L05LinearIndependence

September 18, 2015

1   Linear Independence

In the last lecture, we learned that $A\mathbf{x} = \mathbf{b}$ is consistent if and only if $\mathbf{b}$ lies in the span of the columns of A.

As an example, we saw for the following matrix A, $A\mathbf{x} = \mathbf{b}$ is not consistent for all $\mathbf{b}$:

$$A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$$

We realized that was because the span of A's columns is not $\mathbb{R}^3$, but rather a plane lying within $\mathbb{R}^3$. So, when $\mathbf{b}$ does not lie in that plane, then $A\mathbf{x} = \mathbf{b}$ is not consistent and has no solution.
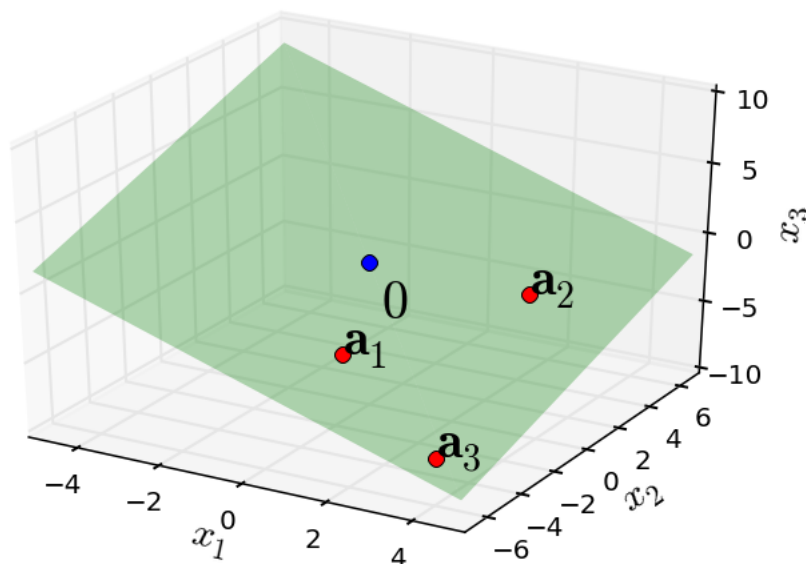
As a reminder, here is the picture of the span of the columns of A.

Clearly, $\mathbf{0}, \mathbf{a_1}, \mathbf{a_2}$, and $\mathbf{a_3}$ have a particular relationship, namely, they all lie within the same plane – even though the vectors are in $\mathbb{R}^3$. Today we will talk about how to define this relationship precisely for vectors of arbitrary dimension, that is, vectors in $\mathbb{R}^n$.

The relationship between these vectors will be called *linear dependence.*

Before stating the definition, let's get a sense intuitively of what we want to capture. We make this observation: the plane defined by $\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}$ happens to include the origin. That's one way of capturing the special relationship among $\mathbf{a_1}, \mathbf{a_2}, \mathbf{a_3}$.

Here is the formal definition:

A set of vectors $\{\mathbf{v_1}, ..., \mathbf{v_p}\}$ all of which are in $\mathbb{R}^n$ is said to be *linearly dependent* if there exist weights $\{c_1, ..., c_p\}$, **not all zero,** such that

$$c_1\mathbf{v_1} + ... + c_p\mathbf{v_p} = 0.$$

Conversely, the set $\{\mathbf{v_1}, ..., \mathbf{v_p}\}$ is said is said to be *linearly independent* if the vector equation

$$c_1\mathbf{v_1} + ... + c_p\mathbf{v_p} = 0.$$

has only the trivial solution $c_1 = 0, ..., c_p = 0$.

This is called a *linear dependence relation* among $\{\mathbf{v_1}, ..., \mathbf{v_p}\}$ when the weights are not all zero. A set of vectors is linearly dependent if and only if it is not linearly independent.

1.1   Example.

Let $\mathbf{v_1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v_2} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}$, and $\mathbf{v_3} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$.

Let's determine (a) if the set $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ is linearly independent, and (b) if not, a linear dependence relation among them.

(a). Are $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ linearly independent?

We must determine if there is a nontrivial solution of the vector equation:

$$x_1\mathbf{v_1} + x_2\mathbf{v_2} + x_3\mathbf{v_3} = 0.$$

Let's row reduce the augmented matrix:

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We can see that x_1 and x_2 are basic variables, and x_3 is free. Each nonzero value of x_3 determines a nontrivial solution of the vector equation. So $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ are linearly dependent.

(b). To find the linear dependence relation among $\mathbf{v_1}, \mathbf{v_2}$, and $\mathbf{v_3}$, we continue the row reduction to obtain the reduced echelon form:

$$\begin{bmatrix} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Which denotes the system of equations:

$$\begin{aligned} x_1 \quad\quad -2x_3 &= 0 \\ x_2 \quad +x_3 &= 0 \\ 0 &= 0 \end{aligned}$$

So $x_1 = 2x_3, x_2 = -x_3$, and x_3 is free.

We can choose any nonzero value for x_3 – say, $x_3 = 5$. Then $x_1 = 10$ and $x_2 = -5$. This gives us the solution:

$$10\mathbf{v_1} - 5\mathbf{v_2} + 5\mathbf{v_3} = \mathbf{0}.$$

This is one (out of infinitely many) linear dependence relations among $\mathbf{v_1}, \mathbf{v_2}$, and $\mathbf{v_3}$.

1.2 Linear Independence of Matrix Columns

The columns of a matrix are a set of vectors, so our definition of linear dependence extends naturally to them. In particular if $A = [\mathbf{a_1} \ \ldots \ \mathbf{a_n}]$ then

$$x_1\mathbf{a_1} + \ldots + x_n\mathbf{a_n} = 0$$

can be written as $A\mathbf{x} = \mathbf{0}$.

So each linear dependence relation among the columns of A corresponds to a nontrivial solution of $A\mathbf{x} = \mathbf{0}$.

In other words, the columns of the matrix A are linearly independent if and only if the equation $A\mathbf{x} = \mathbf{0}$ has **only** the trivial solution.

1.3 Another Interpretation of Linear Dependence

Here is another way of thinking about linear dependence.

Theorem. A set $S = \{\mathbf{v_1}, ..., \mathbf{v_p}\}$ of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others.

Proof.

First, let's consider the "if" part:

Assume $\mathbf{v_p} = c_1\mathbf{v_1} + \cdots + c_{p-1}\mathbf{v_{p-1}}$. Then clearly

$$c_1\mathbf{v_1} + \cdots + c_{p-1}\mathbf{v_{p-1}} - \mathbf{v_p} = 0,$$

and not all the coefficients are zero (the coefficient of $\mathbf{v_p}$ is -1). Thus, the vectors are linearly dependent.

Now, we consider the "only if" part:

Assume S is linearly dependent. Then $c_1\mathbf{v_1} + \cdots + c_p\mathbf{v_p} = 0$ and at least one of the c_i is nonzero. Pick one of the nonzero c_i, and rearranging, we get:

$$\mathbf{v_i} = -(c_1/c_i)\mathbf{v_1} + \cdots + -(c_p/c_i)\mathbf{v_p}$$

Thus, there is at least one vector that is a linear combination of the others.

1.4 Question Time! Q5.1

1.5 Example.

Let us return to the motivation at the start of the lecture, and formalize the connection between spanning sets and linear dependence.

Let $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$. Let's (a) describe the set spanned by $\mathbf{u}$ and $\mathbf{v}$ and (b) explain why a vector $\mathbf{w}$ is in Span$\{\mathbf{u}, \mathbf{v}\}$ if and only if $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is linearly dependent.

Solution. The vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly independent so they span a plane in $\mathbb{R}^3$. In fact, since $x_3 = 0$ in both vectors, they span the x_1x_2 plane.
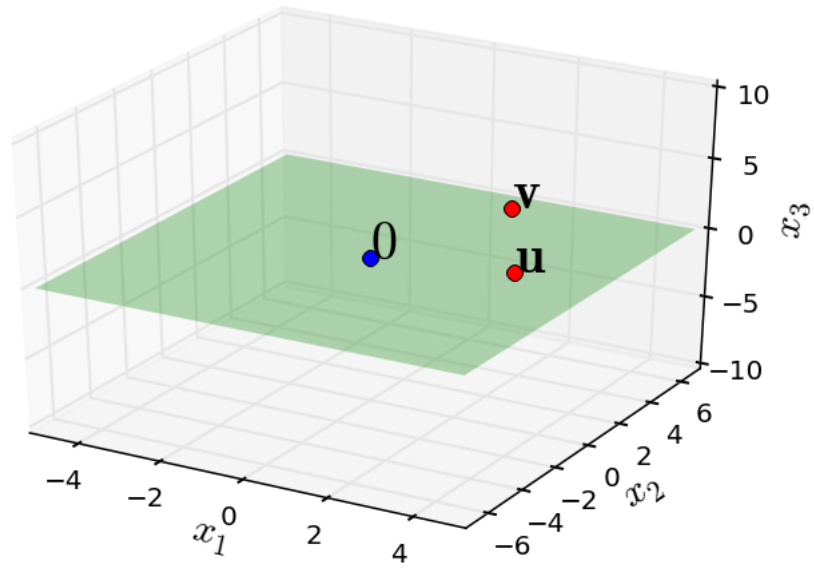
Now, if **w** is in Span{**u**, **v**}, then **w** is a linear combination of **u** and **v**. So then {**u**, **v**, **w**} is linearly dependent (by the Theorem we just proved).

And conversely, if {**u**, **v**, **w**} is linearly dependent, then there exist c_1, c_2, and c_3, not all zero, such that

$$c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = \mathbf{0}.$$

We can see that **u** and **v** are linearly independent, so the only way for $c_1\mathbf{u} + c_2\mathbf{v}$ to be zero is if $c_1 = c_2 = 0$. So $c_1\mathbf{u} + c_2\mathbf{v}$ must be nonzero, and that means c_3 must be different from zero. So

$$\mathbf{w} = -(c_1/c_3)\mathbf{u} - (c_2/c_3)\mathbf{v},$$

that is, **w** is a linear combination of **u** and **v**, and therefore lies in Span{**u**, **v**}.
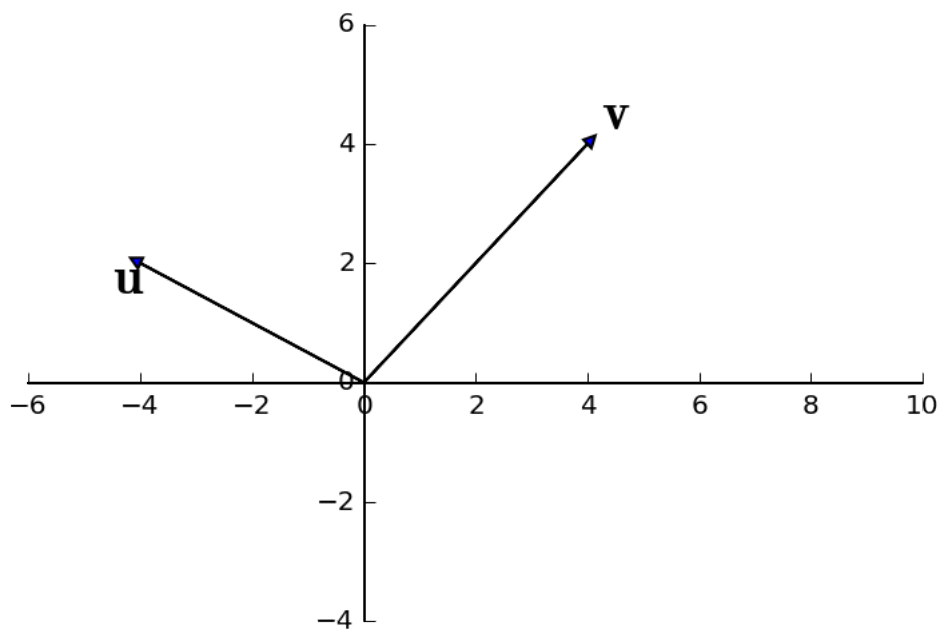
Linearly Dependent
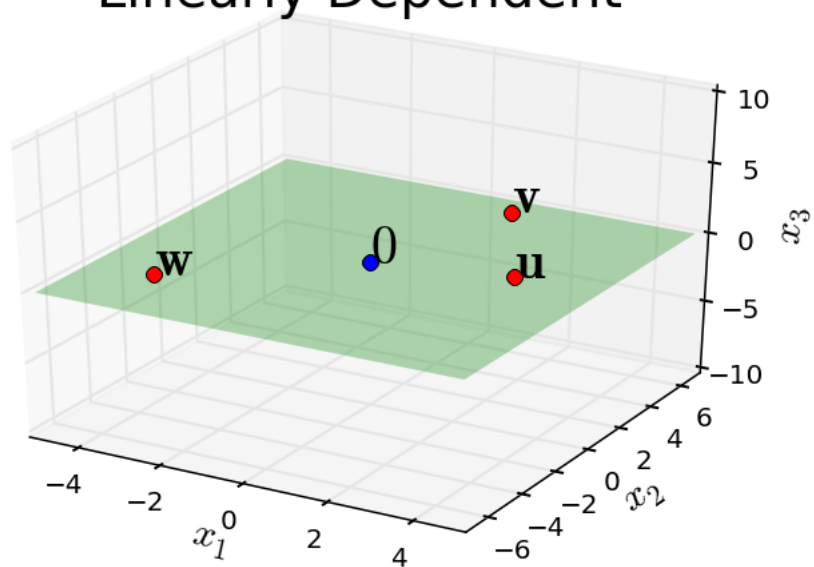
Linearly Independent
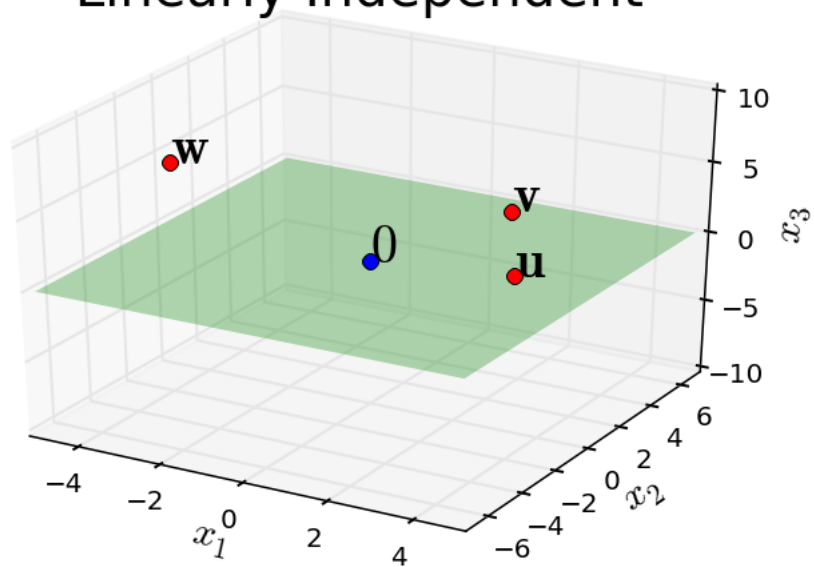
1.6 Linear Dependence in $\mathbb{R}^2$

Let's try to interpret linear dependence directly in terms of vector sums.

u and **v** are independent because there is no nozero combination that yields the origin.

A nonzero combination of **u** and **v** geometrically means moving in the direction of **u** or **v** or both. There is no way to move in the direction of **u** and then in the direction of **v** and arrive back at the origin.

Now let's add another vector **w**:
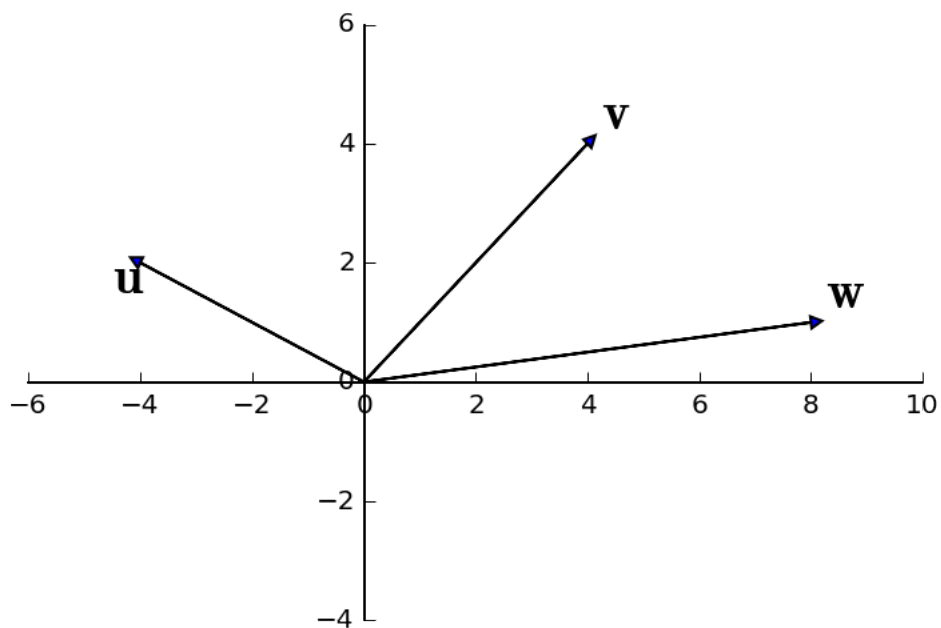
The set $\mathbf{u}, \mathbf{v}, \mathbf{w}$ is linearly dependent. There are nonzero moves along the three directions that can bring you back to the origin:
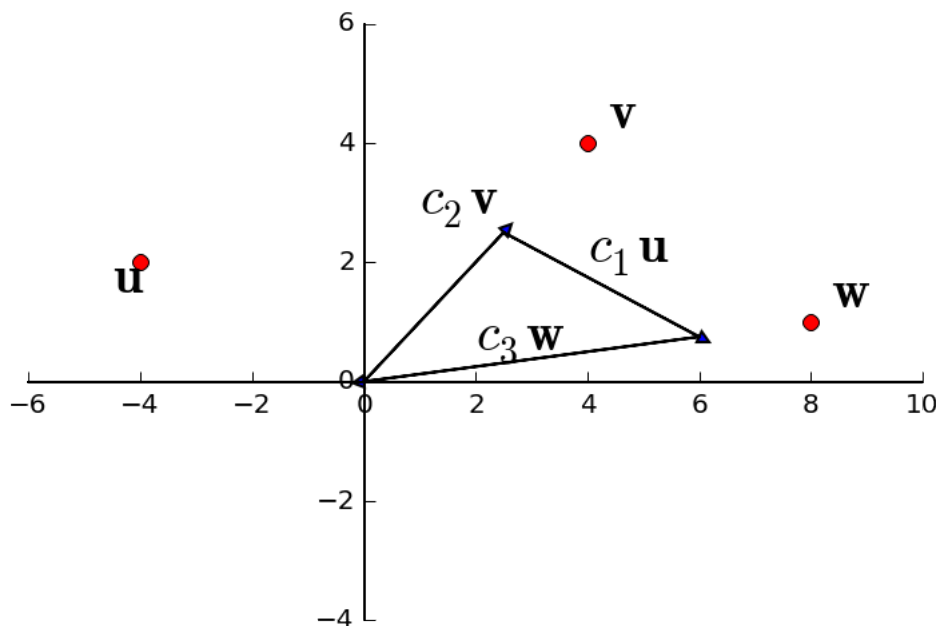
This is a geometric interpretation of the equation

$$c_1\mathbf{u} + c_2\mathbf{v} + c_3\mathbf{w} = 0.$$

1.7 Question Time! Q5.2

The last example suggested that *any* three vectors in $\mathbb{R}^2$ are linearly dependent. This is true, and furthermore, we can generalize to $\mathbb{R}^n$.

Theorem. If a set contains more vectors than there are entries in each vector, then the set is linearly dependent. That is, any set $\{\mathbf{v_1}, \ldots, \mathbf{v_p}\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

Proof. Let $A = [\mathbf{v_1} \ \ldots \ \mathbf{v_p}]$. Then A is $n \times p$, and the equation $A\mathbf{x} = \mathbf{0}$ corresponds to a system of n equations in p unknowns. If $p > n$, there are more variables than equations, so there must be a free variable. Hence $A\mathbf{x} = \mathbf{0}$ has a nontrivial solution, and the columns of A are linearly dependent.

Example.

The vectors $\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ -1 \end{bmatrix}$, and $\begin{bmatrix} -2 \\ 2 \end{bmatrix}$ are linearly dependent because these vectors live in $\mathbb{R}^2$ and there are 3 of them.

Theorem. If a set $S = \{\mathbf{v_1}, ..., \mathbf{v_p}\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

Proof. Let's say $\mathbf{v_1} = 0$. Then $1\mathbf{v_1} + 0\mathbf{v_2} + \cdots + 0\mathbf{v_p} = 0$. The coefficients are not all zero, so the set is linearly dependent.

## 1.8 Question Time Q5.3

1.9 Extended Example: Network Flow

Systems of linear equations arise when considering flow through a network.

A *network* is a set of *nodes* and *links*. Links connect nodes.

Be aware that there is another set of terminology that is used more often in theoretical computer science and mathematics: A *graph*, which consists of *vertices* and *edges*. A network and a graph are exactly the same thing.

Here are some examples of networks:
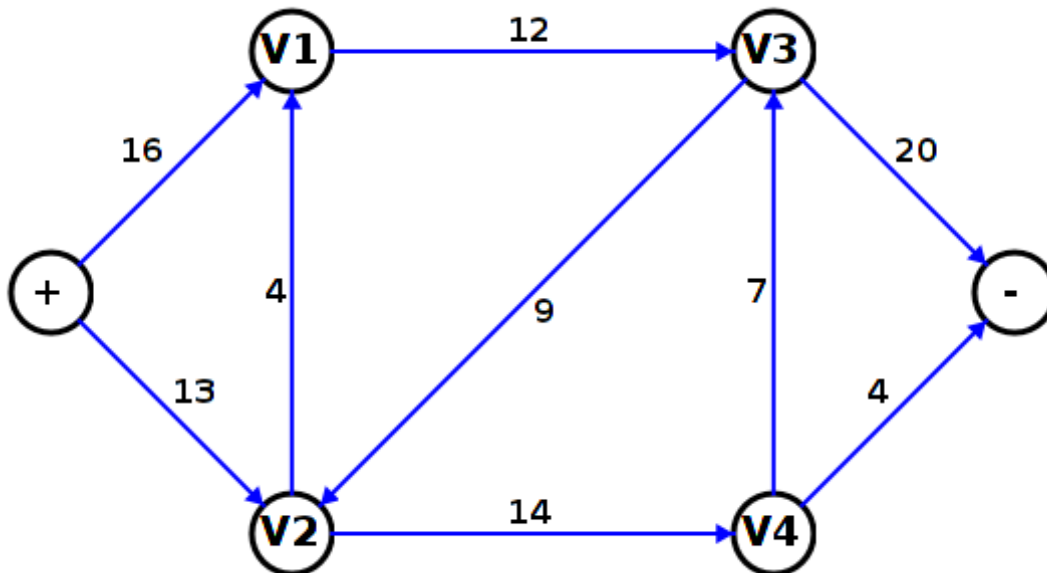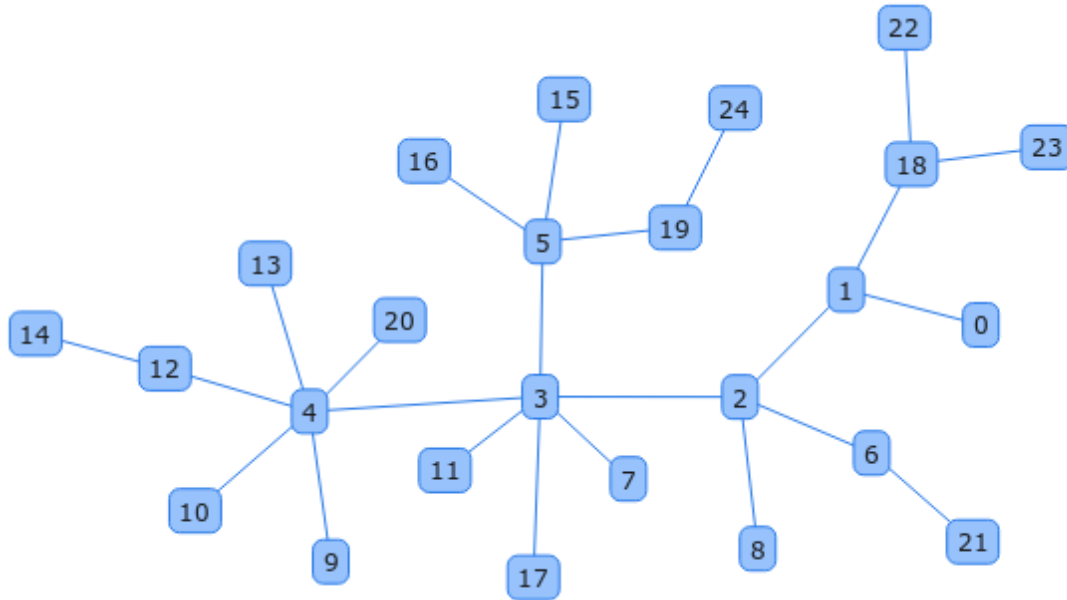
Many times we are interested in *flow* through a network. Flows can represent movement from place to place. For example, consider this map of the MBTA:
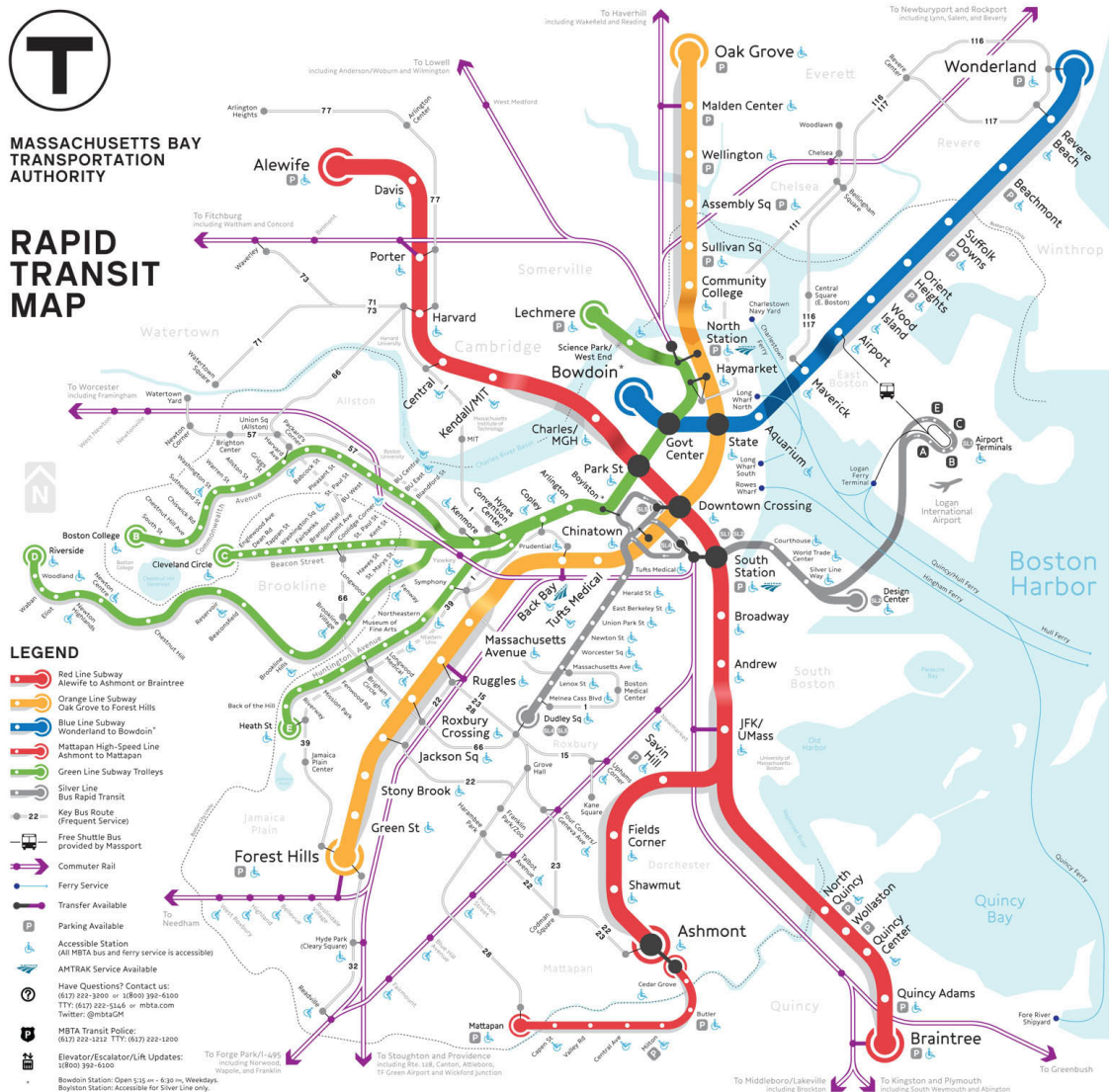
We can think of T stops as nodes and rail connections as links. Flow corresponds to the movement of subway cars from station to station.

Here is another example: this is a representation of the *Abilene* data network, which is used by universities to exchange data traffic (packets) within the US:

Networks usually obey the rule of "flow balance" or "conservation of flow." This simply means that the amount of flow going into a node equals the amount coming out. For example, the number of packets that enter any node of the Abilene network equals the number that leave that node. This simply reflects that packets are not created or destroyed *within the network.*
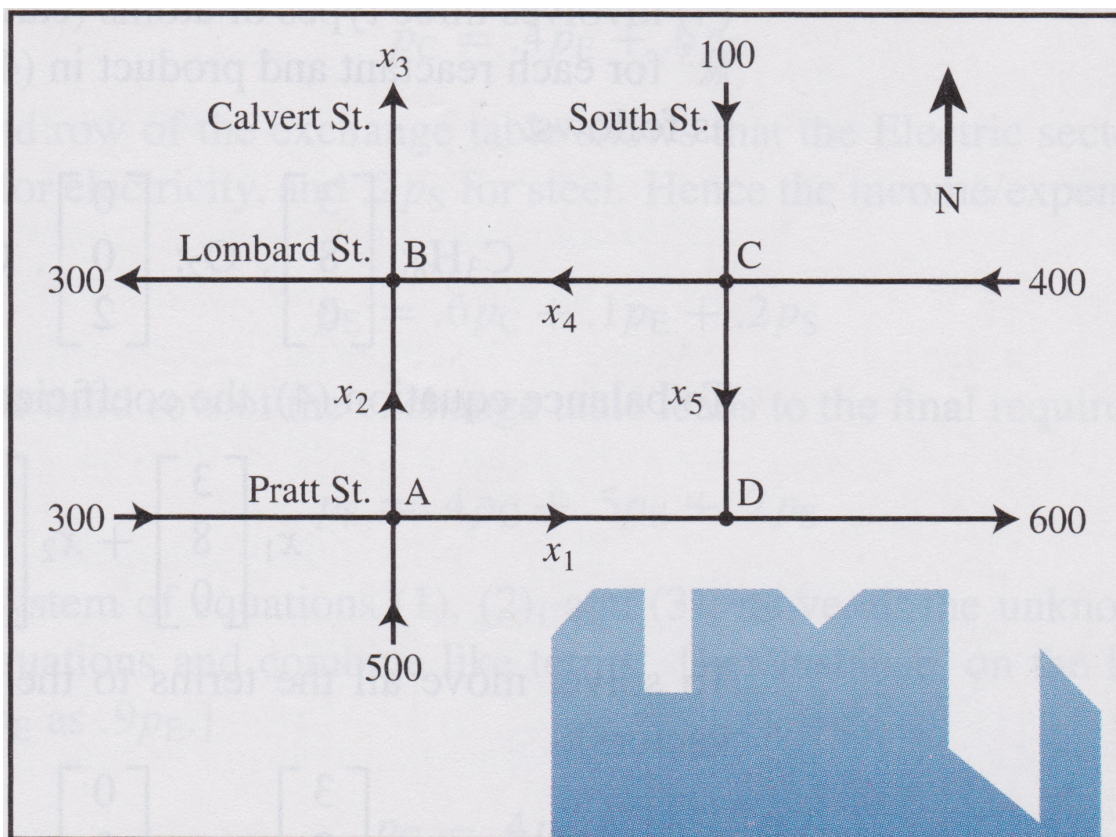
Here is an example (from the text, Section 1.6). The flows are vehicles per hour in downtown Baltimore during rush hour.

Note that some of the traffic flows are measured, and some are not. The unmeasured flows are marked with symbols x_1, x_2, etc.

We'd like to understand the traffic pattern in the city, despite the presence of unmeasured flows. We can do that by using the principle of flow balance.

The key idea is: flow balance dictates that *every node determines a linear equation.* The equation is of the form:

$$\text{Flow in} = \text{Flow out}.$$

Intersection	Flow in	Flow out
A	$300 + 500$	$x_1 + x_2$
B	$x_2 + x_4$	$300 + x_3$
C	$100 + 400$	$x_4 + x_5$
D	$x_1 + x_5$	600
network	x_3	400

The last line indicates that the total flow into the network equations the total flow out of the network, which means that $x_3 = 400$.

This yields the following system of equations:

$$\begin{aligned} x_1 \ +x_2 \qquad\qquad\qquad\qquad &= \ 800 \\ x_2 \ -x_3 \ +x_4 \qquad\qquad &= \ 300 \\ x_4 \ +x_5 &= \ 500 \\ x_1 \qquad\qquad\qquad\quad +x_5 &= \ 600 \\ x_3 \qquad\qquad\qquad\quad &= \ 400 \end{aligned}$$

When we row reduce the associated augmented matrix, the reduced echelon form yields these equations:

$$\begin{aligned} x_1 \qquad\qquad\qquad +x_5 &= \ 600 \\ x_2 \qquad\qquad -x_5 &= \ 200 \\ x_3 \qquad\qquad &= \ 400 \\ x_4 \ +x_5 &= \ 500 \end{aligned}$$

Thus the general flow pattern for the network is described by

$$\begin{cases} x_1 = 600 - x_5 \\ x_2 = 200 + x_5 \\ x_3 = 400 \\ x_4 = 500 - x_5 \\ x_5 \text{ is free} \end{cases}$$
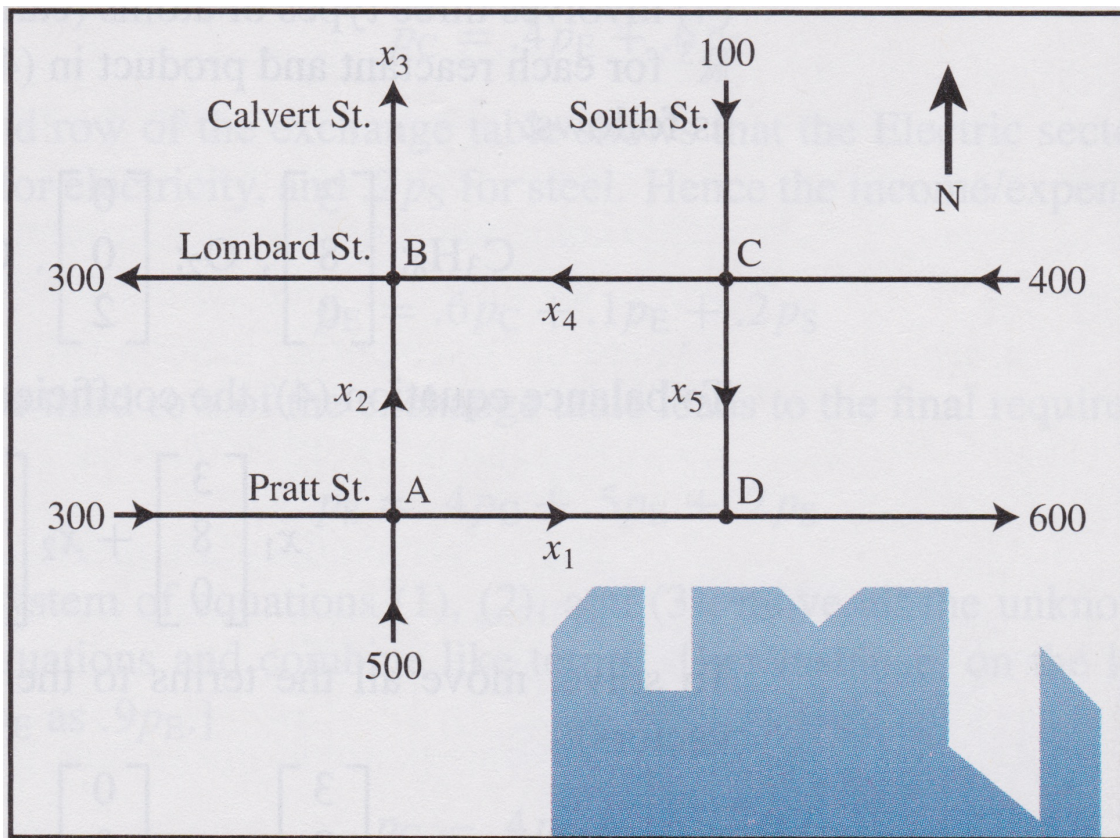
How can we interpret this result? A negative flow corresponds to a flow in the opposite direction assumed in the diagram. Since these streets are one-way, none of the variables here can be negative. So, for example, $x_5 \leq 500$ because x_4 cannot be negative. What else can we infer?