for all i, and there must be only one way to express the vector.

So there's only one way to represent a vector as a weighted sum of ~~an~~ these vectors, and that means we can use those weights as an unambiguous coordinate system. For example with basis $\left\langle \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}0\\2\end{pmatrix} \right\rangle$ we can represent the ~~point~~ vector (or point) $\begin{pmatrix}3\\5\end{pmatrix}$ with the coordinate vector $\begin{pmatrix}3\\1\end{pmatrix}$ since $3\begin{pmatrix}1\\1\end{pmatrix} + 1\begin{pmatrix}0\\2\end{pmatrix} = \begin{pmatrix}3\\5\end{pmatrix}$.

We can in general solve for what these coordinates should be with algebra - for example, for $\begin{pmatrix}4\\5\end{pmatrix}$ we'd have

$$c_1\begin{pmatrix}1\\1\end{pmatrix} + c_2\begin{pmatrix}0\\2\end{pmatrix} = \begin{pmatrix}4\\5\end{pmatrix} \quad \begin{matrix}c_1 = 4\\ c_1 + 2c_2 = 5\end{matrix} \quad 2c_2 = 1 \quad c_2 = \frac{1}{2}$$

We could prove that any basis of the same vector space would have the same # of basis vectors. We can call this the dimension of the space without worrying about which basis we're using. This is also the maximum number of linearly independent vectors we could have.

Rank

When we analyze matrices we'll want to know things like, would equations corresponding to this matrix have a unique solution? The rank is a major tool in this analysis.

We can break a matrix down into vectors in two ways. One is to think of each row as a vector.

$$\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} = \{(1\ 2), (3, 4)\}$$

And the other is to break down by column: $\begin{pmatrix}1\\3\end{pmatrix}, \begin{pmatrix}2\\4\end{pmatrix}$

What would either of these mean to us? In the case of the row representation, notice that in the process of solving a system of equations, we subtract multiples of rows from other rows until we can't anymore.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 3\cdot(1\ 2) = \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix}$$

If all the row vectors are linearly independent, it shouldn't be possible to zero out a row this way, because that would mean the original row could be represented as a linear combo of the others. But if the row vectors aren't linearly independent, then it will be possible to zero out the dependent vectors.

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} - 2\cdot(1\,2) = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

The rank of a matrix is equal to the number of linearly independent row vectors it has. It's the number of nonzero rows that will be left after we perform Gauss-Jordan elimination. It therefore also determines whether there is a unique solution to a system of equations.

Okay, how about the column vector representation? When solving systems of linear equations, we can think of it as determining whether a linear combo of the column vectors can produce the target vector.

$$x_1 + 2x_2 = 5$$
$$3x_1 + 4x_2 = 6 \Rightarrow \left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right]$$
$$\Uparrow$$
$$x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

Whether these vectors are linearly independent matters to the number of solutions as well. But it turns out, the number of independent column vectors is equal to the number of independent row vectors. This number is called the "rank" in both cases.

We can see that both ranks are the same when we find reduced echelon form of the matrix. Doing this doesn't change the rank, since we're just replacing row vectors with other linear combinations of linearly independent vectors. We established that the rank is the number of rows with leading 1's. But now, when everything is fully reduced, the 1's have zeroed out everything else in their column.

eg $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Vectors like $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ with one 1 and the rest 0 belong to the standard basis, so they're ~~orthogonal~~ linearly independent.

If the rank of a matrix with n unknowns is n, then the solution to the associated equations is ~~either~~ unique ~~or nonexistent~~. But when the rank is less, we can say more: the vector space of solutions has n−r dimensions. This is tied to the fact that every "missing" row in the matrix gives another degree of freedom to play with.

In short the following statements (p. 132) are equivalent, if A is $n \times n$:

(1) rank of A is n
(2) A is nonsingular
(3) the rows of A are linearly independent
(4) cols of A are linearly independent
(5) a linear system with matrix of coefficients A has a unique solution

How do we know there is a solution? Because n linearly independent vectors form a basis for the n-dim. space. There must be a solution to

$$x_1 \begin{pmatrix} \cdots \end{pmatrix} + x_2 \begin{pmatrix} \cdots \end{pmatrix} + \dots + x_n \begin{pmatrix} \cdots \end{pmatrix} = \begin{pmatrix} \ \end{pmatrix}$$

vector of constants

because we know the span has n dimensions.

Transpose

We'll occasionally find it convenient to switch how we're thinking of a matrix — treat the column vectors as if they were row vectors & vice versa. Taking the transpose makes this happen. It flips the matrix along its diagonal.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

This will be a handy operation in some more advanced stuff (like least sq's regression)

row vectors $(1\ 2\ 3)$, $(4\ 5\ 6)$, $(7\ 8\ 9)$

col vectors $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$

Week 3

— skip —

This week, we'll emphasize matrices as transformations — put in some points, and they'll be rotated, scaled, or whatever.

Isomorphisms vs Homomorphisms

We can think of matrices as transforming vectors so that they're now functioning in a new space, according to some new coordinates. Some transformations can be thought of as just cosmetic — for example, rotating everything about the origin. If the new space of possible vectors looks basically identical for practical purposes, the transformation is isomorphic. But if something is lost — for example, projecting a 3D scene onto a 2D view — the transformation is only homomorphic.

More specifically, an isomorphism is

- one-to-one and onto: every point gets mapped to just one point, and no point can be hit with two different mappings
- structure preserving: $f(\vec{v}_1 + \vec{v}_2) = f(\vec{v}_1) + f(\vec{v}_2)$ and $f(c\vec{v}) = c f(\vec{v})$

The first point leaves open the possibility we could recover where everything came from and get back to the original space. The second means we could do some math in the new space that reflects how things would have worked before the transformation.

~~If the first pair of requirements isn't satisfied, and the transfor~~

— skip —

A homomorphism is similar, but doesn't require one-to-one and onto. (An isomorphism is a kind of homomorphism.)

A kind of interesting result is that all vector spaces have the same dimension iff isomorphic to each other. We can argue from the fact that if you apply an isomorphism to a basis, you just get a different linearly independent basis, and from the fact that you can always find a map from an n-dimensional vector space to $\mathbb{R}^n$, just by using "real" coordinates, and thus spaces isomorphic to it are isomorphic to each other. What this will mean is, as long as we don't change the number of dimensions, we can probably change our space back to the way it was — doing some math in the meantime with whatever view of the data is most convenient.

~~Transposition is an example of an isomorphism~~ — ~~N doesn't change the rank.~~

What's interesting about homomorphisms? Basically, that all matrices represent homomorphisms, and vice versa. That is, if there is some transformation that preserves structure, we can use a matrix to represent it; and if there is a matrix, we can think of it as

Such a mapping. —skip—

We'll just show one direction of this equivalence, that every matrix is a homomorphism. (This can justify using the distributive property, $A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}$.)

We just need to show $h(c\vec{v}+d\vec{u}) = ch(\vec{v}) + dh(\vec{v})$ proving both properties at once, if $h()$ represents the action of the matrix. If $H = \begin{bmatrix} h_{11} & \cdots & h_{1n} \\ \vdots & & \\ h_{m1} & & h_{mn} \end{bmatrix}$

and $\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$

then $h(c\vec{v}+d\vec{u}) = H(c\vec{v}+d\vec{u}) = \begin{bmatrix} & H & \end{bmatrix} \begin{bmatrix} cv_1 + du_1 \\ \vdots \\ cv_n + du_n \end{bmatrix} =$

$$\begin{bmatrix} h_{11}(cv_1+du_1) + \cdots + h_{1n}(cv_n+du_n) \\ \vdots \\ h_{m1}(cv_1+du_1) + \cdots + h_{mn}(cv_n+du_n) \end{bmatrix} = ch(\vec{v}_1) + dh(\vec{u}_1)$$

$$\underset{\text{distributive!}}{=} \begin{bmatrix} h_{11}cv_1 + \cdots + h_{1n}cv_n \\ \vdots \\ \\ \end{bmatrix} + \begin{bmatrix} h_{11}du_1 + \cdots + h_{1n}du_n \\ \vdots \\ \\ \end{bmatrix}$$

Because the properties needed to be a homomorphism are properties of linearity, matrices can also be said to represent <u>linear maps</u>.

Maintaining & losing dimensions

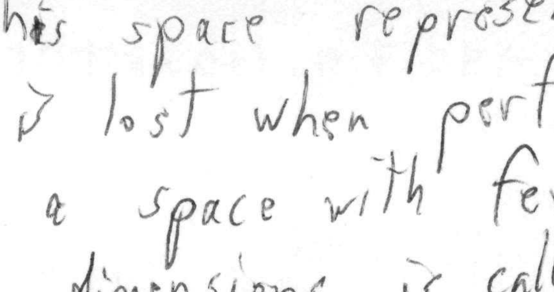
~~Recall that~~ if a transformation maintains the number of dimensions, then it's an "isomorphism". That means matrices where the space of possible vectors they produce has the same dimension as the input vectors, are isomorphisms. Which matrices are those?

Definitely not matrices with fewer rows than columns— they lose dimensions, producing vectors with fewer numbers. And logically, we couldn't be both one-to-one and onto going to a larger number of dimensions either. So the matrix needs to be square to be an isomorphism.

But that's not quite sufficient. The rank of the matrix needs to be equal to n for an $n \times n$ matrix — that is, it needs to be nonsingular. This is true because we can show that a matrix's rank actually gives us the number of dimensions in its "range space", the space of vectors it can produce by multiplying. If the rank isn't n, then the range space has fewer dimensions than the "input" vector's space, so the mapping isn't one-to-one and onto.

Before we go about proving some of this, wk3 II
we'll note some applications. A big one is the idea of a matrix being invertible: being able to create a second matrix A^{-1} that undoes the action of the original matrix. We'll need that idea for least squares regression for example, not to mention graphics apps. And what this is saying is that non-square or singular matrices can't be inverted, because they're not going to be one-to-one and onto. (You can't invert a projection from more dimensions to fewer dimensions, for example — you don't know where it all came from exactly.)

Ok, so a bit more terminology. The kernel is the space of vectors that gets mapped to the zero vector under a transformation. (Also called the null space.) This space represents the information that is lost when performing a transformation to a space with fewer dimensions. Its number of dimensions is called the nullity (or just "dimension of the kernel"). In a projection from 2D to 1D straight down, we'd just lose the y-axis, basically.



(focusing on just $\vec{0}$ gets the # of lost dims right)

A fundamental result is the rank plus nullity equals the dimension of the domain. Whatever the rank doesn't keep around is lost.

So, how does this work? —skip—

- The null space is itself a vector space with some number of dimensions — its nullity. So we can create a basis for it with that many vectors.
- We can keep adding linearly independent basis vectors to form a basis for the whole space of the domain.
- ~~Any vector that is nonzero in the range space must be composed of a linear sum of these new vectors, because the other vectors in the basis of the null spa~~
- For any particular vector in the range space $h(\vec{v})$, we can argue that it's the sum
  $h(c_1\vec{\beta}_1) + h(c_2\vec{\beta}_2) + \cdots h(c_r\vec{\beta}_r)$, where
  $\vec{\beta}_1 \ldots \vec{\beta}_r$ are the basis vectors we added on top of the null space vectors, because when we applied $h(\vec{v})$ we got these terms plus $h(c_{n1}\beta_{n1}) + h(c_{n2}\beta_{n2}) + \cdots$ which all became zero (the null space components).
  So these extra basis vectors span the range space.
- The number of dimensions in the range space

p. 218 is equal to the rank of the matrix.

~~The short and handwavy idea here is that~~

This is because the rank of a matrix is the number of linearly independent columns it has, and we can show that the vectors in the range space must be linear sums of the column vectors.

So the dimension of the range space is no wk318 bigger or smaller than the number of linearly independent columns in the matrix - its rank.

And, zooming out a bit, we therefore have that the dimension of the domain is equal to the nullity - dimensions lost - plus rank - dimensions in the range space (image).

Example

$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 5 \\ 3 & 6 & 6 \end{bmatrix}$ has obviously linearly dependent column vectors $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}$. We also notice that we can write its effect as

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 5 \\ 3 & 6 & 6 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} v_1 + 2v_2 + 4v_3 \\ 2v_1 + 4v_2 + 5v_3 \\ 3v_1 + 6v_2 + 6v_3 \end{bmatrix}$$

Thinking about the range space, the possible results are of the form

$$v_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + v_2 \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix} + v_3 \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$$

Which only has two linearly independent vectors (pick one of the two dependent and $\begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$). So we know there's a one-dimensional space that goes to $\vec{0}$.

This also means there should be multiple solutions. If we go to reduced echelon form, we get

$$\begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 5 \\ 3 & 6 & 6 \end{bmatrix} \xrightarrow{\text{sub } 1} \begin{array}{c} \\ \times 2 \\ \times 3 \end{array} \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & -3 \\ 0 & 0 & -6 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 4 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The rank is 2, and the solutions for $=\vec{0}$ require $v_3 = 0$ and $v_2 = -\frac{1}{2} v_1$. Hence the single-dimensional kernel - there's one degree of freedom.

A short summary for isomorphisms -skip-

The following statements about a ~~homomorphism~~ (in other words, a matrix) linear map h are equivalent.

- It is one-to-one
- It has an inverse from its range space to its domain
- Its kernel consists of just $\vec{0}$
- Its rank is equal to the dimension of its domain

Matrix Operations

We haven't said much yet about what we're allowed to do with matrices from a purely mechanical point of view.

~ Scalar multiplication & addition

We can add matrices $A+B$ by adding their components - if they're the same size matrices - or multiply a matrix by a constant (scalar) by multiplying its entries.