Algebra 1

Homework 2 — Thursday, 20/08/2026

Lilian — Prof. Upendra

Homework

Submit: problems 2(a), 3, 4.

Problem 2(a)

We have seen that RREF(A) allows you to write the solution space of $A\mathbf{x} = \mathbf{0}$ as the span of a "magic" set of vectors. Do this for the following system.

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 2 & 4 & 0 & 4 & 6 \\ 1 & 2 & 2 & 5 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Proof. Call the coefficient matrix A. Row operations do not change the solution set of $A\mathbf{x} = \mathbf{0}$, so we row reduce A (the zero column on the right never changes, so I dropped it).

$$\begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 2 & 4 & 0 & 4 & 6 \\ 1 & 2 & 2 & 5 & 1 \end{pmatrix} \xrightarrow{R_2-2R_1,\,R_3-R_1} \begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 0 & -2 & -2 & 2 \\ 0 & 0 & 1 & 2 & -1 \end{pmatrix} \xrightarrow{-\frac{1}{2}R_2} \begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 2 & -1 \end{pmatrix}$$

$$\xrightarrow{R_3-R_2} \begin{pmatrix} 1 & 2 & 1 & 3 & 2 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \xrightarrow{R_2-R_3,\,R_1-3R_3} \begin{pmatrix} 1 & 2 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \xrightarrow{R_1-R_2} \begin{pmatrix} 1 & 2 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

The last matrix is in RREF, with pivots in columns $1, 3, 4$. So x_2 and x_5 are free, and the system is equivalent to

$$x_1 = -2x_2 - 3x_5, \quad x_3 = x_5, \quad x_4 = 0.$$

Hence every solution has the form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -2x_2 - 3x_5 \\ x_2 \\ x_5 \\ 0 \\ x_5 \end{pmatrix} = x_2 \underbrace{\begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{u}} + x_5 \underbrace{\begin{pmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}}_{\mathbf{w}}, \quad x_2, x_5 \in \mathbb{R}.$$

Conversely, for any choice of x_2, x_5 the vector above satisfies the three RREF equations, hence is a solution. So

$$\{\mathbf{x} \in \mathbb{R}^5 \mid A\mathbf{x} = \mathbf{0}\} = \text{span}\{\mathbf{u}, \mathbf{w}\}.$$

The magic vectors are exactly the ones you get by setting one free variable to 1 and the others to 0. In particular $\mathbf{u}$ and $\mathbf{w}$ are independent: $\mathbf{u}$ has a 1 in the x_2 slot where $\mathbf{w}$ has 0, and vice versa for x_5.

Problem 3

Recall that "the general solution of a consistent system is a translate of a subspace", i.e. of the form $W + \mathbf{u}$ where W is a subspace (of what?) and $\mathbf{u}$ a vector. Demonstrate this explicitly for the following system.

$$\begin{pmatrix} 1 & -1 & 2 & 3 \\ 2 & -2 & 5 & 4 \\ -1 & 1 & -1 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \\ -4 \end{pmatrix}$$

Proof. Call the coefficient matrix A and the right hand side $\mathbf{b}$. Row reducing $(A \mid \mathbf{b})$:

$$\left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 5 \\ 2 & -2 & 5 & 4 & 11 \\ -1 & 1 & -1 & -5 & -4 \end{array}\right) \xrightarrow{R_2-2R_1, R_3+R_1} \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 5 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 & 1 \end{array}\right) \xrightarrow{R_3-R_2, R_1-2R_2} \left(\begin{array}{cccc|c} 1 & -1 & 0 & 7 & 3 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

There is no pivot in the last column, so by Lemma 1 of HW 1 the system is consistent. Pivots are in columns $1, 3$, so x_2, x_4 are free and

$$x_1 = 3 + x_2 - 7x_4, \quad x_3 = 1 + 2x_4.$$

So every solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \underbrace{\begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{u}} + x_2 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -7 \\ 0 \\ 2 \\ 1 \end{pmatrix}, \quad x_2, x_4 \in \mathbb{R},$$

and every such vector is a solution. Setting $W = \operatorname{span}\left\{(1,1,0,0)^T, (-7,0,2,1)^T\right\}$, the solution set is exactly $W + \mathbf{u}$.

Problem 4

For an $r \times c$ matrix A, recall/check that the map $f(\mathbf{x}) = A\mathbf{x}$ from $\mathbb{R}^c$ to $\mathbb{R}^r$ is linear, which means it preserves vector addition and scalar multiplication. Show that every linear map L from $\mathbb{R}^c$ to $\mathbb{R}^r$ comes in this fashion, i.e. for a given linear map L, there is a unique matrix A such that $L(\mathbf{x}) = A\mathbf{x}$.

Proof. f **is linear.** The i-th entry of $A\mathbf{x}$ is $\sum_j a_{ij}x_j$. For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^c$ and $\lambda \in \mathbb{R}$,

$$(A(\mathbf{x}+\mathbf{y}))_i = \sum_j a_{ij}\big(x_j + y_j\big) = \sum_j a_{ij}x_j + \sum_j a_{ij}y_j = (A\mathbf{x})_i + (A\mathbf{y})_i,$$

$$(A(\lambda\mathbf{x}))_i = \sum_j a_{ij}\big(\lambda x_j\big) = \lambda \sum_j a_{ij}x_j = \lambda(A\mathbf{x})_i.$$

Since this holds for every i, $f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ and $f(\lambda\mathbf{x}) = \lambda f(\mathbf{x})$.

Existence. Let $\mathbf{e}_1, \ldots, \mathbf{e}_c$ be the standard basis of $\mathbb{R}^c$, and let A be the $r \times c$ matrix whose j-th column is $L\big(\mathbf{e}_j\big)$. Any $\mathbf{x} \in \mathbb{R}^c$ is $\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_c\mathbf{e}_c$, so by linearity of L

$$L(\mathbf{x}) = x_1L(\mathbf{e}_1) + \cdots + x_cL(\mathbf{e}_c).$$

On the other hand, $A\mathbf{x}$ is exactly the linear combination of the columns of A with coefficients $x_1, \ldots, x_c$ (Tutorial 2, problem 1), which is the same vector. So $L(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x}$.

Uniqueness. Suppose $A\mathbf{x} = B\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^c$. Plugging in $\mathbf{x} = \mathbf{e}_j$ gives $A\mathbf{e}_j = B\mathbf{e}_j$. But $A\mathbf{e}_j$ is the j-th column of A, and likewise for B. So A and B have the same columns, i.e. $A = B$.

Rest of Tutorial 2

Not part of the homework.

All problems except problem 7 are of major conceptual importance. Recall the definitions of: linear combination, span, subspace, linear map.

Problem 1

Let A be an $r \times c$ matrix and $\mathbf{x}$ a vector in $\mathbb{R}^c$. Show that $A\mathbf{x}$ is a linear combination of certain vectors obtained from entries of A.

Proof.

By laws of vector-matrix multiplication,

$$A\mathbf{x} = x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{1r} \end{pmatrix} + \cdots + x_c \begin{pmatrix} a_{c1} \\ \vdots \\ a_{cr} \end{pmatrix}$$

Problem 2(b)

Recall/check that the following are subspaces of $\mathbb{R}^c$: (i) for a subset S of $\mathbb{R}^c$, span(S) = the set of all linear combinations of vectors from S; (ii) the set of solutions of $A\mathbf{x} = \mathbf{0}$ for an $r \times c$ matrix A.

(b) (Optional for now.) Show how to write a subspace of the first type as a subspace of the second type by finding a magic matrix. Setup: let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ be a subset of $\mathbb{R}^n$. Show how to explicitly find a matrix A such that

$$\text{span}(S) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

Proof.

(b) We show how to find a $n \times n$ matrix A, whose null space is span(S). Let $\mathbf{w}_1, \ldots, \mathbf{w}_m$ be a basis of span(S) where $m \leq k$. We can extend it to form a basis of $\mathbb{R}^n$. Let this be $\mathbf{w}_1, \ldots, \mathbf{w}_m, \mathbf{u}_1, \ldots, \mathbf{u}_\ell$, where $\ell + m = n$. Then, we can determine A by just its action on the basis. In particular, $A\mathbf{w}_i = \mathbf{0}$ and $A\mathbf{u}_i = \mathbf{e}_i$ where $\mathbf{e}_i$ are the standard basis vectors.

Note that span(S) $\subset$ null(A) as any vector in the span is a linear combination of the basis, $\mathbf{v} = c_1\mathbf{w}_1 + \cdots + \mathbf{w}_m$, and thus: $A\mathbf{v} = A(c_1\mathbf{w}_1 + \cdots + c_m\mathbf{w}_m) = c_1A_1\mathbf{w}_1 + \cdots + c_mA\mathbf{w}_m = \mathbf{0}$, so $\mathbf{v} \in$ null(A).

Also, null(A) $\subset$ span(S), as $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ form a basis of null(A). This is because, since null(A) $\subset \mathbb{R}^n$, any $\mathbf{v} \in$ null(A) can be written down as a linear combination of our basis of $\mathbb{R}^n$. In particular, assume $\mathbf{v} = \sum_i a_i\mathbf{w}_i + \sum b_i\mathbf{u}_i$. Then, since $A\mathbf{v} = \mathbf{0}$,

$$\mathbf{0} = \sum_i b_i A\mathbf{u}_i = \sum_i b_i\mathbf{e}_i$$

which cannot be $\mathbf{0}$ unless all the b_i are 0. Thus $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ does indeed form a basis. Therefore, we are done.

Problem 5

This exercise has many parts and it will take some time to do all of them. Do what you can at the moment!

When is $f(\mathbf{x}) = A\mathbf{x}$ injective? Surjective? Give an "if and only if" criterion for each of the two conditions in terms of each of the following. (So the final answer will be a table with two columns and five rows, the two entries in the top row being injective and surjective.)

(a) Solutions of a suitable linear system

(b) REF(A)

(c) Column vectors of A

(d) Row vectors of A

Proof. (a) This is just a rephrasing of injective and surjective. It is injective iff there is a unique solution to $A\mathbf{x} = \mathbf{0}$[1], and surjective if there is a solution to $A\mathbf{x} = \mathbf{b}$ for all $\mathbf{b}$ in the co-domain of the function.

(b) It is injective iff every column of the REF has a pivot. It is surjective iff every row has a pivot. I have already proved this in HW 1, for a RREF. Using problem 7, to show that pivots are unchanged across REFs, we will be done.

(c). It is injective iff column vectors are linearly independent, and surjective if they span the co-domain (obvious from Problem 1).

(d). We will use (b). We claim that it is injective iff row vector spans the domain. Note that every column has a pivot implies that the rows span the domain (this can be seen from the RREF. I argue in HW3 how elementary row operations do not change row space.) The reverse is also true, if the rows span the domain, and if there is a column, say the kth one, that does not have a pivot, note that we will have the standard basis vectors $\mathbf{e}_i$ for $i \neq k$ as the rows of A_{RREF}, but $\mathbf{e}_k$ will not exist. Since these form a basis, this shows that the row vectors cannot span the domain.

Similarly, it is surjective iff the row vectors are linearly independent. Again, we use (b). One can see that each row having a pivot will ensure linear independence. For the converse, suppose some row does not have a pivot. It must then be a zero row in the RREF. It remains to show that the row operations preserve linear independence, but that is easy to do.

[1]Since $A\mathbf{x} = A\mathbf{y} \Leftrightarrow A(\mathbf{x} - \mathbf{y}) = 0$.

Problem 6

(a) Show that doing an elementary row operation on a matrix A results in a matrix EA, where E is a suitable matrix, called an *elementary matrix*. Describe all three types of elementary matrices.

(b) Show that E itself can be obtained using an elementary row operation on a very simple matrix.

Proof.

We define each E by applying the row operation to I_m, and then check that the same E works for every $m \times n$ matrix A.

Write $\boldsymbol{a}_k$ for the kth row of A and $\boldsymbol{e}_k^T$ for the kth row of I_m. The one fact we need: for any $m \times m$ matrix B,

$$(BA)_{kj} = \sum_{l=1}^{m} B_{kl} A_{lj},$$

so the kth row of BA is $\sum_l B_{kl}\boldsymbol{a}_l$. In words, row k of BA is the combination of rows of A whose coefficients are read off from row k of B. In particular, if row k of B is $\boldsymbol{e}_l^T$, then row k of BA is $\boldsymbol{a}_l$.

Swap. $E_{i\leftrightarrow j}$ has rows $\boldsymbol{e}_k^T$ for $k \neq i, j$, row i equal to $\boldsymbol{e}_j^T$, and row j equal to $\boldsymbol{e}_i^T$. By the fact above, $E_{i\leftrightarrow j}A$ has rows $\boldsymbol{a}_k$ for $k \neq i, j$, row i equal to $\boldsymbol{a}_j$, and row j equal to $\boldsymbol{a}_i$. That is exactly A with rows i and j swapped.

Scale. For $\lambda \neq 0$, $E_{i\to\lambda i}$ agrees with I_m except that row i is $\lambda\boldsymbol{e}_i^T$. So $E_{i\to\lambda i}A$ agrees with A except that row i is $\lambda\boldsymbol{a}_i$.

Add. For $i \neq j$, $E_{i\to i+\lambda j}$ agrees with I_m except that row i is $\boldsymbol{e}_i^T + \lambda\boldsymbol{e}_j^T$, i.e. it has a λ in position (i, j). So $E_{i\to i+\lambda j}A$ agrees with A except that row i is $\boldsymbol{a}_i + \lambda\boldsymbol{a}_j$.

In each case E depends only on the operation and on m, not on A. This proves (a).

For (b): by construction, each E is the result of applying the corresponding row operation to I_m. Equivalently, taking $A = I_m$ in (a) gives $\mathrm{op}(I_m) = EI_m = E$.

Problem 7

Useful as problem-solving/proof-writing practice, but not conceptually central. Recall:

(i) A matrix is in REF if leading nonzero entries in successive rows move strictly to the right and all zero rows are at the bottom. We can convert any matrix to REF by a forward/downward pass of row operations. We also scale the leading nonzero entries, called pivots, to 1.

(ii) A matrix is in RREF if it is in REF and moreover all entries above any pivot are 0. We can convert any REF matrix to RREF by a backward/upward pass, killing entries above pivots column by column starting from the right.

(a) Show that starting with an REF matrix, we can get to an RREF matrix by treating columns in any order.

(b) Show that any matrix can be converted to a unique matrix in RREF by row operations.

Proof.

(a) Suppose the pivot in row i is in the column ρ_i. Let us work with a specific ordering of the pivot columns, and thus rows, $\sigma : [k] \to [k]$. Starting with $\sigma(1)$, we clear all the entries in $\rho_{\sigma(1)}$ (except in row $\sigma(1)$ of course), by the row operation, row j - $a_{j\rho_{\sigma(1)}}$ row $\sigma(1)$ for the $j < \sigma(1)$ (because we already had a REF, the entries in $\rho_{\sigma(1)}$ are zero below the row $\sigma(1)$).

Now note that doing this procedure on the row $\sigma(2)$ does not change $\rho_{\sigma(1)}$. The row operations we did involve scaling, and subtracting this row from all the rows above it. But, scaling doesn't change that any entry (except the pivot) in $\rho_{\sigma(1)}$ is 0, and neither will subtracting 0 from 0. The pivot is of course unchanged as well, because we only ever subtract 0 from it. Hence, these procedures do not impact any other pivot columns. So, we can independently do them for $\sigma(1), \ldots, \sigma(k)$ and get that the only non-zero entries in $\rho_{\sigma(i)}$ is the pivot, and thus have the matrix in RREF.

(b) We already know that a matrix can be converted to a REF. We also saw that it can be converted to a RREF in (a). It suffices to show that it is unique. We may do this inductively, inducting on the number of columns. It is obvious that this is true for one column, since then the RREF is $\mathbf{0}$ if the column is zero, and $\mathbf{e}_1$ otherwise (row operations are reversible, so they cannot make a non-zero column zero).

Note that row operations do not change the solutions of $A\mathbf{x} = \mathbf{0}$, again because they are reversible. So if R and S are both RREFs of A, then $R\mathbf{x} = \mathbf{0}$ iff $S\mathbf{x} = \mathbf{0}$.

Now let A have n columns, and let R and S be RREFs of A. Deleting the last column of R and S gives RREFs of A with its last column deleted (a row whose pivot was in the last column just becomes a zero row at the bottom). So by induction, R and S can only differ in the last column. Suppose they do. Then for any $\mathbf{x}$ with $R\mathbf{x} = \mathbf{0}$,

$$\mathbf{0} = (R - S)\mathbf{x} = x_n(\mathbf{r}_n - \mathbf{s}_n)$$

where $\mathbf{r}_n, \mathbf{s}_n$ are the last columns, so $x_n = 0$. So x_n is not a free variable, i.e. the last column of R has a pivot, and similarly for S. But then if k is the number of pivots in the other columns, the last column of both is $\mathbf{e}_{k+1}$, which means they do not differ. Thus $R = S$.