

## Math 314

### Topics for first exam

#### Systems of linear equations

$$2x - 3y - z = 6$$

$$3x + 2y + z = 7$$

Goal: find simultaneous solutions: all  $x, y, z$  satisfying both equations.

Most general type of system:

$$a_{11}x_1 + \cdots + a_{1n}x_n = b_1$$

...

$$a_{m1}x_1 + \cdots + a_{mn}x_n = b_m$$

#### Gaussian elimination: basic ideas

$$3x + 5y = 2$$

$$2x + 3y = 1$$

Idea use  $3x$  in first equation to eliminate  $2x$  in second equation. How? Add a multiple of first equation to second. Then use  $y$ -term in new second equation to remove  $5y$  from first!

The point: a solution to the original equations **must** also solve the new equations. The **real** point: it's much less complicated to figure out the solutions of the new equations!

Streamlining: keep only the essential information; throw away unneeded symbols!

$$\begin{array}{ll} 3x+5y=2 & \text{replace} \\ 2x+3y=1 & \text{with} \end{array} \quad \left( \begin{array}{cc|c} 3 & 5 & 2 \\ 2 & 3 & 1 \end{array} \right)$$

We get an **(augmented) matrix** representing the system of equations. We carry out the same operations we used with equations, but do them to the rows of the matrix.

Three basic operations (elementary row operations):

$E_{ij}$  : switch  $i$ th and  $j$ th rows of the matrix

$E_{ij}(m)$  : add  $m$  times  $j$ th row to the  $i$ th row

$E_i(m)$  : multiply  $i$ th row by  $m$

Terminology: first non-zero entry of a row = **leading entry**; leading entry used to zero out a column = **pivot**.

Basic procedure (Gauss-Jordan elimination): find non-zero entry in first column, switch up to first row ( $E_{1j}$ ) (pivot in (1,1) position). Use  $E_1(m)$  to make first entry a 1, then use  $E_{1j}(m)$  operations to zero out the other entries of the first column. Then: find leftmost entry in remaining rows, switch to second row, use as a pivot to clear out the entries in the column below it. Continue (forward solving). When done, use pivots to clear out entries in column above the pivots (back-solving).

Variable in linear system corresponding to a pivot = **bound** variable; other variables = **free** variables

#### Gaussian elimination: general procedure

**The big fact:** After elimination, the new system of linear equations have the exact **same solutions** as the old system. Because: row operations are reversible!

Reverse of  $E_{ij}$  is  $E_{ij}$ ; reverse of  $E_{ij}(m)$  is  $E_{ij}(-m)$ ; reverse of  $E_i(m)$  is  $E_i(1/m)$

So: you can get old equations from new ones; so solution to new equations **must** solve old equations **as well**.

Row echelon form: apply elementary row operations so turn matrix  $A$  into one so that

- (a) each row looks like  $(000 \cdots 0 ** \cdots *)$ ; first  $*$  = leading entry
- (b) leading entry for row below is further to the right

Reduced row **echelon** form: in addition, have

- (c) each leading entry is = 1
- (d) each leading entry is the only non-zero number in its column.

REF can be achieved by forward solving; RREF by back-solving and  $E_i(m)$  's

Elimination: every matrix can be put into RREF by elementary row operations.

Big Fact: If a matrix  $A$  is put into RREF by two different sets of row operations, you get the **same matrix**.

With the RREF of an augmented matrix: can read off solutions to linear system.

$$\left( \begin{array}{cccc|c} 1 & 0 & 2 & 0 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 3 \end{array} \right) \text{ means } \begin{array}{l} x_4=3, x_2=1-x_3 \\ x_1=2-2x_3 \text{ ; } x_3 \text{ is free} \end{array}$$

Inconsistent systems: row of zeros in coefficient matrix, followed by a non-zero number (e.g., 2).  
Translates as  $0=2$  ! System has no solutions.

We can understand the existence and number of solutions using

- number of pivot variables = number of columns in (R)REF containing a pivot
- number of free variables = number of columns in (R)REF not containing a pivot

$A$  = coefficient matrix,  $\tilde{A}$  = augmented matrix ( $A = m \times n$  matrix)

System is consistent if and only if in (R)REF a row containing no pivot consists of all 0's in  $\tilde{A}$ .

A consistent system has a unique solution if and only if there are no free variables; there is a pivot in every row.

## Spanning sets and linear independence

We can interpret an SLE in terms of (column) vectors; writing  $v_i$  =  $i$ th column of the coefficient matrix, and  $b$ =the column of target values, then our SLE really reads as a single equation  $x_1 v_1 + \cdots + x_n v_n = b$ . The lefthand side of this equation is a *linear combination* of the vectors  $v_1, \dots, v_n$ , that is, a sum of scalar multiples. Asking if the SLE has a solution is the same as asking if  $b$  is a linear combination of the  $v_i$ .

This is an important enough concept that we introduce new terminology for it; the *span* of a collection of vectors,  $\text{span}(v_1, \dots, v_n)$  is the collection of all linear combinations of the vectors. If the span of the  $(m \times 1)$  column vectors  $v_1, \dots, v_n$  is all of  $\mathbb{R}^m$ , we say that the vectors span  $\mathbb{R}^m$ . Asking if an SLE has a solution is the same as asking if the target vector is in the span of the column vectors of the coefficient matrix.

The flipside of spanning is *linear independence*. A collection of vectors  $v_1, \dots, v_n$  is linearly independent if the only solution to  $x_1 v_1 + \cdots + x_n v_n = 0$  (the 0-vector) is  $x_1 = \cdots = x_n = 0$  (the "trivial" solution). If there is a non-trivial solution, then we say that the vectors are linearly dependent. If a collection of vectors is linearly dependent, then choosing a non-trivial solution and a vector with non-zero coefficient, throwing everything else on the other side of the equation expresses one vector as a linear combination of all of the others. Thinking in terms of an SLE, the columns of a matrix are linearly dependent exactly when the SLE with target 0 has a non-trivial solution, i.e., has more than one solution. It has the trivial (all 0) solution, so it is consistent, so to have more than one, we need the the RREF for the matrix to have a free variable, i.e., the rank of the coefficient matrix is less than the number of columns.

## Matrix addition and scalar multiplication

Idea: take our ideas from vectors. Add entry by entry. Constant multiple of matrix: multiply entry by entry.

$\mathbf{0}$  = matrix all of whose entries are 0

Basic facts:

$A+B$  makes sense only if  $A$  and  $B$  are the same size ( $m \times n$ ) matrix

$$A+B = B+A$$

$$(A+B)+C = A+(B+C)$$

$$A+\mathbf{0} = A$$

$$A+(-1)A = \mathbf{0}$$

$cA$  has the same size as  $A$

$$c(dA) = (cd)A$$

$$(c+d)A = cA + dA$$

$$c(A+B) = cA + cB$$

$$1A = A$$

## Matrix equations

We can also think of a system of equations as a single matrix equation  $A\vec{x} = \vec{b}$ , by interpreting the product of our coefficient matrix  $A$  and our column vector of variables  $\vec{x}$  as the  $\vec{x}$ -linear combination of the columns of  $A$ . This multiplication behaves well with vector addition and scalar multiplication

$$A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v} \quad A(a\vec{u}) = a \cdot A\vec{u}$$

(We say that this matrix multiplication is linear.) Because of this, any two solutions to  $A\vec{x} = \vec{b}$ , say  $\vec{u}$  and  $\vec{v}$  have  $A(\vec{u} - \vec{v}) = \vec{b} - \vec{b} = \vec{0}$ , the all-0 vector. If we treat  $\vec{u}$  as a “particular” solution to  $A\vec{x} = \vec{b}$ , then any other solution  $\vec{v} = \vec{u} + (\vec{v} - \vec{u}) = \vec{u} + \vec{n}$  is our particular solution plus a solution to  $A\vec{x} = \vec{0}$ , a “homogeneous” solution.

But we can show that the solutions to  $A\vec{x} = \vec{0}$  are precisely the span of a collection of vectors, one for each free variable of the RREF of  $A$ . This is done by row reducing the system  $A\vec{x} = \vec{0}$  and writing the solutions in terms of the free variables; this can be expressed as a linear combination of the solutions obtained by setting each free variable equal to 1, in turn (and all of the others to 0).

## Some applications

A model of an economy:

An economy consists of ‘sectors’, whose output is used as materials for other sectors (including its own). The exchange of outputs is carried out by paying for it; each sector has a price for its output. At equilibrium, the amount paid by each sector for the resources it needs will equal the income it earns from the sale of its output. We assume all output is consumed by the sectors (no surplus, no shortage). To determine what prices should be charged by each sector in this equilibrium, we solve a system of equations which asserts that (for sectors 1 through  $n$ )

the amount earned,  $p_i$  by a sector equals the total amount paid  $a_{i,1}p_1 + \cdots + a_{i,n}p_n$  by the sector for its resources (where the  $a_{i,j}$  are (known) fractions of sector  $j$ ’s output allocated to sector  $i$  (and so, for each  $j$ ,  $a_{1,j} + \cdots + a_{n,j} = 1$ ).

Solving the system will always yield a free variable; setting that variable to 1 yields the relative prices to set (we can always double all prices and still solve the equations!).

Balancing chemical reactions:

In a chemical reaction, some collection of molecules is converted into some other collection of molecules. The proportions of each can be determined by solving an SLE:

E.g., when ethane is burned,  $x$   $C_2H_6$  and  $y$   $O_2$  is converted into  $z$   $CO_2$  and  $w$   $H_2O$ . Since the number of each element must be the same on both sides of the reaction, we get a system of equations

$$C : 2x = z ; H : 6x = 2w ; O : 2y = w .$$

which we can solve. More complicated reactions, e.g.,  $PbO_2 + HCl \rightarrow PbCl_2 + Cl_2 + H_2O$ , yield more complicated equations, but can still be solved using the techniques we have developed.

Network Flow:

We can model a network of water pipes, or traffic flowing in a city's streets, as a graph, that is, a collection of points = vertices (=intersections=junctions) joined by edges = segments (=streets = pipes). Monitoring the flow along particular edges can enable us to know the flow on every edge, by solving a system of equations; at every vertex, the net flow must be zero. That is, the total flow into the vertex must equal the total flow out of the vertex. Giving the edges arrows, depicting the direction we think traffic is flowing along that edge, and labeling each edge with either the flow we know (monitored edge) or a variable denoting the flow we don't, we have a system of equations (sum of flows into a vertex) = (sum of flows out of the vertex). Solving this system enables us to determine the value of every variable, i.e., the flow along every edge. A negative value means that the flow is opposite to the one we expected.

## Linear transformations

From a different point of view, the matrix product  $A\vec{x} = \vec{b}$  can be viewed as a function  $A$  takes a vector  $\vec{x}$  and produces a vector  $A\vec{x}$  (the vector of dot products of the rows of  $A$  with  $\vec{x}$ ). This function  $T_A(\vec{x}) = A\vec{x}$  is linear:

$$T_A(\vec{u} + \vec{v}) = T_A(\vec{u}) + T_A(\vec{v}) \quad T_A(a\vec{u}) = aT_A(\vec{u})$$

more generally, any function  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  satisfying these properties is called a *linear transformation*. We can find them in many of the operations we carry out on vectors; reflections, rotations, and many other geometric operations (that leave the origin fixed) can be viewed as linear transformations. [Note that since  $T(\vec{0}) = T(0 \cdot \vec{0}) = 0 \cdot T(\vec{0}) = \vec{0}$ , any linear transformation must send  $\vec{0}$  to  $\vec{0}$ .]

Matrix multiplication is our model for a linear transformation. In fact, every lin. transf. can be viewed as multiplication by some  $A$ , since, letting  $\vec{e}_i$  stand for the  $i$ -th standard coordinate vector (all 0's, except for a 1 in the  $i$ -th place),  $A\vec{e}_i$  is the  $i$ -th column of  $A$ . So if we build a matrix  $A$  with  $i$ -th column equal to  $T(\vec{e}_i)$ , then  $A\vec{x}$  and  $T(\vec{x})$  must be equal, since using linearity both are equal to the  $\vec{x}$  - *linear combination of the columns of*  $A = T(\vec{x})$ .

For example, we can find the matrix corresponding to rotation  $R_\theta$  by an angle  $\theta$  by determining that  $R_\theta(\vec{e}_1) = (\cos \theta, \sin \theta)^T$  and  $R_\theta(\vec{e}_2) = (-\sin \theta, \cos \theta)^T$ .

## Matrix multiplication

Idea:  $A\vec{x}$  is a function  $T_A$ , so  $A(B\vec{x}) = T_A(T_B(\vec{x})) = (T_A \circ T_B)(\vec{x})$  is the composition of two linear functions, which is linear! So  $T_A \circ T_B$  should be equal to multiplication by some matrix, which we will call  $AB$ ! To figure it out, we use the fact that the columns of  $AB$  should be the vectors  $A(B\vec{e}_i)$ , which is  $A$  times the  $i$ -th column of  $B$ . So we define matrix multiplication this way!

In  $AB$ , each row of  $A$  is 'multiplied' by each column of  $B$  to obtain an entry of  $AB$ . Need: the length of the rows of  $A$  (= number of columns of  $A$ ) = length of columns of  $B$  (= number of rows of  $B$ ). I.e, in order to multiply,  $A$  must be  $m \times n$ , and  $B$  must be  $n \times k$ ;  $AB$  is then  $m \times k$ .

Formula:  $(i,j)$ th entry of  $AB$  is  $\sum_{k=1}^n a_{ik}b_{kj}$  (the dot product of the  $i$ -th row of  $A$  and the  $j$ -th column of  $B$ ).

I = identity matrix; square matrix ( $n \times n$ ) with 1's on diagonal, 0's off diagonal

Basic facts:

$$AI = A = IA$$

$$(AB)C = A(BC)$$

$$c(AB) = (cA)B = A(cB)$$

$$(A+B)C = AC + BC$$

$$A(B+C) = AB + AC$$

In general, however it is **\*\*not\*\*** true that  $AB$  and  $BA$  are the same; they are almost always different! \*\*\*\*

### Special matrices and transposes

Elementary matrices:

A row operation ( $E_{ij}$ ,  $E_{ij}(m)$ ,  $E_i(m)$ ) applied to a matrix  $A$  corresponds to multiplication (on the left) by a matrix (also denoted  $E_{ij}$ ,  $E_{ij}(m)$ ,  $E_i(m)$ ). The matrices are obtained by applying the row operation to the identity matrix  $I_n$ . E.g., the  $4 \times 4$  matrix  $E_{13}(-2)$  looks like  $I$ , except it has a  $-2$  in the  $(1,3)$ th entry.

The idea: if  $A \rightarrow B$  by the elementary row operation  $E$ , then  $B = EA$ .

So if  $A \rightarrow B \rightarrow C$  by elementary row operations, then  $C = E_2 E_1 A \dots$

Row reduction **is** matrix multiplication!

A scalar matrix  $A$  has the same number  $c$  in the diagonal entries, and 0's everywhere else (the idea:  $AB = cB$ )

A diagonal matrix has all entries zero off of the (main) diagonal

A upper triangular matrix has entries =0 below the diagonal, a lower triangular matrix is 0 above the diagonal. A triangular matrix is either upper or lower triangular.

A strictly triangular matrix is triangular, and has zeros **on** the diagonal, as well. They come in upper and lower flavors.

The **transpose** of a matrix  $A$  is the matrix  $A^T$  whose columns are the rows of  $A$  (and vice versa).  $A^T$  is  $A$  reflected across the main diagonal.  $(a_{ij})^T = (a_{ji})$ ;  $(m \times n)^T = (n \times m)$

Basic facts:

$$(A+B)^T = A^T + B^T$$

$$(AB)^T = B^T A^T$$

$$(cA)^T = cA^T$$

$$(A^T)^T = A$$

Transpose of an elementary matrix is elementary:

$$E_{ij}^T = E_{ij}, E_{ij}(m)^T = E_{ji}(m), E_i(m)^T = E_i(m)$$

### Matrix inverses

One way to solve  $A\vec{x} = \vec{b}$ : 'divide' by  $A$ ! That is, find a matrix  $B$  with  $BA = I$ . When is there such a matrix?

This tells us that the solution is unique; it must be  $B\vec{b}$ . So in (R)REF  $A$  must have no free variables. So there is a pivot in every column. But it also tells us that the system has a solution! So there must be a pivot in every row, as well. This implies that the RREF of  $A$  must be the identity matrix (in part,  $A$  must be square).

A matrix  $B$  is an **inverse** of  $A$  if  $AB=I$  and  $BA=I$ ; it turns out, the inverse of a matrix is always unique. We call it  $A^{-1}$  (and call  $A$  invertible).

Finding  $A^{-1}$ : row reduction! (of course...)

Build the "super-augmented" matrix  $(A|I)$  (the matrix  $A$  with the identity matrix next to it). Row reduce  $A$ , and carry out the operations on the entire row of the S-A matrix (i.e., carry out the

identical row operations on I). When done, if invertible+ the left-hand side of the S-A matrix will be I; the right-hand side will be  $A^{-1}$  !

I.e., if  $(A|I) \rightarrow (I|B)$  by row operations, then  $I=BA$  .

We can also understand this via row operations, since, it turns out, each row operation can be viewed as multiplication (on the left) by an ‘elementary’ matrix. So if we multiply A (on the left) by the correct elementary matrices, you get I; call the product of those matrices B and you get  $BA=I$  !

Basic facts:

$$(A^{-1})^{-1} = A$$

if A and B are invertible, then so is AB, and  $(AB)^{-1} = B^{-1}A^{-1}$

$$(cA)^{-1} = (1/c)A^{-1}$$

$$(A^T)^{-1} = (A^{-1})^T$$

If A is invertible, and  $AB = AC$ , then  $B = C$ ; if  $BA = CA$ , then  $B = C$ .

Inverses of elementary matrices:

$$E_{ij}^{-1} = E_{ij}, E_{ij}(m)^{-1} = E_{ij}(-m), E_i(m)^{-1} = E_i(1/m)$$

Highly useful formula: for a 2-by-2 matrix,

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ and } D=ad-bc, \quad A^{-1} = \frac{1}{D} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

(Note: need  $D=ad-bc \neq 0$  for this to work....)

Some conditions for/consequences of invertibility: the following are all equivalent ( $A = n$ -by- $n$  matrix).

1. A is invertible,
2. A has  $n$  pivots (the *rank* of A is  $n$ ).
3. The RREF of A is  $I_n$ .
4. Every linear system  $Ax = b$  has a unique solution.
5. For one choice of  $b$ ,  $Ax = b$  has a unique solution (i.e., if one does, they all do...).
6. The equation  $Ax = 0$  has only the solution  $x = 0$ .
7. There is a matrix B with  $BA = I$ .
8. There is a matrix B with  $AB = I$ .

The equivalence of 4. and 6. is sometimes stated as **Fredholm’s alternative**: Either every equation  $Ax=b$  has a unique solution, or the equation  $Ax=0$  has a **non-trivial** solution (and only one of the alternatives can occur).