mathcal{B} \circ \mathcal{A})) \leq dim(\textsc{Range}(\mathcal{B})).$$

   Next, assume $x \in \textsc{Null}(\mathcal{A})$, i.e., $\mathcal{A}(x) = 0$. Then, we have $x \in \textsc{Null}(\mathcal{B} \circ \mathcal{A})$, since $\mathcal{B}(\mathcal{A}(x)) = 0$. This implies that $\textsc{Null}(\mathcal{A}) \subseteq \textsc{Null}(\mathcal{B} \circ \mathcal{A})$ and

   $$dim(\textsc{Null}(\mathcal{A})) \leq dim(\textsc{Null}(\mathcal{B} \circ \mathcal{A})).$$

   Subtracting n and changing sign, we obtain

   $$n - dim(\textsc{Null}(\mathcal{A})) \geq n - dim(\textsc{Null}(\mathcal{B} \circ \mathcal{A})).$$

   We can use the rank-nullity theorem and conclude

   $$dim(\textsc{Range}(\mathcal{B} \circ \mathcal{A})) \leq dim(\textsc{Range}(\mathcal{A})).$$

   Finally, we have that:

   $$dim(\textsc{Range}(\mathcal{B} \circ \mathcal{A})) \leq \min\{dim(\textsc{Range}(\mathcal{A})), dim(\textsc{Range}(\mathcal{B}))\}.$$

Solution 3. (Linear maps and matrix representations [20 **points in total**]**)**

1. [10 **points in total**] For $i = 1, \ldots, n$, the representation of ν_i in the basis $\{\nu_1, \ldots, \nu_n\}$ is $e_i = [0, \ldots, 0, 1, 0, \ldots, 0]^T$, where the "1" is in the i-th position.

   Since $\mathcal{A}(\nu_i) = \lambda\nu_i + \nu_{i+1}$, the representation A of $\mathcal{A}$ must satisfy $Ae_i = \lambda e_i + e_{i+1} = [0, \ldots, 0, \lambda, 1, 0, \ldots, 0]^T$, where the "$\lambda$" and the "1" are in the i-th and $(i+1)$-th positions, respectively. Moreover, $\mathcal{A}(\nu_n) = \lambda\nu_n$ implies that $Ae_n = \lambda e_n$. Finally, since Ae_i is the i-th column of A,

$$A = \begin{bmatrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{bmatrix},$$

where the entries not shown are zeroes [10 **points**].

2. [10 **points in total**] For $i = 1, \ldots, n$, let us define $b_i = \mathcal{A}^{i-1}(b)$, so that the basis becomes $\{b_1, \ldots, b_n\}$. In vector form, the representation of b_i is given by $e_i = [0, \ldots, 0, 1, 0, \ldots, 0]^T$, where the "1" is in the i-th position.

   For $i = 1, \ldots, n-1$, it holds that $\mathcal{A}(b_i) = \mathcal{A}\big(\mathcal{A}^{i-1}(b)\big) = \mathcal{A}^i(b) = b_{i+1}$, hence the representation A of $\mathcal{A}$ must satisfy $Ae_i = e_{i+1}$. Finally, since Ae_i is the i-th column of A,

$$A = \begin{bmatrix} 0 & & & & \beta_1 \\ 1 & 0 & & & \beta_2 \\ & \ddots & \ddots & & \vdots \\ & & 1 & 0 & \beta_{n-1} \\ & & & 1 & \beta_n \end{bmatrix},$$

   where the entries not shown are zeroes and $\beta = [\beta_1, \ldots, \beta_n]^T$ is determined as follows: since $\beta = Ae_n$ and e_n is the representation of b_n, β is the representation of $\mathcal{A}(b_n)$ with respect to the basis $\{b_1, \ldots, b_n\}$ [10 **points**].