Notes on Chapter 3

1. The eigenvalues λ_i of an N by N matrix A are the (possibly complex) roots of the characteristic polynomial det($\lambda I - A$). This N^{th} degree polynomial can be factored as

   $(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}...(\lambda - \lambda_k)^{m_k}$

2. m_i is called the algebraic multiplicity of the eigenvalue λ_i. The sum of the algebraic multiplicities is always N.

3. For each eigenvalue λ_i the set of solutions to $(\lambda_i I - A)z = 0$ (the set of eigenvectors corresponding to λ_i) is a subspace of R^N, called the eigenspace for that eigenvalue. The dimension of the eigenspace corresponding to λ_i is called the geometric multiplicity, n_i, of this eigenvalue.

4. Now let's choose a basis for the eigenspace of λ_1, and place these n_1 vectors in the first n_1 columns of a matrix S (n_i is the geometric multiplicity of λ_i). Choose the remaining columns of this N by N matrix so that they form a basis for R^N (recall the extension to a basis theorem). Then AS = SE, where E has the form:

$$\begin{bmatrix} \lambda_1 & 0 & ... & 0 & x & ... & x \\ 0 & \lambda_1 & ... & 0 & x & ... & x \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_1 & x & ... & x \\ 0 & 0 & ... & 0 & x & ... & x \\ 0 & 0 & ... & 0 & x & ... & x \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & x & ... & x \end{bmatrix}$$

5. Since S has an inverse, $A = SES^{-1}$ (A is "similar" to E), so
   det($\lambda I - A$) = det ($\lambda SS^{-1} - SES^{-1}$) = det (S($\lambda I - E)S^{-1}$) =
   det(S)det($\lambda I - E$)det(S^{-1}) = det(SS^{-1})det($\lambda I - E$) = det($\lambda I - E$)
   so A and E have exactly the same characteristic polynomial, and the characteristic polynomial for E contains the factor $(\lambda - \lambda_1)^{n_1}$. Thus the algebraic multiplicity m_1 is as least as large as the geometric multiplicity n_1. The same is clearly true for all eigenvalues, so in general we see that $1 \leq n_i \leq m_i$. If $n_i < m_i$, the eigenvalue λ_i is said to be "defective" (it's missing some of its eigenvalues).

6. Any set of eigenvectors z_i corresponding to distinct eigenvalues λ_i are independent. To show this, assume $\sum_{i=1}^{k} \alpha_i z_i = 0$. Then by multiplying both sides of this equation by A^j we get:
   $\sum_{i=1}^{k} \alpha_i A^j z_i = \sum_{i=1}^{k} \alpha_i \lambda_i^j z_i = 0$
   If we take j=0,...,k-1, we get k equations for the k unknowns $\alpha_i z_i, i = 1, ..., k$:

$$\begin{bmatrix} 1 & 1 & 1 & ... & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & ... & \lambda_k \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & ... & \lambda_k^2 \\ \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & ... & \lambda_k^3 \\ ... & ... & ... & ... & ... \\ \lambda_1^{k-1} & \lambda_2^{k-1} & \lambda_3^{k-1} & ... & \lambda_k^{k-1} \end{bmatrix} \begin{bmatrix} \alpha_1 z_1 \\ \alpha_2 z_2 \\ \alpha_3 z_3 \\ ... \\ \alpha_k z_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ ... \\ 0 \end{bmatrix}$$

   The matrix is the Vandermonde matrix, and we showed in a homework problem that if all the λ_i are distinct, the determinant of this matrix is nonzero, thus $\alpha_i z_i = 0$ for each i, and since $z_i \neq 0, \alpha_i = 0.$

7. Since the sum of the algebraic multiplicities is N, the sum of the geometric multiplicities is less than or equal to N. If it is equal to N, then A has a complete set of N linearly independent eigenvectors, because there are n_i independent eigenvectors for each λ_i, and we saw in (6) that eigenvectors for different eigenvalues are independent. So load up all N linearly independent eigenvectors in the columns of a new matrix S. Then AS = SD, where D is a diagonal matrix with the eigenvalues of A along the diagonal, and $A = SDS^{-1}$ and we say that A is diagonalizable (similar to a diagonal matrix).

8. If all eigenvalues are distinct, then $1 \leq n_i \leq m_i = 1$ and so $n_i = m_i$ for each i, and A has a complete set of eigenvectors and is therefore diagonalizable. If A has eigenvalues of algebraic multiplicity greater than 1, it will be diagonalizable only if no eigenvalues are defective.