Andrew Shao

Linear Algebra I

Unit 3: Matrices

Matrices:

A matrix is defined as a m x n rectangular array, where a_{ij} represents the i'th row and the j'th column.
Key Operations with Matrices:

1) A = B, iff $A_{ij} = B_{ij}$ for all i and j
2) $(A + B)_{ij} = A_{ij} + B_{ij}$
3) $(tA)_{ij} = t(A_{ij})$, where t is a scalar value.

Just like vectors, matrices also have the 10 fundamental properties from Theorem 1.1.1, and you can still take a linear combination of various matrices. If these 10 properties aren't already ingrained in your mind, take the time to memorize them, as they will appear again.

Transposing Matrices: To transpose a matrix, you swap the columns and rows of the matrix, where for all ${A_{ij}}^T$, the corresponding index into the matrix is A_{ji}.

Some matrices have the property that their transpose is exactly the same as the regular matrix, and such matrices are called *symmetric*. Note that only square matrices can be symmetric.

Properties of the Transpose:

1) $(A^T)^T = A$
2) $(A + B)^T = A^T + B^T$
3) $(cA)^T = c(A^T)$

Multiplying Matrices:

Matrix Vector Multiplication: matrix vector multiplication is really the same as matrix matrix multiplication, if you consider a vector to be a matrix of size n x 1(rows x columns). So, we jump right into matrix matrix multiplication and then return back to define matrix vector multiplication.

Matrix Matrix Multiplication: This operation is defined when the inner dimension of the two matrices are the same. Ex. when a matrix is size (a x b) and the right one is (b x c), the operation is defined and the resulting matrix is (a x c).

We perform matrix matrix multiplication by taking the dot product between each row of the left matrix with each column of the right one. This is best shown with an example.

Ex. Calculate $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

We first check to see that the matrix multiplication operation is valid, and we see that the first matrix has dimensions (2 x 3) and the second (3 x 2). Thus, it will work and will yield a matrix sized (2 x 2).
We first take the dot product between the first row([1 2 3]) and the first column([1 3 5]) and yield 22 as the first entry into the matrix(first row, first column).

We then repeat this for the second row and first column, yielding 49(second row, first column).

We then take the first row([1 2 3]) and dot it with the second column([2 4 6]), yielding 28(first row, second column).

Finally, we take the second row and take the dot product with the second column, yielding 64(second row, second column).

This gives us our final answer of $\begin{bmatrix} 22 & 28 \\ 49 & 64 \end{bmatrix}$

We can extend this definition of matrix matrix multiplication by thinking of a vector that is just a matrix n x 1, and performing matrix matrix multiplication as usual.

Properties of Matrix Multiplication:
1) A(B + C) = AB + AC
2) (B + C)A = BA + CA
3) t(AB) = (tA)B = A(tB)
4) (AB)C = A(BC)
5) $(AB)^T = B^T A^T$
6) $iff\ Ax = Bx$ for all x vectors, then A = B

Not Properties of Matrix Multiplication:
1) AB != BA
2) if AC = BC, A != B, there is no cancellation law.

Identity Matrix:

The Identity Matrix is a matrix *I(nxn)* that for any Matrix *M*, sized n x n, MI = M = IM. The identity matrix is a matrix with the basis vectors as it's columns.

Proof: AI = I

We want to find a matrix I where $[a_1, a_2, ..., a_n] = [a_1, a_2, ..., a_n][i_1, i_2, ..., i_n]$
Since $[a_1, a_2, ..., a_n]$ = A,

$[a_1, a_2, ..., a_n] = A[i_1, i_2, ..., i_n]$, and by definition of matrix matrix multiplication:
$[a_1, a_2, ..., a_n] = [Ai_1, Ai_{2,} ..., Ai_n]$

The way to satisfy this is to use the basis vectors, since $a_1 = Ae_1$

Proof: I is unique

We want to prove that the Identity matrix I is unique. We know that if there existed two different Identity Matrices, $I_1 and I_2$, $I_2 = I_1 I_2 = I_1$.

Thus, it's impossible that more than one identity matrix exists.

Block Matrices:

To make proofs or computation easier, we might break down matrices into "blocks", and we can see that the properties are the same.

Ex. When we split up A into $[a_1, a_2, ..., a_i]$, this is equivalent to breaking up matrix A into its column vectors.

Block Matrix Multiplication: If we break up matrices into their blocks, we can still use matrix multiplication on the blocks and treat them the same.

Matrix Mappings:

Basic Function Definitions: Functions map one element from set a to a unique *f(a)* in set b. We say that *f(a)* is the image of a under f. If a function changes the dimensionality of a problem, We can also represent a function mapping the dimension A to dimension B as so:

F: A -> B, where A is called the domain and B is the codomain.

We define a Matrix Mapping to be a function where $f(x) = Ax$. Matrix mappings can change the dimensionality of the problem($\mathbb{R}^n$ -> $\mathbb{R}^m$).

Observations regarding Matrix Mappings:

1) let *f(x)* represent Ax, [*f(1, 0), f(0, 1)*] is just the matrix A
2) Let *f(x)* represent Ax, $f(x_1, x_2) = x_1 f(1, 0) + x_2 f(0, 1)$
3) If A is a matrix sized m x n, then the domain is $\mathbb{R}^n$ and the codomain is $\mathbb{R}^m$.

Linearity of Matrix Mappings:

If *f(x) = Ax,* then *f(sx + ty) = sf(x) + tf(y)*
Proof:

$f(sx + ty) = A(sx + ty)$

$A(sx + ty) = sAx + tAy$

$f(sx + ty) = sf(x) + tf(y)$

Thus, *f(sx + ty) = sf(x) + tf(y)*

Linear Mappings:

We saw that matrix mappings have the property where *f(sx + ty) = sf(x) + tf(y).* Linear mappings are the class of functions that have this property. Thus, any linear mapping has the property that *L(sx + ty) = sL(x) + L(y).* This property continues for all terms in a linear combination(however long). Keep in mind that a linear mapping is equivalent to a linear transformation.

Definition: A linear mapping is called a linear operator if the domain and codomain are the same, i.e if the function maps from $\mathbb{R}^n$ to $\mathbb{R}^n$.

Proof that $f(x_1, x_2) = (2x_1 + x_2, -3x_1 + 5x_2)$ is linear.

By inspection, we can already just see that the function is linear, but we need to provide a formal proof, and thus we want to use the linearity property.

$$f(sx + ty) = f(s(x_1, x_2) + t(y_1, y_2))$$
$$f(sx + ty) = f((sx_1, sx_2) + (ty_1, ty_2))$$
$$f(sx + ty) = f((sx_1 + ty_1), (sx_2 + ty_2))$$
$$f(sx + ty) = f(sx_1 + ty_1, sx_2 + ty_2)$$
$$f(sx + ty) = (2(sx_1 + ty_1) + (sx_2 + ty_2), -3(sx_1 + ty_1) + 5(sx_2 + ty_2))$$
$$f(sx + ty) = (2sx_1 + 2ty_1 + sx_2 + ty_2, -3sx_1 - 3ty_1 + 5sx_2 + 5ty_2)$$
$$f(sx + ty) = (2sx_1 + sx_2, -3sx_1 + 5sx_2) + (2ty_1 + ty_2, -3ty_1 + 5ty_2)$$
$$f(sx + ty) = sf(x) + tf(y)$$

Thus, this function is linear.

For some functions, it might be too difficult to get a direct proof of why a given function is linear or not. Thus, if you look by inspection and see that it shouldn't be linear, an alternative solution is to find a counterexample that proves that it isn't linear.

Are all linear mappings matrix mappings?

Previously, we figured out that all matrix mappings are linear mappings, but is the converse true? Yes! To find the matrix representation of a linear function, we simply compute the function on all standard basis vectors, and this forms the matrix mapping needed. This is called the standard matrix of the function.

Proof:

$$L(x) = L(x_1e_1 + x_2e_2 + ... + x_ne_n)$$

By Linearity(x_1 is a constant applied to the basis vector):

$$L(x) = x_1L(e_1) + x_2L(e_2) + ... + x_nL(e_n)$$

This is just a dot product:

$$L(x) = [L(e_1) + L(e_2) + ... + L(e_n)]\, x$$

Thus,

$$L(x) = [L]x$$

Before constructing a Standard matrix, *always* check if the function is linear first. This condition must be true before the standard matrix works.

Geometric Linear Mappings:

Some linear mappings have geometric properties, and when they are applied, vectors change in some constant geometric ways.

Rotations:

To rotate a vector around a point, we need to do a few basic trigonometric operations. If we first consider the simplest case of rotation, where the first vector is on the x-axis(x, 0), we can rotate it with the following operations(If (x', y') is the desired point).

$$cos(\theta) = x' / r$$
$$x' = rcos(\theta) = xcos(\theta)$$

Similar logic can be applied for y(except it's with the sine function instead)

For any given point, the logic now becomes:

$$cos(\alpha) = x / r$$
$$x = rcos(a)$$

$$cos(\alpha + \beta) = x' / r$$
$$x' = rcos(\alpha + \beta)$$

With the trigonometric identity that cos(a + b) = cosAcosB - sinAsinB, we get:

$$x' = r(cos\alpha cos\beta - sin\alpha sin\beta)$$
$$x' = rcos\alpha cos\beta - rsin\alpha sin\beta$$
$$x' = xcos\beta - ysin\beta$$

Similar derivations can occur for y'(which turns out to be $y' = ycos\beta + xsin\beta$))

With this, we can use a matrix to represent any given rotation. Being $\begin{bmatrix} cosA & sinB \\ sinB & cosB \end{bmatrix}$, which is a matrix mapping that will map any vector to it's rotated equivalent.

Furthermore, this *rotation matrix* can be applied like any other linear function(ex. tR(x) = R(tx), R(a) + R(b) = R(a + b), etc.). This matrix form only allows for counterclockwise rotation, and for clockwise rotation, swap the sign of the two sines in the matrix.

Formal Proof that the rotation operation is linear:

$$R(sx + ty) = s(x_x, x_y) + t(y_x, y_y)$$
$$s(x_x, x_y) + t(y_x, y_y) = (sx_x, sx_y) + (ty_x, ty_y)$$
$$s(x_x, x_y) + t(y_x, y_y) = (sx_x, sx_y) + (ty_x, ty_y)$$
$$(sx_x, sx_y) + (ty_x, ty_y) = (sx_x + ty_x, sx_y + ty_y)$$

Plug these x and y coord into the rotation operation($R(x, y) = (xcosB - ysinB, xsinB + ycosB)$

$$R(sx_x + ty_x, sx_y + ty_y) = ((sx_x + ty_x)cosB - (sx_y + ty_y)sinB, (sx_x + ty_x)sinB + (sx_y + ty_y)cosB)$$
$$R(sx_x + ty_x, sx_y + ty_y) = (sx_xcosB + ty_xcosB - sx_ysinB - ty_ysinB, sx_xsinB + ty_xsinB + sx_ycosB + ty_ycosB)$$

Group Terms and Rearrange

$$R(sx_x + ty_x, sx_y + ty_y) = (s(x_xcosB - x_ysinB) + t(y_xcosB - ty_ysinB), s(x_xsinB + x_ycosB) + t(y_xsinB + y_ycosB))$$
$$R(sx_x + ty_x, sx_y + ty_y) = s((x_xcosB - x_ysinB), s(x_xsinB + x_ycosB)) + t((y_xcosB - ty_ysinB), (y_xsinB + y_ycosB))$$

Thus we get:

$$R(sx_x + ty_x, sx_y + ty_y) = sR(x) + tR(y)$$

Reflection Matrices:

We can reflect a vector across a plane using the formula $refl_P x = x - 2perp_P$ or equivalently:

$$refl_P x = x - 2proj_n$$

At the current moment, we cannot compute the standard matrix of this operation simply due to how the equation is defined based on the plane that you reflect over, but the process would be the same(Plug in the Normal vector and use the function on the basis vectors).

Range: The set of all outputs that the linear mapping can produce, for any vector inputted. From this set, you could either find a *basis* for the set or find if a vector exists inside a given set.

Theorem 3.3.2: If a Linear Mapping from domain $\mathbb{R}^n$ to codomain $\mathbb{R}^m$, then Range(L) is a subspace of the codomain.

Proof:

We must first check that the 0 vector exists, and since it's a linear mapping, inputting the 0 vector will yield the 0 vector back.

We must then check for closure under addition: y + z = L(x) + L(w) = L(x + w). Since L is linear, this is pretty simple.

Finally, we check for closure under scalar multiplication, and find that tL(x) = L(tx), which is again a simple proof since L is linear.

Column Space: The same as range, but now a linear combination of the standard matrix vectors. This means that Col([A]) = Range(A).

Kernel:

The kernel of a linear mapping is the set of all vectors where the output of the linear mapping is the zero vector.

Theorem 3.3.3: The kernel of a linear mapping is always a subspace of the domain of the linear mapping. A proof of this is super simple since all vectors in the kernel result in the output of the 0 vector.

NullSpace: The same as kernel, but on the standard matrix of the Linear Mapping

Left NullSpace: The Nullspace of A^T

Row Space: The column space of A^T

Operations on Linear Mappings:

Note: *Any* linear combination of linear mappings is another linear mapping.

Equality: L = M iff L(x) = M(x)
Addition: (L + M)(x) = L(x) + M(x)
Scalar Multiplication: (tL)(x) = tL(x)

These three properties exist for the standard matrices of the linear mappings.

Equality: L(x) = M(x) iff [L]x = [M]x
Addition: (L + M)(x) = [L + M]x
Scalar Multiplication: (tL)(x) = t[L]x

Since we just showed that the linear mappings are pretty much equivalent to the standard matrices, the 10 properties that apply to vectors also apply to linear mappings.

Composition of Linear Mappings: We can compose linear mappings by nesting linear functions together.
$(M \circ L) = M(L(x))$
It's not surprising that this operation is linear, but what is surprising is that:
$M \circ L = [M][L]$

Invertible Mappings: if a linear mapping has the following property, we call L and M to be invertible.
$(M \circ L)x = x \textit{ iff } M = L^{-1} \textit{ and } L = M^{-1}$

Next Unit: Vector Spaces.