Andrew Shao

Linear Algebra I

Unit 6: Diagonalization

Linear Mappings and Bases:

Many linear mappings have nice geometrical properties, but show none of these properties in the standard matrix. Projecting the matrices to a basis can help reveal what properties the matrix has.

$L(x) = [L]_B[x]_B$, we want to project the standard matrix onto the vector space, in the same way we had done it in previous units.

Recall that we can compute the standard matrix by applying the linear mapping onto all basis vectors, when we compute it for a specific basis, we apply the linear mapping onto that specific basis' vectors

Geometric Interpretation of this basis linear mapping: If we have (ordered) vectors:
$\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-1\end{bmatrix}\}$ as a basis and our linear mapping is computed to be the matrix:
$\begin{bmatrix}5 & 0 & 0\\0 & -1 & 0\\0 & 0 & -1\end{bmatrix}$, this refers to us stretching any given vector 5 times in the direction of the first vector, and reflecting over the second and third vectors.

Diagonal and Similar Matrices:

Diagonal matrices: Matrices where $A_{ij} = 0 \; if \; i \, != \, j$. Note that diagonal matrices are both upper and lower triangular.

Similar Matrices provide us with another method of projecting a linear mapping into a different basis. We know the following property:

$A_B = [\,[Av_1]_{B'}\, [Av_2]_{B'}\, ...]$
We also know that we can construct a change of basis matrix to perform the same task:
$A_B = [BP_S[Av_1].\; BP_S[Av_2],\; ...]$
Furthermore, we know that $BP_S = [SP_B]^{-1}$, thus:

$A_B = [[SP_B]^{-1}[Av_1],\; ...]$

$A_B = [[SP_B]^{-1}A[v_1,\; v_{2,}...]]$
Thus,
$A_B = [SP_B]^{-1}A[BP_S]$

If there exists an invertible matrix $P^{-1}AP = B$, the following matrices are *similar* and have a few key properties:

1) rank(A) = rank(B)
2) det(A) = det(B)
3) tr(A) = tr(B), where tr = the trace of a matrix and it's the sum of diagonal entries.

Eigenvectors and Eigenvalues:

When computing Eigenvectors and Eigenvalues, we have the goal of taking the original standard matrix and writing it wrt a basis where the new basis will be diagonal.

Considering we don't even know what the basis is, we need to work in reverse.

$[L]_B = diag(\lambda_1, ..., \lambda_n)$, where lambda is just some future value of the matrix. Then, we replace it with the similar matrix $P^{-1}LP$.

$P^{-1}[L]P = diag(\lambda_1, ..., \lambda_n)$
$LP = Pdiag(\lambda_1, ..., \lambda_n)$
Now, let's break down the matrices.
$L[v_1, ..., v_n] = [v_1\lambda_1, ..., v_n\lambda_n]$, the matrix multiplication on the right side acts like a dot product due to how the second matrix is diagonal.

$[Lv_1, ..., Lv_n] = [v_1\lambda_1, ..., v_n\lambda_n]$

Thus, for all i, $Lv_i = v_i\lambda_i$.

We write this in terms of the basis B, and since we know that P is just the basis vectors
$[Lv_i]_b = \lambda I$
Thus,
$[L]_B = diag(\lambda)$

We basically are trying to solve for the equation $Lv = \lambda v$, where we call λ an eigenvalue and v an eigenvector. Note that the vector v cannot be 0, and eigenvalues simply scale a given eigenvector up in some direction, which has some nice geometric properties.

You can easily check if an eigenvalue is valid by substituting in for the value lambda and solving for the vector v. And similarly, you can check if an eigenvector is valid by solving for lambda.

But, how can we find all eigenvectors and eigenvalues? We can perform this with some basic algebra.

$Lv - \lambda v = 0$
$(L - \lambda I)v = 0$
For this system to be consistent, $det(L - \lambda I) = 0$, meaning it cannot be invertible. Why? Otherwise, we could simply perform the following

$P = (L - \lambda I)$
$P^{-1}Pv = P^{-1}0$
$v = 0.$

Thus, we can simply solve for when det(L - λI) = 0, to get the eigenvalues, and use this to compute eigenvectors.

We call this determinant the *characteristic polynomial,* and an nxn matrix will have n roots(By the fundamental theorem of Algebra)

Note that we call the dimension of the null space of the characteristic polynomial has to be at least dimension 1, meaning there is minimally 1 row of zeros for there to be an eigenvector.

The number of times the lambda appears as a root is called the *Algebraic Multiplicity,* and the number of null rows in the RREF forms is the *Geometric Multiplicity.* Note that the geometric multiplicity(dim of Null space) is at least 1 and always less than or equal to the Algebraic Multiplicity for lambda to be an eigenvalue(*Theorem 6.2.3)*

Theorem 6.2.4: if $\lambda_1, \ldots, \lambda_n$ are all eigenvalues, then det(A) = $\lambda_1 \ldots \lambda_n$, tr(A) = $\lambda_1 + \ldots + \lambda_n$

Diagonalization Theory:

We have been working with the formula $P^{-1}AP = D$, where A is considered to be a diagonalizable matrix, P is an invertible matrix consisting of eigenvectors, and D is a diagonal matrix with eigenvalues. Just as terminology, we also say that A is a *similar matrix* to D.

Lemma 6.3.1: If we have n distinct eigenvalues, there will be n linearly independent eigenvectors.

Proof By Induction:

Clearly, if we only have one eigenvector, it is linearly independent. However, for larger amounts of eigenvectors, we see the following:
$Av = \lambda v$
$(A - \lambda)v = 0$
$(\lambda_j - \lambda_i)v_i$

This is what each vector represents in the linear independence equation
$c_1(\lambda_{k+1} - \lambda_1)v_1 + \ldots + c_{k+1}(\lambda_{k+1} - \lambda_{k+1})v_{k+1} = 0$
We know from previous steps(induction) that c_1 -> c_k = 0, thus:
$c_{k+1}v_{k+1} = 0$
Since v_{k+1} cannot be the 0 vector, by eigenvectors, c_k+1 also has to be 0. QED.

Eigenspace: All vectors A where $Ax = \lambda x$

Theorem 6.3.2: The basis vectors for every eigenspace(Joined set) is linearly independent.

Theorem 6.3.3(Diagonalization Theorem): A given matrix is diagonalizable if every eigenvector's geometric multiplicity is equal to its algebraic multiplicity. Otherwise, we wouldn't have n unique eigenvectors. Note this simply means that if we have n distinct eigenvalues, the given matrix A is diagonalizable.

Algorithm to Diagonalize a Matrix A:

1) Solve for the characteristic Polynomial $C = Det(A - \lambda I)$
2) Solve for the roots, they should all be real and there should be n unique eigenvalues, otherwise it is not diagonalizable
3) Using the formula: $(A - \lambda I)x = 0$, find a basis for each of the eigenspaces to solve for x, which is an eigenvector
4) The geometric multiplicity should be greater or equal to it's algebraic multiplicity. If not, it is not diagonalizable.
5) Form a basis of the eigenvectors, and use this as a basis for the matrix A, using the formula: $P^{-1}AP$

Note that we can "refactorize" back to A using PLP^{-1}.

Application of Diagonalization: Powers of Matrices

Computing powers of matrices can be tedious and sometimes impossible, if you are asked to compute A^{1000}. Diagonalization can help with this, as the exponentiation of a diagonal matrix is simply the exponent of the diagonal row. Thus, if we are able to project the matrix into a diagonal space using eigenvalues and eigenvectors, we can easily compute the exponentiated matrix and then project it back to it's original form.

Ex. Compute A^{1000} if $A = \begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$:

We first compute the eigenvalues, which are 2 and 3.

Using these eigenvalues, we compute the eigenvectors, which are $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

We then compute the inverse $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ and use this to compute the diagonal matrix: $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$.

We exponentiate it and get: $\begin{bmatrix} 2(1000) & 0 \\ 0 & 3(1000) \end{bmatrix}$. After converting back to the original matrix, we get: $\begin{bmatrix} 2(1001) - 3(1000) & 2 * 3(1000) - 2(1001) \\ 2(1000) - 3(1000) & 2 * 3(1000) - 2(1000) \end{bmatrix}$.

Note that LATEX wasn't formatting properly, so brackets imply exponents(2(1000) = 2^1000).

Next Unit: Linear Algebra 2!