1. Inhomogeneous equations:

[we should briefly mention the relationship between homogeneous and inhomogeneous equations:]

Consider general system:

\[ Ax = b. \]  \hspace{1cm} (1)

Suppose \( p \) is a particular solution of (1), so \( Ap = b \). If \( x \) is any other solution of (1), we still have \( Ax = b \). Subtracting the two equations:

\[ Ax - Ap = 0 \quad \Rightarrow \quad A(x - p) = 0. \]

So \( v_h = x - p \) satisfies the homogeneous equation. Generally:

**Theorem 1:** If \( p \) is a particular solution of (1), then for any other solution \( x \), we have that \( v_h = x - p \) solves the homogeneous equation (i.e., with \( b = 0 \)). Thus every solution \( x \) of (1) can be written \( x = p + v_h \), where \( v_h \) is a solution of the homogeneous equation.

2. Application: Network flows

Traffic pattern at Drummond Square:
Quantities in cars/min. What are the flows on the inside streets? One equation for each node:

\[ x_1 - x_3 - x_4 = 40 \]
\[ -x_1 - x_2 = -200 \]
\[ x_2 + x_3 - x_5 = 100 \]
\[ x_4 + x_5 = 60 \]

\[
\begin{bmatrix}
1 & 0 & -1 & -1 & 0 & | & 40 \\
-1 & -1 & 0 & 0 & 0 & | & -200 \\
0 & 1 & 1 & 0 & -1 & | & 100 \\
0 & 0 & 0 & 1 & 1 & | & 60 \\
\end{bmatrix}
\]
\[
\begin{bmatrix}
1 & 0 & -1 & -1 & -1 & 0 & | & 40 \\
0 & -1 & -1 & -1 & 0 & | & -160 \\
0 & 1 & 1 & 0 & -1 & | & 100 \\
0 & 0 & 0 & 1 & 1 & | & 60 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & -1 & -1 & 0 & | & 40 \\
0 & 1 & 1 & 1 & 0 & | & 160 \\
0 & 1 & 0 & -1 & 0 & | & 100 \\
0 & 0 & 0 & 1 & 1 & | & 60 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & -1 & -1 & 0 & | & 40 \\
0 & 1 & 1 & 1 & 0 & | & 160 \\
0 & 0 & 0 & -1 & -1 & | & -60 \\
0 & 0 & 0 & 1 & 1 & | & 60 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & -1 & -1 & 0 & | & 40 \\
0 & 1 & 1 & 1 & 0 & | & 160 \\
0 & 0 & 0 & 1 & 1 & | & 60 \\
0 & 0 & 0 & 1 & 1 & | & 60 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & -1 & 0 & 1 & | & 100 \\
0 & 1 & 1 & 0 & -1 & | & 100 \\
0 & 0 & 0 & 1 & 1 & | & 60 \\
0 & 0 & 0 & 0 & 0 & | & 0 \\
\end{bmatrix}
\]

So:

\[x_1 = 100 + x_3 - x_5\]

\[x_2 = 100 - x_3 + x_5\]

\[x_4 = 60 - x_5,\]

where \(x_3, x_5\) are free.
Constraint: if for example all flows have to be positive; then we require \( x_i \geq 0 \) for all \( i \). Therefore:

\[
x_3, x_5 \geq 0
\]

\[
-100 \leq x_3 - x_5 \leq 100
\]

\[
x_5 \leq 60
\]

This corresponds to a region in the \( x_3, x_5 \) plane - can be plotted if desired.

If they closed off road \( x_3 \) and \( x_5 \), then we have \( x_3 = x_5 = 0 \), so that

\[
x_1 = 100, \ x_2 = 100, \ x_4 = 60
\]

note that then traffic flow becomes uniquely determined.

**Definition 1:**
A collection of vectors \( \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \) is *linearly independent* if no vector in the collection is a linear combination of the others.

Equivalently,

**Definition 2:** A collection of vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_n \) is *linearly independent* if the only way we can have \( c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_n \mathbf{v}_n = \mathbf{0} \) is if all of the \( c_i = 0 \).

**Equivalence of the definitions:**

Def 1 \( \Rightarrow \) Def 2

If no vector is a linear combination of the others, then if

\[
c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_n \mathbf{v}_n = \mathbf{0}
\]

we will show that \( c_1, \ldots, c_n \) have also to be \( \mathbf{0} \).

*Proof:* Suppose not (for contradiction). Without loss of generality, assume \( c_1 \neq 0 \) (proof works same way otherwise). Then we have:

\[
\mathbf{v}_1 = -\frac{c_2}{c_1} \mathbf{v}_2 - \ldots - \frac{c_n}{c_1} \mathbf{v}_n,
\]
contradicting that no vector is a combination of the others. Thus the \( c_i \)
all have to be 0 as desired. □

Note: If \( S_2 \) is a collection of vectors and \( S_1 \) is a subcollection of \( S_2 \), then

If \( S_2 \) is linearly independent

\[ \Rightarrow \text{no vector in } S_2 \text{ is a linear combination of the others} \]

\[ \Rightarrow \text{no vector in } S_1 \text{ is a linear combination of the others (since every vector in } S_1 \text{ is also in } S_2 \)\]

\[ \Rightarrow S_1 \text{ is linearly independent.} \]

Logically equivalent [contrapositive]

If \( S_1 \) is linearly dependent (i.e., not independent)

\[ \Rightarrow S_2 \text{ is linearly dependent} \]

[These are stated more formally in the book as theorems.]

**Theorem 2:** Let \( S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \) be a collection of vectors in \( \mathbb{R}^d \). Then \( S \) is linearly dependent if and only if one of the vectors \( \mathbf{v}_i \) is a linear combination of the previous ones \( \mathbf{v}_1, \ldots, \mathbf{v}_{i-1} \).

**Proof:** (\( \Rightarrow \)) If \( S \) is linearly dependent, then there is a set of constants \( c_i \) not all 0 such that

\[ c_1 \mathbf{v}_1 + \ldots + c_n \mathbf{v}_n = \mathbf{0}. \]

Let \( c_k \) be the last non-zero coefficient. Then the rest of the coefficients are zero, and

\[ c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \ldots + c_{k-1} \mathbf{v}_{k-1} + c_k \mathbf{v}_k = \mathbf{0} \]

\[ \Rightarrow \mathbf{v}_k = -c_1/c_k \mathbf{v}_1 - c_2/c_k \mathbf{v}_2 - \ldots - c_{k-1}/c_k \mathbf{v}_{k-1} \]

i.e. one of the vectors is a linear combination of the previous ones.

(\( \Leftarrow \)) Obvious.
3. Checking for linear independence:

Example 2: Consider the vectors

\[ v_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \]

Are they linearly independent?

\[ c_1v_1 + c_2v_2 + c_3v_3 = 0 \]

\[ \Rightarrow \]

\[ c_1 + c_2 - c_3 = 0 \]

\[ c_1 - c_2 = 0 \]

Reduced matrix:

\[ \begin{bmatrix} 1 & 1 & -1 & | & 0 \\ 1 & -1 & 0 & | & 0 \end{bmatrix} \]

\[ \rightarrow \begin{bmatrix} 1 & 1 & -1 & | & 0 \\ 0 & -2 & 1 & | & 0 \end{bmatrix} \]

\[ \rightarrow \begin{bmatrix} 1 & 1 & -1 & | & 0 \\ 0 & 1 & -1/2 & | & 0 \end{bmatrix} \]

Conclude: there are free variables. By theorem on homogeneous equations there is a nontrivial solution, so \( c_i \) need not be 0.

Thus not all \( c_i \) must be 0 \( \Rightarrow \) not linearly independent.

[note that if number of vectors is greater than the size of the vectors, this will always happen].

More generally thus:
Theorem 3: In $\mathbb{R}^n$, if we have more than $n$ vectors, they cannot be linearly independent.

From above we have:

Algorithm: To check whether vectors are linearly independent, form a matrix with them as columns, and row reduce.
   (a) If reduced matrix has free variables (i.e., $\exists$ a non-pivot column), then they are not independent.
   (b) If there are no free variables (i.e., there are no nonpivot columns), they are independent.

[material from section 1.8 starts here]
Lecture 4

1. A note on linear transformations: Single variable

Simple idea: \( y = ax = T(x) \), where \( x, y \in \mathbb{R}^1 \)

[e.g., \( y = \) distance travelled by light, \( x = \) time, \( a = \) speed of light]

This is a linear function.

Properties:

\[
T(x + y) = T(x) + T(y)
\]

\[
T(cx) = cT(x)
\]

2. Note on linear functions: more variables

**NOTE:** Linear transformations are a simple and often accurate way of describing the transition from input to output. Here is an example.

In a chemical reaction three chemicals \( C_1, C_2, C_3 \) are mixed in weights \( x_1, x_2, x_3 \) grams. After one minute the amounts of the chemicals are

\[
y_1, y_2, y_3, \text{ where} \]

\[
y_1 = .5x_1 + .25x_2 + .25x_3
\]

\[
y_2 = .25x_1 + .25x_2 + .5x_3
\]

\[
y_3 = .25x_1 + .5x_2 + .25x_3
\]

Notice that variables \( y \) are a linear function of variables \( x \)

We write \( y = T(x) \) if above true.

Of course we have matrix \( A \) such that \( y = Ax = T(x) \).

There is a linear relationship between \( y \) and \( x \).

[Can check properties of linear relationship using \( T(x) = A(x) \)]
A linear law relates inputs and outputs of chemical reaction]

After 2 minutes the amounts are $z = A^2x$. After $k$ minutes the amounts are $A^kx$.

Have linear transformation from $y$ to $x$. A linear transformation comes from a linear system of equations in this case.

**Definition 1**: A transformation (or function or mapping) $T$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $x$ in $\mathbb{R}^n$ another vector $T(x)$. The set $\mathbb{R}^n$ is called the domain of $T$ and the set $\mathbb{R}^m$ is called the codomain of $T$.

Notation: $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

**Definition 2**: The vector $T(x)$ is called the image of $x$. The set of all images $T(x)$ is called the range of $T$.

**Observation**: A general linear transformation from a vector $x$ to a vector $y$ can always be written out as a system of equations where the $x_i$ are linear combinations of the $y_i$ (no squares, etc.); to be proved later.

**Example 1**: Multiplication by any matrix is linear in $\mathbb{R}^n$ (check properties).

$$A(x + y) = Ax + Ay$$

$$A(cx) = cAx.$$
Figure: $A$ viewed as a linear transformation: $Ax = T(x)$.

**Fact** (to be proved later): *For every linear transformation $T(x)$, there exists a unique matrix $A$ such that $T(x) = Ax$."

A is called the *standard matrix* of $T(x)$.

**Definition 3:**

**Example:** American Automotive market - three major kinds of cars: American, Japanese, European. Every 5 years (average time between car purchases),

- American market keeps 80% of previous buyers, and gets 20% of Japanese market, and 10% of European market
- Japanese market keeps 70% of previous buyers, gets 10% of American market, and gets 30% of European market
- European market keeps 60% of previous buyers, and gets 10% of American and 10% of Japanese market

Assuming non-steady state model, what are the percentage shares of the market?

$A : x_1 =$ number of American car owners (in millions)

$x_2 =$ number of Japanese car owners

$x_3 =$ number of European car owners

Then:

$y_1 =$ number of American cars in 5 years

$= .8x_1 + .2x_2 + .1x_3$

$y_2 =$ number of Japanese cars in 5 years

$= .1x_1 + .7x_2 + .3x_3$

$y_3 =$ number of European cars in 5 years

$= .1x_1 + .1x_2 + .6x_3$
$$y = \begin{bmatrix} .8 & .2 & .1 \\ .1 & .7 & .3 \\ .1 & .1 & .6 \end{bmatrix} x$$

again the input-output mapping is linear for this system.

Can again iterate system by squaring matrix, etc.

Domain: $\mathbb{R}^3$. Codomain: $\mathbb{R}^3$.

[later on vectors will have more general meanings; in that case we have:]

Example: Differentiation of polynomials is also a linear input-output map.

$$L(P(x)) = \frac{d}{dx} P(x)$$ is linear:

$$L(P_1(x) + P_2(x)) = \frac{d}{dx} (P_1(x) + P_2(x)) = \frac{d}{dx} P_1(x) + \frac{d}{dx} P_2(x)$$

$$= L(P_1(x)) + L(P_2(x))$$

$$L(cP(x)) = \frac{d}{dx} (cP(x)) = c \frac{d}{dx} P(x) = cL(P(x)).$$

3. Matrix of a Linear Transformation

Def. 4: The unit vector $e_j$ in $\mathbb{R}^n$ is
\[ \mathbf{e}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \]

with 1 in the \( j^{th} \) entry, and zeroes elsewhere.

**Theorem 4:** Let \( T : \mathbb{R}^n \to \mathbb{R}^m \) be a linear transformation. Then there exists a unique matrix \( A \) such that
\[ T(\mathbf{x}) = A\mathbf{x} \]
\[ \forall \mathbf{x} \in \mathbb{R}^n. \]

Furthermore, the \( j^{th} \) column of \( A \) is just the vector \( T(\mathbf{e}_j) \), so that
\[ A = [T(\mathbf{e}_1) \; T(\mathbf{e}_2) \; \ldots \; T(\mathbf{e}_n)] \]

**Proof:** The short proof of this is in the book.

[We now cover material from section 2.1]

**Example:** Do matrix for rotation in two dimensions.

In space shuttle, a particle for experimental accelerometer is moving at 5 meters per second in direction \( \mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \). What is the velocity vector for someone in a rotated coordinate system by clockwise angle \( \theta \)?

[How does someone rotate with respect to what you see in that vector?]

Note that vector will be rotated counterclockwise by \( \theta \).
First: counterclockwise rotation by $\pi/4 = 45^\circ$:

\[
\begin{array}{c}
\gamma \\
\downarrow \\
\vec{v} \\
\downarrow \\
\vec{x}
\end{array}
\]

Note that if $T$ is the operation which rotates vectors counterclockwise by $45^\circ$, then:

\[
T(\vec{v}_1 + \vec{v}_2) = \text{rotation of } (\vec{v}_1 + \vec{v}_2)
= \text{rotation of } \vec{v}_1 + \text{rotation of } \vec{v}_2
= T(\vec{v}_1) + T(\vec{v}_2)
\]

Similarly, show that $T(c\vec{v}) = cT(\vec{v})$;

Conclude: $T$ is linear.

Use fact that any linear transformation on a finite dimensional space is given by a matrix. Thus there exists a matrix $A$ such that

\[
T(\vec{x}) = A\vec{x} = \text{rotation of } \vec{x} \text{ by } 45^\circ \text{ counterclockwise}
\]

To find $A$, note that if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, then
\[
\begin{bmatrix}
a & b \\
c & d
\end{bmatrix}
\begin{bmatrix} 
1 \\
0
\end{bmatrix} =
\begin{bmatrix}
\frac{\sqrt{2}}{2} \\
\frac{\sqrt{2}}{2}
\end{bmatrix}
\]

\[\Rightarrow \quad a = c = \frac{\sqrt{2}}{2}\]

Similarly, get

\[
\begin{bmatrix}
a & b \\
c & d
\end{bmatrix}
\begin{bmatrix} 
0 \\
1
\end{bmatrix} =
\begin{bmatrix}
-\frac{\sqrt{2}}{2} \\
\frac{\sqrt{2}}{2}
\end{bmatrix}
\]

\[\Rightarrow \quad b = -\frac{\sqrt{2}}{2} \quad d = \frac{\sqrt{2}}{2}\]

\[\Rightarrow \quad A =
\begin{bmatrix}
\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\
\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2}
\end{bmatrix}
\]

Generally if \( A \) rotates by \( \theta \), can let:

\[A = \begin{bmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{bmatrix}\]

where the entries are determined by action of \( A \) on \( e_1 \) and \( e_2 \)

**Preliminary:** We will see later that if \( I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \) = identity matrix, then for any matrix \( A \), we have \( AI = IA = A \).

**Ex:** Look at the shear transformation (to be discussed in more detail later) from graphics:

Operation \( P = \) perspective operation
What is the transformation which takes picture 1 to picture 2?

**Answer:** It is a linear transformation $T$ (called $P$ above)

Just compute $T(e_1)$ and $T(e_2)$, and all else is determined. Assume that

$$T(e_1) = \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}; \quad T(e_2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$ 

Thus the matrix for this transformation is $A$ where:

$$A = A_1 \begin{bmatrix} 1/2 & 1 \\ 0 & 1 \end{bmatrix} = [T(e_1) \ T(e_2)].$$

[this is how we transform points using linear transformations]

Recall the linear transformation $T(v)$ on the vector space $\mathbb{R}^2$ above. How do we find the matrix $A$ of $T$? In above example,

$$A = [T(e_1) \ T(e_2)]$$

This is true in general with the same proof:
Example: Let $T$ be the linear transformation which rotates vectors in $\mathbb{R}^3$ by 90 degrees counterclockwise about the $x$ axis. What is the matrix $A$ of $T$?

\[
T(e_1) = e_1 \\
T(e_2) = e_3 \\
T(e_3) = -e_2
\]

Therefore

\[
A = [T(e_1) \ T(e_2) \ T(e_3)] = [e_1 \ e_3 - e_2] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}
\]

Thus rotation of vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ counterclockwise about $x$ axis gives:

\[
\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.
\]
4. Geometric Linear Transformations

[Material through the end of this lecture will be elaborated on the board]

[The above rotation moves vectors; equivalently it moves points]

Consider the rotation with matrix \( A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \)

[90° counterclockwise rotation]

Image of the unit square:

[What is the image of unit square under the above rotation?]

Other transformations:

1. \( A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \)

[reflection about x axis]

[Note image of unit square under the rotation]

2. \( A = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \)

[Stretch by factor k horizontally on the unit square; note effect on e_1 and e_2]

3. \( A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \)

[Shear transformation; again note effect on unit vectors]

4. \( A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \)
5. Domain and Range

Recall the definitions of the domain and range of a LT $T$.

[We first want to consider some intuition behind the notions of one to one and onto]

[Now the formal definitions]

**Def. 5:**
A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **one to one** if $T(x) = T(y) \Rightarrow x = y$

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is **onto** if $\forall y \in \mathbb{R}^m \exists x \in \mathbb{R}^n$ s.t.

$$T(x) = y.$$  

**Theorem 5:** $T : \mathbb{R}^n \to \mathbb{R}^m$ is one to one iff the equation

$$T(x) = 0$$  

has only the trivial solution $x = 0$.

*Equivalently, $T$ is one to one if its matrix $A$ has a pivot position in every column.*

*Proof:* In book

**Theorem 6:** The mapping $T$ is onto iff the equation

$$T(x) = b$$  

has a solution for all $b$.

*Equivalently, $T$ is onto if its matrix $A$ has a pivot entry in every row.*

*Proof:* In book
Example 2:

\[ A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 1 & 0 \end{bmatrix}. \]