Homework #4 Solutions

Thayer Anderson, Davis Lazowski, Handong Park, Rohil Prasad
Eric Peterson

1 For submission to Thayer Anderson

Problem 1.1. Suppose that V is a vector space over a field of scalars K with $\dim V = 1$. Show that each element $f \in \mathcal{L}(V, V)$ has the form $f(v) = \lambda_f \cdot v$ for some scalar λ_f (which depends on f but not on v). Conclude from this that there is an isomorphism of vector spaces

$$\mathcal{L}(V, V) \simeq K$$

which *does not depend* upon choosing a basis for V.

Solution. The set $\mathcal{L}(V, V)$ has vector-space structure over K and dimension $(\dim V)^2 = 1$. The identity map, id_V is a linear map in $\mathcal{L}(V, V)$. It follows that $\{\mathrm{id}_V\}$ is a basis for $\mathcal{L}(V, V)$. As a basis, this set is spanning, and thus if $f \in \mathcal{L}(V, V)$ then $f = \lambda_f \cdot \mathrm{id}_V$ for some $\lambda \in K$. It follows that $f(v) = \lambda_f \cdot \mathrm{id}_V(v) = \lambda_f \cdot v$ for all $v \in V$.

Let the map $\varphi : \mathcal{L}(V, V) \to K$ be defined by its action on a basis: $\mathrm{id}_V \mapsto 1$. We see that $\varphi(f) = \lambda_f$. We must prove this map is linear, surjective, and injective. To show linearity, first consider $f + g$ for linear maps f and g.

$$\begin{aligned} (f + g)(v) = f(v) + g(v) = \lambda_f(v) + \lambda_g(v) = (\lambda_f + \lambda_g)(v) \\ \Rightarrow \lambda_{f+g} = \lambda_f + \lambda_g \end{aligned}$$

Then we see

$$\varphi(f + g) = \lambda_{f+g} = \lambda_f + \lambda_g.$$

Similarly for scalar multiplication, consider the map $c \cdot f$ for $c \in K$ and $f \in \mathcal{L}(V, V)$. Then

$$(c \cdot f)(v) = c \cdot f(v) = c\lambda_f \cdot v = (c\lambda_f)v$$

Thus

$$\varphi(cf) = c\lambda_f.$$

We conclude that φ is linear and since it maps a basis to a basis it is an isomorphism. Alternatively, we can see that $\varphi(f) = 0$ only if $f = 0$. Then it follows that φ is a bijection. (TA)

Solution. Here's an alterative solution. Take some non-zero vector $v \in V$. Then $\{v\}$ forms a basis for V. Since the vector $f(v) \in V$ it follows that $f(v) \in \mathrm{span}(\{v\})$. Thus $f(v) = \lambda_{v,f} \cdot v$. Suppose that $\lambda_{v',f} \neq \lambda v, f$ for two non-zero vectors v and v'. Then, by the fact that v spans, we can say $v' = c \cdot v$. We have

$$\lambda_{v',f} v' = f(v') = f(cv) = cf(v) = c\lambda_{v,f} v = \lambda_{v,f} cv = \lambda_{v,f} v'$$

We conclude that $\lambda_{v',f} = \lambda_{v,f}$ and thus the λ_f is determined by f and is basis independent. In addition, the definition of $f(v) = \lambda_f(v)$ is consistent at $v = 0$. Then we prove the existence of the desired isomorphism as in previous solution. (TA)

Problem 1.2. Suppose that $f : V \to W$ and $g : W \to X$ are linear maps

1. Show that $\dim\ker(g \circ f) \le \dim\ker f + \dim\ker g$.
2. Show that $\dim\operatorname{im}(g \circ f) \le \min\{\dim\operatorname{im} f, \dim\operatorname{im} g\}$.

Solution.

1. Consider the map $h := f|_{\ker(g\circ f)} : \ker(g \circ f) \to W$. Since kernels are subspaces, this map makes sense. Apply the fundamental theorem for linear maps to h:

   $$\dim\ker(g \circ f) = \dim\ker h + \dim\operatorname{im} h$$

   Note that $\ker h \subset \ker f$ and so $\dim\ker h \le \dim\ker f$. Similarly, suppose $w \in \operatorname{im} h$. Then $w = h(v) = f(v)$ for some $v \in \ker(g \circ f)$. Therefore $g(w) = (g \circ f)(v) = 0$. Thus $w \in \ker g$. Therefore, $\dim\operatorname{im} h \le \dim\ker g$. Combining these inequalities, we have

   $$\dim\ker(g \circ f) \le \dim\ker f + \dim\ker g.$$

   This completes the proof.

2. To prove this claim, I will prove that $\dim\operatorname{im}(g \circ f) \le \dim\operatorname{im} g$ and $\dim\operatorname{im}(g \circ f) \le \dim\operatorname{im} f$. For the first inequality, note that $\operatorname{im}(g \circ f) \subset \operatorname{im} g$ and thus the inequality follows.

   For the second inequality, note that $\operatorname{im} f$ is a subspace of W. Thus we can consider the linear map $h := g|_{\operatorname{im} f}$. Applying the fundamental theorem for linear maps,

   $$\dim\operatorname{im} f = \dim\operatorname{im} h + \dim\ker h$$

   I claim that $\operatorname{im} h = \operatorname{im}(g\circ f)$. To prove this, I prove both inclusions. Suppose $x \in \operatorname{im} h$. Then $x = g(w)$ for some $w \in \operatorname{im} f$. Then we see that $w = f(v)$ for some $v \in V$. Thus $x = g(f(v))$ and therefore $x \in \operatorname{im}(g \circ f)$. For the other inclusion, suppose that $x \in \operatorname{im}(g \circ f)$. Then $x = g(f(v))$ for some $v \in V$. Then $x = g(w)$ where $w = f(v) \in \operatorname{im} f$. This proves the other inclusion. Thus we have proven that $\operatorname{im} h = \operatorname{im}(g \circ f)$. Now we have the following:

   $$\dim\operatorname{im} f - \dim\ker h = \dim\operatorname{im}(g \circ f)$$

   Since $\dim\ker h \ge 0$ it follows that

   $$\dim\operatorname{im} f \ge \dim\operatorname{im}(g \circ f)$$

   This second inequality completes the proof. (TA)

Problem 1.3. Let $f, g\colon V \to W$ be two linear maps.

1. Suppose that W is finite dimensional. Show that $\ker f \subset \ker g$ if and only if there exists a third linear map $h\colon W \to W$ satisfying $g = h \circ f$.
2. Suppose that V is finite dimensional. Show that $\operatorname{im} f \subset \operatorname{im} g$ if and only if there exists a third linear map $h\colon V \to V$ satisfying $f = g \circ h$.

Solution.

Lemma 1. Suppose X is a K vector-space such that $\dim X$ is finite. Suppose that $U \le X$ is a subspace. There exists a linear map $\varphi\colon X \to U$ such that $\varphi|_U = \mathrm{id}_U$.

Note. One might call φ the projection map onto the subspace U. You can clearly see that $\varphi^2 = \varphi$. We will deal a lot with these projection maps in the future, so it may be worthwhile to study their construction.

Proof. Take a basis $u_1, \ldots u_n$ for U. Extend this to $S = \{u_1, \ldots, u_n, x_1, \ldots x_m\}$, a basis for X. Then let $\varphi \colon X \to U$ be defined as follows:

$$\varphi(a_1u_1 + \cdots + a_nu_n + a_{n+1}x_1 + \cdots + a_{n+m}x_m) = a_1u_1 + \cdots + a_nu_n.$$

We see that this defines a single, well-defined, value for φ on all $x \in X$ because S is a basis. To prove linearity, let

$$\begin{aligned} x &= a_1u_1 + \cdots + a_nu_n + a_{n+1}x_1 + \cdots + a_{n+m}x_m \\ cy &= ca'_1u_1 + \cdots + ca'_nu_n + ca'_{n+1}x_1 + \cdots + ca'_{n+m}x_m \end{aligned}$$

for arbitrary $x, y \in X$ and $c \in K$. Then consider

$$\begin{aligned} &\varphi(a_1u_1 + \cdots + a_nu_n + a_{n+1}x_1 + \cdots + a_{n+m}x_m \\ &+ca'_1u_1 + \cdots + ca'_nu_n + ca'_{n+1}x_1 + \cdots + ca'_{n+m}x_m) \\ &= (a_1 + ca'_1)u_1 + \cdots + (a_n + ca'_n)u_n = \varphi(x) + c\varphi(y). \end{aligned}$$

This proves that φ is a linear map. By inspection, $\varphi|_U = \mathrm{id}_U$.

1. First, the forwards direction. Suppose there exists a linear map $h \colon W \to W$ satisfying $g = h \circ f$. Suppose that $v \in \ker f$. Then $g(v) = h(f(v)) = h(0) = 0$. Therefore $v \in \ker g$. It follows that $\ker f \subset \ker g$.

   For the reverse direction, I present two proofs. One is very "algebra-y" and the other is very "linear-algebra-y". I think the first is more elegant, but the second has better intutions for you at this point. Take whichever you prefer.

   Proof. Suppose that $\ker f \subset \ker g$. Let $\varphi \colon W \to \operatorname{im} f$ be a linear map whose existence is given by Lemma 1. Then let $\pi \colon V \to V/\ker f$ be the canonical quotient map. Then let $\tilde{f} = \pi \circ f$. We note that $\tilde{f} \colon V/\ker f \to W$ is an isomorphism. (If this construction is confusing to you, refer to Axler 3.91). Then the following diagram commutes:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \big\downarrow \pi & & \big\downarrow \varphi \\ V/\ker f & \xleftarrow{\tilde{f}^{-1}} & \operatorname{im} f \end{array}$$

   Namely, I claim $\tilde{f}^{-1} \circ \varphi \circ f = \pi$. This is in fact a fairly trivial claim as $\tilde{f} = \pi \circ f$ and $\varphi \circ f = f$ because $f \colon V \to W$ embeds V in W as $\operatorname{im} f$. Let $\tilde{g} \colon V/\ker f \to W$ be the map defined by $\tilde{g}(v + \ker f) = g(v + \ker f) = g(v)$. Such a map is well defined because $\ker f \subset \ker g$, as represented in the last equality. This is essential so it bears repeating. If $x, y \in v + \ker f$ then $x - y \in \ker f \Rightarrow x - y \in \ker g$. Then $g(x) - g(y) = g(x - y) = 0$. Thus $g(x) = g(y)$. So the map is well defined on the affine subsets.

   Then we see that $\tilde{g}((v + cv') + \ker f) = g((v + cv') + \ker f) = g(v + cv') = g(v) + cg(v')$. Thus $\tilde{g}$ is a linear map. We see by inspection that $g = \tilde{g} \circ \pi$.

   Then let $h = \tilde{g} \circ \tilde{f}^{-1} \circ \varphi$. As a composition of linear maps, h is a linear map. It only remains to be shown that $h \circ f = g$, but this was the point of the commutative diagram!

$$h \circ f = \tilde{g} \circ \tilde{f}^{-1} \circ \varphi \circ f = \tilde{g} \circ \pi = g$$

   Proof. Now for the second proof. This time with bases and more explicit constructions. We are given that W is finite dimensional. Construct a basis for $\operatorname{im} f$. Such a basis has the form $S =$

$\{f(v_1), f(v_2), \ldots, f(v_n)\}$ for $v_i \in V$. (As a basis for $\operatorname{im} f$, each basis element is in the image). Extend S to a basis for W as follows $S' = \{f(v_1), \ldots, f(v_n), w_1, \ldots, w_m\}$. Then let $\varphi\colon W \to V$ be defined as

$$\varphi(a_1 f(v_1) + \cdots + a_n f(v_n) + a_{n+1} w_1 + \cdots + a_{n+m} w_m) = a_1 v_1 + \cdots + a_n v_n$$

To show linearity of this map, suppose we have w and w' arbitrary as follows:

$$\begin{aligned} w &= a_1 f(v_1) + \cdots + a_n f(v_n) + a_{n+1} w_1 + \cdots + a_{n+m} w_m \\ w' &= b_1 f(v_1) + \cdots + b_n f(v_n) + b_{n+1} w_1 + \cdots + b_{n+m} w_m \end{aligned}$$

Then we have, for $c \in K$ arbitrary

$$\begin{aligned} \varphi(w + cw') =& \varphi((a_1 + cb_1) f(v_1) + \cdots + (a_n + cb_n) f(v_n) \\ &+ (a_{n+1} + cb_{n+1}) w_1 + \cdots + (a_{n+m} + cb_{n+m}) w_m) \\ =& (a_1 + cb_1) v_1 + \cdots + (a_n + cb_n) v_n = \varphi(w) + c\varphi(w') \end{aligned}$$

Thus φ is a linear map. Then let $h\colon W \to W$ be defined by $h = g \circ \varphi$. As a composition of linear maps, h is a linear map. Then I claim $h \circ f = g$. Consider $\varphi \circ f$. I claim that $\varphi \circ f(v) \in v + \ker f$. This must be the case as $f \circ \varphi(w) = w$ for $w \in \operatorname{im} f$ by inspection. Then, since $\ker f \subset \ker g$ it follows that

$$g \circ \varphi \circ f(v) = g(v')$$

for $v' \in v + \ker f$. Thus

$$\begin{aligned} g(v') &= g(v) \\ \Rightarrow g \circ \varphi \circ f &= g \end{aligned}$$

This completes the latter proof.

2. For the second part, we proceed very similarly, and I will only present one of the proof techniques (but note that a similar style proof could be constructed with another commutative diagram).

   For the forwards direction, suppose that there exists a linear map $h\colon V \to V$ satisfying $f = g \circ h$. Then for $w \in \operatorname{im} f$ there exists a $v \in V$ such that $g \circ h(v) = f(v) = w$. But $h(v)$ is a vector like any other, and is the preimage of w under g. Thus the claim is proven.

   For the reverse direction, suppose that $\operatorname{im} f \subset \operatorname{im} g$. Let $w_1, \ldots, w_n$ be a basis for $\operatorname{im} f$ (finite dimensional because V is finite dimensional). Then extend this to a basis for V by adding $w_{n+1}, \ldots, w_{n+m}$. We see that $w_i = f(v_i)$ for $1 \leq i \leq n$ and $w_i = g(v_i')$ for $1 \leq i \leq n$. Note that here we have applied the crucial assumption that $\operatorname{im} f \subset \operatorname{im} g$.

   The vectors $v_1, \ldots, v_n$ span a subspace $U \leq V$. Then construct $\varphi\colon V \to U \hookrightarrow V$ (the hooked arrow just means "naturally included in"), the projection map from Lemma **??**. Then define the map $h\colon V \to V$ by its action on the basis of V:

$$h(v_i) = \begin{cases} v_i' & 1 \leq i \leq n \\ 0 & \text{else} \end{cases}$$

   This map is linear by analogy to the previous part. I claim that $f = g \circ h$. Suppose $v \in V$ arbitrary such that $f(v) = a_1 f(v_1) + \cdots + a_n f(v_n)$. Then we consider $g \circ h(v)$:

$$\begin{aligned} g \circ h(v) = g(a_1 v_1' + \cdots + a_n v_n') &= a_1 w_1 + \cdots + a_n w_n \\ &= a_1 f(v_1) + \cdots + a_n f(v_n) = f(v). \end{aligned}$$

   This completes the proof. (TA)

2 For submission to Davis Lazowski

Problem 2.1. Actually check the distributivity of linear maps. Suppose you are given linear maps

$$f, g\colon V \to W$$

and a third linear map

$$h\colon W \to X.$$

Demonstrate $h(f+g) = (hf) + (hg)$.

Solution. It's enough to show that $\forall x : h(f+g)(x) = [(hf)(x) + hg(x)]$.

But $(f+g)(x)$ is just a vector, and $(f+g)(x) = f(x) + g(x)$, which is also just a vector. So by linearity:

$$\begin{aligned} &h((f+g)(x)) \\ &= h(f(x) + g(x)) \\ &= h(f(x)) + h(g(x)) \\ &= [hf](x) + [hg](x) \end{aligned}$$

(DL)

Problem 2.2. Let $f\colon V \to W$ be a linear map and suppose that V is finite dimensional. Show that there exists a subspace $U \leq V$ with $U \oplus \ker f = V$ and $f(U) = \operatorname{im} V$.

Solution. $\ker f$ is a subspace. So by the algorithm for finding the complement of a subspace discussed in class, there exists $U : \ker f \oplus U = V$. So it's enough to show that $f(U) = \text{imV}$.

Let $i \in \text{imV}$. Then there exists $v \in V : f(v) = i$. By the direct sum relation, there are $k \in \ker f, u \in U$ such that $v = k + u$. Also, $f(k) = 0$ by definition of the kernel. So

$$f(u) = f(v-k) = f(v) - f(k) = f(v) - 0 = f(v)$$

So $f(U) \supset \text{im(V)}$. But also $f(U) \subset \text{im(V)}$ by definition of the image. Therefore $f(U) = \text{im(V)}$ (DL)

Problem 2.3. Suppose that $f\colon V \to W$ is a linear function of vector spaces V and W over a scalar field K, and let $(w_1, \dots, w_n)$ be a basis for $\operatorname{im} f$. Show that there exist $\varphi_1, \dots, \varphi_n \in \mathcal{L}(V, K)$ such that

$$f(v) = \varphi_1(v) \cdot w_1 + \cdots + \varphi_n(v) \cdot w_n.$$

Solution. Define $\pi_j(w)$ as the coefficent of w_j when w is written in terms of the basis $w_1 ... w_n$, so that $\pi_j : W \to K$. By linear independence of the basis, this definition is well-defined.

Claim. π_j is linear.

Proof. Let $w = \sum_{i=1}^n \alpha_i w_i$, $u = \sum_{i=1}^n \beta_i w_i$.

Then $\pi_j(w) = \alpha_j, \pi_j(u) = \beta_j$. Also, $w + \lambda u = \sum_{i=1}^n (\alpha_i + \lambda\beta_i) w_i$.

Therefore,

$$\pi_j(w + \lambda u) = \alpha_j + \lambda\beta_j = \pi_j(w) + \lambda\pi_j(u)$$

So $\pi_j \in \mathcal{L}(W, K)$. Therefore, $\pi_j \circ f \in \mathcal{L}(V, K)$.

Also, by definition of the π_j,

$$f(v) = \sum_{j=1}^n [\pi_j \circ f](v) w_j$$

(DL)

Problem 2.4. Prove that the intersection of any finite collection of affine subsets of V is either the empty set or yet another affine subset.

Solution. Recall that an affine subset of V is $v+U$, with $v \in V$ and $U \subset V$ a subspace. If W is a complement subspace of W, $\forall v+U \exists w \in W : w+U = v+U$.

This is because $v = \tilde{w} + \tilde{u}$, with $\tilde{w} \in W, \tilde{u} \in U$. Clearly $\tilde{u} \in U$, so can be dropped, so $\tilde{w} + U = v + U$.

Recall also that the solution space of $f(v) = w$ is always an affine subset of V or empty. Let's show the reverse implication, that every affine subset of V is also the solution space of an equation.

Let U a subspace, $u_1...u_m$ a basis of this subspace. Extend this to a basis $u_1...u_m, r_1...r_{n-m}$ of V. Write $x \in V$ as $x = \sum_{j=1}^{m} \alpha_j u_j + \sum_{i=1}^{n-m} \beta_i r_i$.

Define $\pi_R \in \mathcal{L}(V,V)$ as $\pi_R(x) = \sum_{j=1}^{n-m} \beta_i r_i$. Then $\ker \pi_R = U$. The equation $\pi_R(u) = w$, with $w \in W$, then clearly has solution space $w + U = v + U$.

Let $v_i + U_i$ be some collection of affine subspaces. Denote their associated equations as $\pi_{R_i}(u) = w_i$.

Observe that $(v_i + U_i) \cap (v_{i'} + U_{i'})$ is just the set of u that simultaneously solve the associated equations:

$$\pi_{R_i}(u) = w_i$$
$$\pi_{R_{i'}}(u) = w_{i'}$$

And a simlar relation holds true for the intersection of any collection of affine subspaces.

Let $(v_i + U_i), i \in I$ a collection of affine subspaces. Define $\pi \in \mathcal{L}(V, V^I)$ by $\pi(u) = (\pi_{R_1}(u), \pi_{R_2}(u)....)$

Then the equation
$$\pi(u) = (w_1, w_2....)$$
Has u as a solution if any only if it solves all the equations $\pi_{R_j}(u) = w_j$. So $u \in \bigcap_{i \in I}(v_i + U_i)$. But if $u \in \bigcap_{i \in I}(v_i + U_i)$, then u solves each equation $\pi_{R_j}(u) = w_j$, so solves $\pi(u) = (w_1, w_2...)$. But the solution space of $\pi(u) = (w_1, w_2...)$ is an affine subset or empty, therefore done. (DL)

3 For submission to Handong Park

Your task is to reinvent Gaussian elimination, since this topic is not covered in the book.

Problem 3.1. Gaussian elimination involves three *elementary row operations* performed on the entries of a matrix:

- Swap the j^{th} and k^{th} rows.
- For indices $j \neq k$ and some scalar c, take the k^{th} row, scale all its entries by c, and add the result to the j^{th} row.
- For an index j and a *nonzero* scalar c, scale the j^{th} row by c.

Our first goal is to understand some features of these row operations.

1. Each of these operations can be encoded by matrices $S(j,k)$, $A(j,k,c)$, and $M(j,c)$ so that $S(j,k) \cdot X$, $A(j,k,c) \cdot X$, and $M(j,c) \cdot X$ are the respective results of the operations applied to a matrix X. These matrices $S(j,k)$, $A(j,k,c)$, and $M(j,c)$ are called *elementary matrices*. Find descriptions of the elementary matrices. (This could mean formulas, or English descriptions, or...)
2. Check that the elementary matrices are all invertible.
3. You can think of $S(j,k)$, $A(j,k,c)$, and $M(j,c)$ as describing a change of basis. If X encodes the behavior of a linear map $f\colon V \to W$ on a basis $(w_1, \ldots, w_n)$ of W, then in what bases do $S(j,k) \cdot X$, $A(j,k,c) \cdot X$, and $M(j,c) \cdot X$ encode f?

Solution.

1. Here are three formulaic descriptions:

$$S(j,k)_{yx} = \begin{cases} 1 & y = x \text{ and } y \neq j, k, \\ 1 & y = j \text{ and } x = k, \\ 1 & y = k \text{ and } x = j, \\ 0 & \text{otherwise.} \end{cases}$$

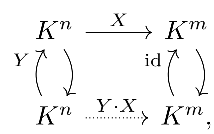
$$A(j,k,c)_{yx} = \begin{cases} 1 & y = x, \\ c & y = j \text{ and } x = k, \\ 0 & \text{otherwise.} \end{cases}$$

$$M(j,c)_{yx} = \begin{cases} 1 & y = x \text{ and } y \neq j, \\ c & y = x = j, \\ 0 & \text{otherwise.} \end{cases}$$

   (Another way to think of this intuitively is that you just perform the operation you want on the identity matrix of appropriate size - i.e., to swap row j and k, swap these rows in the identity matrix.)

2. All three matrices are invertible because the three row operations they encode are invertible. The operation $S(j,k)$ is its own inverse, $M(j,c)$ is inverted by $M(j,c^{-1})$, and $A(j,k,c)$ is inverted by $A(j,k,-c)$.

3. We consider the product $Y \cdot X$ in terms of the diagram



   and to ask what basis Y encodes in the case that Y is an elementary matrix.

   - For $Y = S(j,k)$ (assuming, without any loss of generality, that $j < k$), the standard basis is traded for
   $$(v_1, \ldots, v_{j-1}, v_k, v_{j+1}, \ldots, v_{k-1}, v_j, v_{k+1}, \ldots, v_n).$$
   - For $Y = A(j,k,c)$, v_j is replaced by $v_j + cv_k$.
   - For $Y = M(j,c)$, v_j is replaced by cv_j. (ECP)

Problem 3.2. A matrix X is said to be *upper-triangular* when the entries satisfy $X_{ij} = 0$ whenever $i > j$. (When written as a block of numbers, all the entries below the main diagonal are zero.) Describe an algorithm which modifies a matrix to be upper-triangular using the elementary row operations. (Be sure to argue that your algorithm actually succeeds at this goal.) (Hint: work one column at a time.)

Solution. We do, indeed, work one column at a time. In the j^{th} column, consider all rows at and below the j^{th} position. If all their entries in the j^{th} column are zero, this column passes the upper-triangularity test, and we proceed to the next column. If some row in this region contains a nonzero entry in the j^{th} column, use a swap operation to place it into the j^{th} row. (Note that this does not disturb any upper-triangularity test in columns before the j^{th} column, because the swap trades 0 entries for 0 entries.) Then, use the rescaling operation to make the (new) j^{th} row have a 1 in the j^{th} column. Finally, consider every other row at position $i \neq j$: scale up the j^{th} row by $-X_{ij}$, and add it to the i^{th} row so that the $(i,j)^{\text{th}}$ position new reads as zero.[1] Now the j^{th} column passes the upper-triangularity test, so we proceed to the next column. (ECP)

[1] If you're just aiming for upper-triangularity, you might only zero out the rows below j^{th}. That's fine, but it makes the next problem harder.

Problem 3.3. 1. Problem 3.2 can be used to calculate the inverses of matrices. Suppose that your algorithm row-reduces a matrix X to the identity matrix, i.e.,

$$E_n \cdot \cdots \cdot E_2 \cdot E_1 \cdot X = I$$

for some sequence of elementary matrices (E_j). Supposing further that X is invertible, i.e.,

$$X \cdot X^{-1} = I$$

for other some matrix X^{-1}, solve for X^{-1} in terms of the elementary matrices (E_j).

2. Use your algorithm to calculate the matrix inverse of

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

(If necessary, modify your algorithm above to handle this case — that is, make sure it gives you the identity matrix rather than merely an upper triangular matrix.)

Solution. 1. This equation is easy enough to solve:

$$\begin{aligned} X \cdot X^{-1} &= I \\ (E_n \cdot \cdots \cdot E_2 \cdot E_1) \cdot X \cdot X^{-1} &= (E_n \cdot \cdots \cdot E_2 \cdot E_1) \cdot I \\ X^{-1} &= (E_n \cdot \cdots \cdot E_2 \cdot E_1) \cdot I. \end{aligned}$$

Our conclusion is that the inverse of X is given by the same sequence of elementary row operations applied to the identity matrix.

2. We run the row-reduction algorithm on the left and apply the same row operations to the identity matrix on the right:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \xrightarrow{A(2,1,-1)} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{A(2,1,-1)} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \xrightarrow{A(1,2,-1)} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \xrightarrow{A(1,2,-1)} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

We conclude that $\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ is inverse to $\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$. (ECP)

Problem 3.4. 1. Describe the effect of *right-multiplying* some matrix X by an elementary matrix.

2. Using both the original elementary row operations and the operations uncovered in the previous part, devise a variation on your answer to Problem 3.2 that gives an algorithm that rewrites *any* matrix as a *diagonal* matrix. (Again, be sure to argue that your algorithm actually succeeds at this goal.)

3. Conclude that for any linear map $f\colon V \to W$ between finite dimensional vector spaces, we can find bases of V and W such that the resulting matrix X expressing f is *diagonal*.

4. Conclude rank-nullity from this form X for f:

$$\dim V = \dim \operatorname{im} f + \dim \ker f.$$

Solution. 1. Right-multiplication by an elementary matrix encodes analogous elementary *column* operations. They modify the basis on the target similarly to how the row operations modify the basis on the source.

2. We augment the algorithm described above as follows: before ever moving from the j^{th} column to the $(j+1)^{\text{st}}$ column, we examine the j^{th} row. If it is entirely empty, we proceed to the next column as before. If the (j, j) entry is nonzero, we can use the scale-and-add column operations to zero out the other entries in the j^{th} row, then proceed to the next column as before. If the (j, j) entry is zero and there is some other nonzero entry in the j^{th} row, we swap the j^{th} column for that column and re-process the new j^{th} column using the elementary row operation algorithm from before. In the end, this gives a diagonal matrix with only 0s and 1s on the main diagonal.
3. Beginning with *any* bases for V and W, we get a matrix presentation of f. Applying the row-and-column reduction algorithm described above, we can modify both bases for V and W so that the matrix presentation for f becomes diagonal (with only 0s and 1s on the main diagonal).
4. The kernel and image of such a matrix are extremely easy to describe: the image of the matrix consists of the span of those w_j in which the j^{th} diagonal entry is zero, and the kernel of the matrix consits of the span of whose v_j in which the j^{th} diagonal entry is zero. Since there are $\dim V$ columns in all, this gives
$$\dim V = \dim \operatorname{im} f + \dim \ker f. \tag{ECP}$$

4 For submission to Rohil Prasad

Problem 4.1. Suppose that $x, y \in V$ are vectors in a vector space V and $U, W \leq V$ are subspaces of V, altogether satisfying the relation
$$x + U = y + W$$
Show that $U = W$.

Solution. Since $0 \in W$, $y \in y + W$. Therefore, there exists $u \in U$ satisfying
$$x + u = y$$
so $u = y - x$ and $y - x \in U$.

For any vector $w' \in W$, there exists $u' \in U$ such that $x + u' = y + w'$ which implies $w' = x - y + u'$. Since $x - y, u' \in U$, it follows that $w' \in U$ and therefore $W \subset U$.

In the other direction, we have since $0 \in U$, $x \in x + U$. Therefore, there exists $w \in W$ satisfying
$$y + w = x$$
so $w = x - y$ and $x - y \in W$.

For any vector $u' \in U$, there exists $w' \in W$ such that $x + u' = y + w'$ which implies $u' = y - x + w'$. Since $y - x, w' \in W$, it follows that $u' \in W$ and therefore $U \subset W$.

Since $W \subset U$ and $U \subset W$, we must have $U = W$. (RP)

Problem 4.2. Prove that a nonempty subset $A \subseteq V$ of a vector space V is an affine subset if and only if for all $v, w \in A$ and all $\lambda \in K$ it is also the case that
$$\lambda \cdot v + (1 - \lambda) \cdot w \in A$$
(Side remark: for $K = \mathbb{R}$ and $0 \leq \lambda \leq 1$, this property is called *convexity*.)

Solution. First we will show that if A is affine, then the property described in the problem statement holds. Assume $A = v + U$ for a vector $v \in V$ and a subspace $U \subset V$.

Then any two elements in A can be expressed in the form $v + u_1, v + u_2$ for $u_1, u_2 \in U$. Expanding out the sum $\lambda(v + u_1) + (1 - \lambda)(v + u_2)$, we get:
$$\begin{aligned}
\lambda \cdot (v + u_1) + (1 - \lambda) \cdot (v + u_2) &= \lambda \cdot v + \lambda \cdot u_1 + (1 - \lambda) \cdot v + (1 - \lambda) \cdot u_2 \\
&= (\lambda + 1 - \lambda) \cdot v + \lambda \cdot u_1 + (1 - \lambda) \cdot u_2) \\
&= v + \lambda \cdot u_1 + (1 - \lambda) \cdot u_2
\end{aligned}$$

Since U is a subspace, $\lambda \cdot u_1 + (1-\lambda) \cdot u_2 \in U$ and therefore $\lambda \cdot (v+u_1) + (1-\lambda) \cdot (v+u_2) \in A$ as desired.

Now assume that A satisfies the property. Fix some $a \in A$. Showing that A is affine is equivalent to showing the subset $U = \{x - a | x \in A\}$ is a subspace. To do so, we will show that $0 \in U$ and that U is closed under scalar multiplication.

Since $a \in A$, we have $0 \in U$.

Given some scalar λ and an element $x - a \in U$, we have

$$\begin{aligned} \lambda \cdot (x-a) &= \lambda x - \lambda a \\ &= \lambda x + (1-\lambda)a - a \\ &= (\lambda x + (1-\lambda)a - a \end{aligned}$$

By the convexity property, $\lambda x + (1-\lambda)a \in A$, and therefore $\lambda \cdot (x-a) \in U$ and U is closed under scalar multiplication.

Given two elements $x_1 - a, x_2 - a \in U$, we have their sum is $x_1 + x_2 - 2a = 2 \cdot (\frac{1}{2}x_1 + \frac{1}{2}x_2 - a)$. By the convexity property, $\frac{1}{2}x_1 + \frac{1}{2}x_2 \in A$, so $\frac{1}{2}x_1 + \frac{1}{2}x_2 - a \in U$. Since U is closed under scalar multiplication, it follows that $x_1 + x_2 - 2a \in U$ and U is closed under addition.

Since U is a subspace, it follows that A is an affine subset of the form $a + U$. (RP)

Problem 4.3. Suppose that $U \leq V$ is a subspace such that the quotient space V/U is finite dimensional. Show that V is isomorphic to $U \times (V/U)$.

Solution. We will construct this isomorphism directly. Note that a dimension argument does not work here since V, U are not necessarily finite dimensional.

Assume that V/U has dimension n. Pick $v_1, \dots, v_n \in V$ such that the affine subsets $v_1 + U, \dots, v_n + U \in V/U$ form a basis of V/U.

For any affine subset $v+U \in V/U$, there exists constants $k_1, \dots, k_n \in K$ such that $\sum_{i=1}^n k_i v_i + U = v + U$.

Thus, we can define φ as the map sending a pair $(u, v+U) \in U \times V/U$ to the vector $u + \sum_{i=1}^n k_i v_i$. It is clear by definition that φ is linear, so we must show it is injective and surjective.

If there exists $(u, v+U)$ such that $\varphi(u, v+U) = 0$, then we require $u + \sum_{i=1}^n k_i v_i = 0$, which implies $\sum_{i=1}^n k_i v_i = -u \in U$. However, since the $v_i + U$ form a basis of V/U, $\sum_{i=1}^n k_i v_i \in U$ implies that $\sum_{i=1}^n k_i (v_i + U) = U$. Since U the zero element of V/U, it follows that all the k_i must equal 0. Therefore, we have that $u = -(\sum k_i v_i) = 0$ and $v + U = U$, so only the zero element maps to 0 and φ is injective.

For any $v \in V$, there exists constants k_i such that $\sum_{i=1}^n k_i v_i + U = v + U$. It follows that $\sum_{i=1}^n k_i v_i - v \in U$. Denoting this vector by u, we have that $\varphi(u, v+U) = v$ and therefore φ is surjective. (RP)