Lecture 3: Linear Algebra I

Robert Garrard

Systems of Linear Equations

A **linear** equation in n variables is one in which the highest order of a variable is unity. We're often concerned with solving a set of m linear equations in n variables *simultaneously*.

$$\begin{array}{ccccccccc}
a_{11}x_{11} & + & a_{12}x_{12} & + & \cdots & + & a_{1n}x_{1n} & = & b_1 \\
a_{21}x_{21} & + & a_{22}x_{22} & + & \cdots & + & a_{2n}x_{2n} & = & b_2 \\
a_{31}x_{31} & + & a_{32}x_{32} & + & \cdots & + & a_{3n}x_{3n} & = & b_3 \\
\vdots & + & \vdots & + & \cdots & + & \vdots & = & \vdots \\
a_{m1}x_{m1} & + & a_{m2}x_{m2} & + & \cdots & + & a_{mn}x_{mn} & = & b_m
\end{array}$$

Systems of Linear Equations

Subscripts are ordered by row then column. That is, x_{24} denotes the second equation and fourth variable.

It's convenient to write this system in **augmented matrix** form, where we drop superfluous objects.

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right)$$

Systems of Linear Equations

An **elementary row operation** on a system of linear equations is one of the following.

1. Multiplying a row by a non-zero scalar.
2. Adding two rows together.
3. Switching two rows.

Proposition

The solution of a system of linear equations is preserved by elementary row operations.

Systems of Linear Equations

- So we can solve a system by reducing it to something simpler using row operations.
- There are two good candidates for the form of that simplification: row echelon form and reduced row echelon form.
- Row echelon form is obtained in fewer row operations which, as you may know from experience, are prone to arithmetic errors when performed by a human.
- Reduced row echelon form has the cost of additional row ops, but has the distinct advantage that the solution set is more easily characterised when there are infinitely many solutions.

Systems of Linear Equations

The **pivot** of a non-zero row is its leftmost non-zero element.[1]

A system of linear equations is said to be in **row echelon form** if the pivot of a non-zero row is strictly to the right of the pivot of the row above it.

A system is said to be in **reduced row echelon form** if:

1. It is in row echelon form.
2. Each pivot is the only non-zero entry in its column.
3. Each pivot is a one.

[1] Some texts require that this element be a 1.

Systems of Linear Equations

Exercise 3.1

Which of the following systems are in row echelon form? For any that aren't, which row operations would you use?

$$\left(\begin{array}{ccc|c} 2 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{array}\right) \quad \left(\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 4 & 0 & 0 \end{array}\right) \quad \left(\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Which of the following systems are in reduced row echelon form? For any that aren't, which row operations would you use?

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \quad \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Systems of Linear Equations

If the ith column of the reduced row echelon form contains a pivot, we call x_i a **basic variable**.

If it does not contain a pivot, we call x_i a **free variable**.

Exercise 3.2
Place the following systems in reduced row echelon form. State if each variable is basic or free.

$$\left(\begin{array}{ccc|c} 2 & 1 & 0 & 1 \\ 2 & 1 & 1 & 2 \\ 4 & 3 & 0 & 3 \end{array}\right) \quad \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 2 & 4 & 4 & 2 & 1 \\ 0 & 1 & 2 & 4 & 1 \end{array}\right)$$

Systems of Linear Equations

The **rank** of a matrix is the number of non-zero rows in its row echelon form.[2]

Let A be an $m \times n$ matrix.

- If $\text{rank}(A) = \min\{m, n\}$, A is said to be **full rank**.
- Otherwise it is **rank deficient**.
- A matrix whose rank is m is said to have **full row rank** and whose rank is n is said to have **full column rank**.

Now that we can reduce systems to something simpler, how do we go about characterising the solution?

[2] While a system does not necessarily have a unique row echelon form, the number of non-zero rows is the same in every row echelon form.

Systems of Linear Equations

Recipe for Solving a System of Linear Equations

1. Use row operations to put the augmented system in reduced row echelon form.
2. Is the last column of the augmented system a pivot column? If yes, there is no solution. If no, proceed.
3. Are there any free variables? If no, read the unique solution off of each row. If yes, proceed.
4. Solve for each basic variable in terms of the free variables and a constant.

Systems of Linear Equations

Example

Solve the following system of equations.

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

System is already in RREF. The last column does not contain a pivot, so there's at least one solution. There are two free variables, x_3 and x_4, so there are infinitely many solutions. Solving in terms of free variables and constants gives

$$\begin{aligned} x_1 &= 1 - x_3 \\ x_2 &= -1 - 2x_3 - 2x_4 \\ x_3 &= x_3 \\ x_4 &= x_4 \qquad \text{for any } x_3, x_4 \in \mathbb{R} \end{aligned}$$

Systems of Linear Equations

Example

$$
\begin{aligned}
x_1 &= 1 - x_3 \\
x_2 &= -1 - 2x_3 - 2x_4 \\
x_3 &= x_3 \\
x_4 &= x_4 \qquad \text{for any } x_3, x_4 \in \mathbb{R}
\end{aligned}
$$

Can be written more compactly in vector notation

$$
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} 0 \\ -2 \\ 0 \\ 1 \end{pmatrix} x_4
$$

We can see that the solution set is a plane in $\mathbb{R}^4$.

Vector Spaces

A **vector space** is a structure formed by a set V (whose elements are called vectors) together with two operations: addition and scalar multiplication.

We will only be considering scalars to be the set of real numbers, $\mathbb{R}$, but in general they may be any **field**.

For V to be a vector space over some field F equipped with the binary operations of vector addition and scalar multiplication, then $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, and for all scalars $r, s \in F$, the following conditions must hold:

Vector Spaces

Vector space axioms

1. *Commutativity of vector addition* $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
2. *Associativity of vector addition* $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
3. *Additive identity:* There is a $\mathbf{0}$ s.t $\mathbf{0} + \mathbf{x} = \mathbf{x}$
4. *Additive inverse:* $\forall \mathbf{x}$ there exists a $-\mathbf{x}$ s.t $\mathbf{x} + -\mathbf{x} = \mathbf{0}$
5. *Associativity of scalar multiplication* $r(s\mathbf{x}) = (rs)\mathbf{x}$
6. *Distributivity of scalar sums* $(r + s)\mathbf{x} = r\mathbf{x} + s\mathbf{x}$
7. *Distributivity of vector sums* $r(\mathbf{x} + \mathbf{y}) = r\mathbf{x} + r\mathbf{y}$
8. *Scalar identity* $1 \cdot \mathbf{x} = \mathbf{x}$

Vector Spaces

Exercise 3.3 Which of the following, if any, constitute a vector space?

1. $(x_1, \ldots, x_n) \in \mathbb{R}^n$ with componentwise addition.
2. The set of square matrices under matrix addition.
3. The set of 3×3 matrices under matrix addition.
4. The set of 3×3 matrices under matrix multiplication.
5. The set of functions from $\mathbb{R}$ to $\mathbb{R}$ under $(f + g)(x) = f(x) + g(x)$.

Vector Spaces

Let V be a vector space over $\mathbb{R}$ and let $v_1, \ldots, v_n$ be vectors in V.

The vectors $v_1, \ldots, v_n$ are **linearly independent** if, for scalars $c_1, \ldots, c_n \in \mathbb{R}$, the only solution to the equation

$$c_1 v_1 + \cdots + c_n v_n = 0$$

is $c_1 = 0, \ldots, c_n = 0$.

The vectors $v_1, \ldots, v_n$ are **spanning** if for every vector $v \in V$, we can find $c_1, \ldots, c_n \in \mathbb{R}$ such that

$$v = c_1 v_1 + \cdots + c_n v_n$$

In which case we may write $V = \operatorname{span}\{v_1, \ldots, v_n\}$.

Vector Spaces

The vectors $v_1, \ldots, v_n$ form a **basis** for V if they are linearly independent and spanning.

A vector space is finite dimensional with dimension n if we can find a vectors $v_1, \ldots, v_n \in V$ which form a basis for V.

Exercise 3.4
Show that $\mathbb{R}^n$ is n-dimensional.

Subspaces

A subset $W \subset V$ is called a **subspace** if and only if

1. $\mathbf{0} \in W$
2. If $\mathbf{u}, \mathbf{v} \in W$ then $\mathbf{u} + \mathbf{v} \in W$
3. If $\mathbf{u} \in W$ then $\alpha \mathbf{u} \in W$ for any scalar $\alpha \in \mathbb{R}$

Let $v_1, \ldots, v_n \in V$. The **span** of the list of vectors is

$$\text{span}\{v_1, \ldots, v_n\} = \{c_1 v_1 + \cdots + c_n v_n \mid \forall c_1, \ldots, c_n \in \mathbb{R}\}$$

Note that the span of a list of vectors forms a subspace. Moreover, the list is a spanning set for that subspace.

Subspaces

Exercise 3.5

1. What's the difference between a spanning set and a basis?

2. Let $v_1 = [1, 1, 0]'$, $v_2 = [1, 0, 1]'$, $v_3 = [1, 1, 1]'$.
   a) Find $\text{span}\{v_1, v_2, v_3\}$.
   b) Find a basis for the spanning subspace.
   c) What's the dimension of the spanning subspace?

3. Under what conditions is a plane a subspace of $\mathbb{R}^3$?

Linear Mappings

Consider the vector spaces $\mathbb{R}^m$ and $\mathbb{R}^n$. A function $f : \mathbb{R}^m \to \mathbb{R}^n$ is a **linear map** if:

- $f(x + y) = f(x) + f(y) \;\; \forall x, y \in V.$
- $f(\alpha x) = \alpha f(x) \;\; \forall \alpha \in \mathbb{R},\; x \in V.$

Proposition

Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be a linear map. Further, let the vectors of $\mathbb{R}^m$ and $\mathbb{R}^n$ be column vectors. Then there exists a matrix A such that

$$f(\mathbf{v}) \; = A\mathbf{v}$$

That is to say that matrices are not just a convenience for solving systems of linear equations, instead they may be thought of as linear functions between vector spaces.

Linear Mappings

But if they're functions, do they have analogues of certain interesting structures functions have?

Recall that for a function $f : X \to Y$, the **image** of that function is the set

$$\text{im}(f) = \{y \in Y \mid \exists x \in X s.t\ f(x) = y\}$$

Equations like $f(x) = b$ have a solution if and only if $b \in im(f)$.

Furthermore, if the function is **surjective**, meaning the image of the function equals its codomain, then for any b there is a solution to $f(x) = b$. How might this look in the special case of *linear* functions.

The Column Space

Consider the function $A\mathbf{v}$ for some arbitrary $\mathbf{v}$. We can decompose it to

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} v_1 + \begin{pmatrix} a_{12} \\ \vdots \\ a_{m2} \end{pmatrix} v_2 + \cdots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} v_n$$

$$= \mathbf{a}_1 v_1 + \mathbf{a}_2 v_2 + \cdots + \mathbf{a}_n v_n$$

The output of the function is a linear combination of the columns of A. So the image of the function must be the subspace spanned by the columns of A!

The Column Space

The **column space** of an $m \times n$ matrix A, $\text{Col}(A)$, is the subset of $\mathbb{R}^n$ spanned by its columns.

Exercise 3.6
Consider the following matrices.

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

What do you notice about the relationship between these matrices? Find a basis for the col space of each, what can you conclude?

The Column Space

Theorem

The basic columns of A form a basis for $\mathrm{Col}(A)$.

While row ops don't preserve the span of the columns, they do preserve any dependence relationship between them.

That is, linearly independent columns will still be linearly independent in a row echelon form. This lets us easily refine a spanning set (the whole set of columns) down to a basis (a minimal spanning set).

The Column Space

Theorem

Let A be an $m \times n$ matrix.

a) *The system of equations $A\mathbf{x} = \mathbf{b}$ has a solution for a particular $\mathbf{b} \in \mathbb{R}^m$ if and only if $\mathbf{b} \in \mathrm{Col}(A)$.*
b) *The system $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ if and only if $rank(A) = m$. That is, A has full row rank.*
c) *dim Col(A) = Rank(A).*

The Null Space

Another interesting property functions have is the set of zeros: $\{x \mid f(x) = 0\}$.(i.e., the pre-image of 0)

Define the **null space** (elsewhere called the **kernel**) of an $m \times n$ matrix A, denoted $\text{Null}(A)$, to be the set of vectors that solve $A\mathbf{x} = \mathbf{0}$.

Exercise: Show that the null space is a subspace of $\mathbb{R}^n$.

Affine Subspaces

Let V be a subspace of $\mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^n$ be a fixed vector. A set of the form

$$\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{c} + \mathbf{v} \text{ for some } \mathbf{v} \in V\}$$

is called an **affine subspace** of $\mathbb{R}^n$.

Affine subspaces are not subspaces! (why?)

It does preserve linearity, and we say its dimension is the dimension of V.

The Solution Set is an Affine Subspace

Let $A\mathbf{x} = \mathbf{b}$ be an $m \times n$ system of equations and let $\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \, | A\mathbf{x} = \mathbf{b}\}$ be the set of solutions to the system.

Theorem

If $\mathcal{S}$ is non-empty, such that there is at least one particular solution, $\mathbf{x}^*$, such that $A\mathbf{x}^* = \mathbf{b}$. Then $\mathcal{S}$ is the affine subspace

$$\mathcal{S} = \{\mathbf{x}^* + \mathbf{v} \mid \mathbf{v} \in \mathit{Null}(A)\}$$

Proof (tutorial excercise)

The Fundamental Theorem of Linear Algebra

Theorem (Rank-Nullity theorem)

Let A be an $m \times n$ matrix. Then

$$rank(A) + dimNull(A) = n$$

Recall that $\text{rank}(A) = dim\text{Col}(A)$.

The column space tells us whether the equation has a solution.

The null space tells us how large the solution set is.

Number of solutions to a system of equations

- If $\text{rank}(A) = m$, the number of rows, then $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$.
- If $\text{rank}(A) < m$, then $A\mathbf{x} = \mathbf{b}$ will only have a solution for $b \in \text{Col}(A)$.
- If $\text{rank}(A) = n$, the matrix is full column rank, then $\text{Null}(A) = \{\mathbf{0}\}$, and $A\mathbf{x} = \mathbf{b}$ will have *at most* one solution for any $\mathbf{b}$.
- If $\text{rank}(A) < n$, then if $A\mathbf{x} = \mathbf{b}$ has a solution, it will have an affine subspace of solutions of dimension $n - \text{rank}(A)$.

Square Matrices and Invertibility

Let A be an $n \times n$ (square) matrix.

Suppose A is full rank, $\text{rank}(A) = n$.

- Since its rank equals the number of rows, $A\mathbf{x} = \mathbf{b}$ has *at least* one solution for every $\mathbf{b}$.
- Since its rank equals the number of columns, $A\mathbf{x} = \mathbf{b}$ has *at most* one solution for every $\mathbf{b}$.
- Therefore, $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$.

Square Matrices and Invertibility

So consider the function $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(\mathbf{v}) = A\mathbf{v}$.

- if $A\mathbf{x} = \mathbf{b}$ and $A\mathbf{y} = \mathbf{b}$, then it must be that $\mathbf{x} = \mathbf{y}$, since solutions are unique. So f is an injection!
- For every $\mathbf{b} \in \mathbb{R}^n$ there is an $\mathbf{x}$ such that $f(\mathbf{x}) = \mathbf{b}$, so f is a surjection!
- f is a bijection! f must have an inverse function, $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$!

Exercise 3.7

Show that f^{-1} must be a linear mapping.

Square Matrices and Invertibility

If $f^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is a linear mapping, it can be represented by a matrix!

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

This gives us that

$$f^{-1}(f(\mathbf{v})) = A^{-1}A\mathbf{v} = \mathbf{v}$$
$$f(f^{-1}(\mathbf{v})) = AA^{-1}\mathbf{v} = \mathbf{v}$$

Which must mean that $A^{-1}A = AA^{-1} = I$.

The Determinant

Let A be an $n \times n$ matrix. Let A_{ij} be the $(n-1) \times (n-1)$ submatrix obtained by deleting row i and column j from A.

The (i,j)th **minor** of A is the scalar

$$M_{ij} = det\,(A_{ij})$$

The (i,j)th **cofactor** of A is the scalar

$$C_{ij} = (-1)^{i+j} M_{ij}$$

The **determinant** of A is given by

$$det(A) = \sum_{i=1}^{n} a_{1i} C_{1i}$$

Observe that the definition of the determinant is recursive. Computing a determinant requires that we compute the determinants of submatrices. This is resolved by defining the determinant of a scalar to be

$$\det(a_{ij}) = a_{ij}$$

The Determinant

Example

Compute the determinant of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Expand along the top row, find the minors and cofactors.

$$M_{11} = d \qquad\qquad M_{12} = c$$
$$C_{11} = (-1)^2 d = d \qquad\qquad C_{12} = (-1)^3 c = -c$$

$$\det(A) = a_{11}C_{11} + a_{12}C_{12} = ad - bc$$

The Determinant

Let A be an $n \times n$ matrix.

Theorem

A is invertible iff $\det(A) \neq 0$.

Let the matrix of cofactors be C,

$$C = \begin{pmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & \ddots & \vdots \\ c_{n,1} & \cdots & c_{n,n} \end{pmatrix}$$

The **adjugate** (elsewhere called adjoint) of A is

$$adj(A) = C^T$$

The Inverse

Theorem

Let A be an invertible matrix. Then

$$A^{-1} = \frac{1}{\det(A)}\, adj(A)$$

Learning Outcomes

You should be able to:

- Transform an augmented matrix into row echelon and reduced row echelon form.
- Solve a system of equations.
- Determine the number of solutions a system of linear equations has given appropriate information.
- Apply the axioms of a vector space in proofs.
- Apply the definitions of a subspace and the span operator.
- Find a basis for the column space of a given matrix.
- Calculate the determinant of a matrix of size 2×2 and 3×3.