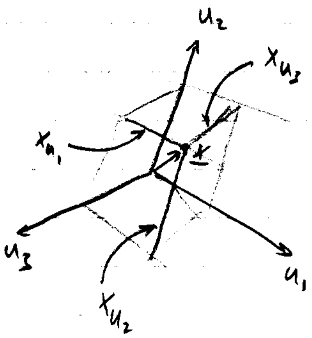
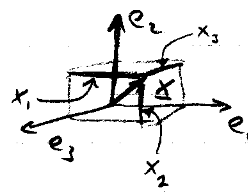
Linear Algebra and Matrices (a few highlights for c.g.)

A bit of motivation: matrices as a shorthand for repetitive algebra.





We discussed eariler how $\underline{x}$ can be represented by coordinates (x_1, x_2, x_3) w.r.t. the canonical basis $\{e_1, e_2, e_3\}$ or by coordinates $(x_{u_1}, x_{u_2}, x_{u_3})$ w.r.t. the basis $\{u_1, u_2, u_3\}$. We can convert between these representations assuming we know the coordinates $(u_1)_1, (u_1)_2, (u_1)_3, (u_2)_1, \dots$ of u_1, u_2, u_3 w.r.t. the canonical basis.

$$\underline{x} = x_{u_1} u_1 + x_{u_2} u_2 + x_{u_3} u_3$$ $\leftarrow$ $u \to e$ is by linear combination

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_{u_1}(u_1)_1 + x_{u_2}(u_2)_1 + x_{u_3}(u_3)_1 \\ x_{u_1}(u_1)_2 + x_{u_2}(u_2)_2 + x_{u_3}(u_3)_2 \\ x_{u_1}(u_1)_3 + x_{u_2}(u_2)_3 + x_{u_3}(u_3)_3 \end{bmatrix} \quad \text{or} \quad x_i = \sum_j \underbrace{(u_j)_i}_{3\times 3 \text{ array}} \underbrace{x_{u_j}}_{3\times 1 \text{ vector}}$$

$$x_{u_1} = \underline{x} \cdot \underline{u}_1 \ ; \quad x_{u_2} = \underline{x} \cdot \underline{u}_2 \ ; \quad x_{u_3} = \underline{x} \cdot \underline{u}_3$$ $\leftarrow$ $e \to u$ is by dot products (assumption: $\{u_1, u_2, u_3\}$ is ONB)

$$\begin{bmatrix} x_{u_1} \\ x_{u_2} \\ x_{u_3} \end{bmatrix} = \begin{bmatrix} x_1(u_1)_1 + x_2(u_1)_2 + x_3(u_1)_3 \\ x_1(u_2)_1 + x_2(u_2)_2 + x_3(u_2)_3 \\ x_1(u_3)_1 + x_2(u_3)_2 + x_3(u_3)_3 \end{bmatrix} \quad \text{or} \quad x_{u_i} = \sum_j \underbrace{(u_i)_j}_{3\times 3 \text{ array}} \underbrace{x_j}_{3\times 1 \text{ vector}}$$

These transformations boil down to operations combining a 3x3 array of numbers with a 3x1 vector of numbers to produce a 3x1 vector.

motivation for matrices, contd.

To summarize this common operation compactly, we introduce matrices and matrix multiplication.

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{\underline{x}} = \underbrace{\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{bmatrix}}_{A\underline{x}} \qquad (A\underline{x})_i = \sum_j a_{ij} x_j$$

This can be interpreted as the two operations we just saw:

$A\underline{x}$ is a linear combination of the columns of A with the elements x_1, x_2, x_3 as the coefficients.

$$\begin{bmatrix} | & | & | \\ \underline{c}_1 & \underline{c}_2 & \underline{c}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\underline{c}_1 + x_2\underline{c}_2 + x_3\underline{c}_3$$

The elements of $A\underline{x}$ are dot products of the rows of A with $\underline{x}$:

$$\begin{bmatrix} — & r_1 & — \\ — & r_2 & — \\ — & r_3 & — \end{bmatrix} \underline{x} = \begin{bmatrix} \underline{r}_1 \cdot \underline{x} \\ \underline{r}_2 \cdot \underline{x} \\ \underline{r}_3 \cdot \underline{x} \end{bmatrix}$$

The rule for matrix - vector multiplication extends to arbitrary dimensions.

So this is why we have matrices: they provide a convenient shorthand for several common operations we need to do often. (Later we'll introduce more uses beyond re-representing vector coordinates)

Definitions: transpose: $\begin{bmatrix} | & | & | \\ a & b & c \\ | & | & | \end{bmatrix}^T = \begin{bmatrix} — & a & — \\ — & b & — \\ — & c & — \end{bmatrix}$ or $(A)_{ij} = (A^T)_{ji}$ reflect

diagonal: $\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$ ← the vector of elements a_{ii}

matrix defns, contd.

if $\{u_1, u_2, u_3\}$ is a basis, $\overset{U}{\begin{bmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{bmatrix}}$ is the basis-to-canonical matrix

if $\{u_1, u_2, u_3\}$ is ONB, $\underset{U^T}{\begin{bmatrix} - u_1 - \\ - u_2 - \\ - u_3 - \end{bmatrix}}$ is the canonical-to-basis matrix.

$$\left.\begin{array}{l} x_e = U x_u \\ x_u = U^T x_e \end{array}\right\} \text{U^T undoes the effect of U. (true only for ONBs)}$$

A matrix that has orthonormal rows ($\Leftrightarrow$ orthonormal columns) is an orthonormal matrix.

(so the matrix of an ONB is orthonormal)

Why should the columns being orthonormal have anything to do with the rows?

re-write e_1, e_2, e_3 in the $\{u_1, u_2, u_3\}$ basis:

$$\overset{U^T}{\begin{bmatrix} - u_1 - \\ - u_2 - \\ - u_3 - \end{bmatrix}} \overset{e_1}{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}} = \begin{bmatrix} (u_1)_1 \\ (u_2)_1 \\ (u_3)_1 \end{bmatrix} \leftarrow \text{first column of } U^T$$

Those vectors are still orthonormal when re-written in another ONB. (this would not be true if rewritten in an arbitrary basis)

So: the columns of U are the basis vectors, expressed in the canonical basis. The rows of U are the canonical basis vectors, expressed in the $\{u_1, u_2, u_3\}$ basis.

Matrix multiplication

Often we have a basis, say V, written in terms of another basis, say U. So if we have x_v, then Vx_v is x_u.

Matrix multiplication cont.

Once we have $\underline{x}_u$ we can compute $\underline{x}_e = U\underline{x}_u = U(V\underline{x}_v)$.
It would be nice to be able to have a single matrix UV that satisfies $(UV)\underline{x}_v = \underline{x}_e = U(V\underline{x}_v)$. This is matrix multiplication.

Very simple: to go directly from $\underline{x}_v$ to $\underline{x}_e$ we just need to write the basis vectors $\{\underline{v}_1, \underline{v}_2, \underline{v}_3\}$ in the canonical basis (not in the U basis). We know that each vector can be transformed individually:

$(\underline{v}_i)_e = U(v_i)_u$

what we have in V.

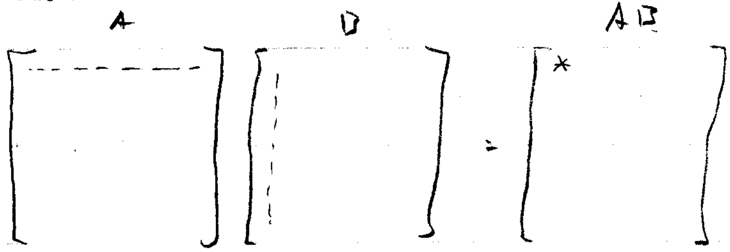
$$\text{so } U\begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} = \underbrace{\begin{bmatrix} | & | & | \\ U\underline{v}_1 & U\underline{v}_2 & U\underline{v}_3 \\ | & | & | \end{bmatrix}}_{\text{just transform the three columns.}}$$

* note: AB means "first B, then A" This can be confusing!

Hence the rule for matrix multiplication.

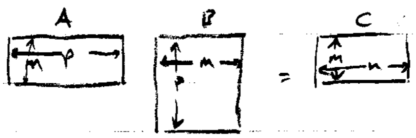
$$(AB)_{ik} = \sum_j a_{ij} b_{jk}$$

one interpretation: (row i of A) dot (col j of B)



This works for any dimension of matrix (even non-square)

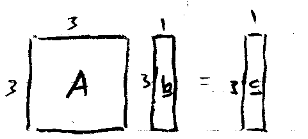
*note: $(AB)^T = B^T A^T$



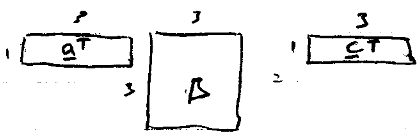
← * note dim p disappeared!

... BUT only as long as the rows of A and the columns of B are the same length! (otherwise how will we compute dot products?)

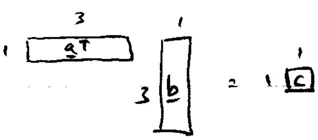
Special cases



mat-vect. multiply is just mat-mat multiply with only one column in B



vectors can also be treated as single-row matrices and multiplied on the left (for us only occasionally useful



dot product is just $\underline{a}\cdot\underline{b} = \underline{a}^T\underline{b}$

Matrix inverses

$\boxed{\text{given ONB } \mathcal{U}}$

Take x_u rewrite in canon. basis: $\underline{x}_e = \mathcal{U}\,\underline{x}_u$

now rewrite back in $\mathcal{U}$ basis: $\underline{x}_u = \mathcal{U}^T \underline{x}_e = (\mathcal{U}^T\mathcal{U})\,\underline{x}_u$

This holds for all $\underline{x}_u$, so multiplying by $(\mathcal{U}^T\mathcal{U})$ has no effect. We say $(\mathcal{U}^T\mathcal{U})$ acts as an <u>identity</u> matrix. ← note only true for ONB

There is only one identity matrix. To see what it is, think of transforming from the canonical basis to the canonical basis. The matrix E has, as its columns, the coordinates of $\{e_1, e_2, e_3\}$ in the canonical basis, or $(1,0,0)$, $(0,1,0)$, $(0,0,1)$. So:

$$E = \begin{bmatrix} (e_1)_e & (e_2)_e & (e_3)_e \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Actually, we call this I.

note only true for ONB

Since $\mathcal{U}^T\mathcal{U} = I$, this means $\mathcal{U}^T$ "undoes" the effect of $\mathcal{U}$.
A similar argument, taking $\underline{x}_e \to \underline{x}_u \to \underline{x}_e$, concludes that $\mathcal{U}\mathcal{U}^T = I$.

We say this means $\mathcal{U}^T$ is the <u>inverse</u> of $\mathcal{U}$, because it has the reverse effect:

$$y = \mathcal{U}x \iff x = \mathcal{U}^T y$$

← note only true for ONB

We write the inverse of A as A^{-1}, and for <u>orthonormal matrices only</u>, the inverse is the transpose. Note: $(AB)^{-1} = B^{-1}A^{-1}$.

Some matrices do not have inverses. These are the same ones for which the columns do not form a basis, so that all our discussions about coordinate transforms are null and void. This happens when the columns ($\Leftrightarrow$ the rows) are linearly dependent.

A matrix that does not have an inverse is <u>singular</u>.