Abstract. This document summarizes diagonalization. A fuller account can be found in Holt, §6.4.

1. Diagonalization

Definition 1. An $n \times n$ matrix $A$ is **diagonalizable** if there is a basis of $\mathbb{R}^n$ consisting of eigenvectors of $A$.

Example 2. A diagonal matrix $D = \text{diag}(d_1, \ldots, d_n)$ has eigenvectors $e_1, \ldots, e_n$ with eigenvalues $d_1, \ldots, d_n$, so $D$ is diagonalizable.

Example 3. If $A$ has $n$ distinct eigenvalues, their corresponding eigenvectors are linearly independent, so they form a basis, so $A$ is diagonalizable.

Theorem 4. An $n \times n$ matrix $A$ is diagonalizable if and only if there is a diagonal matrix $D$ and an invertible matrix $P$ such that

$$A = PDP^{-1}.$$ 

In this case, if the columns of $P$ are given by $P = [v_1 \cdots v_n]$ and the diagonal entries of $D$ are given by $D = \text{diag}(d_1, \ldots, d_n)$, then $\{v_1, \ldots, v_n\}$ is a basis of $\mathbb{R}^n$ consisting of eigenvectors of $A$, and the eigenvalue of $v_i$ is $d_i$.

Proof. First suppose $A = PDP^{-1}$ with $D$ diagonal and $P$ invertible, as above. Since $P$ is invertible, its columns form a basis for $\mathbb{R}^n$, namely $\{v_1, \ldots, v_n\}$. We can pick off the columns of $P$ via $Pe_i = v_i$, so $e_i = P^{-1}v_i$. We verify that $v_i$ indeed is an eigenvector of $A$ with eigenvalue $d_i$:

$$Av_i = PDP^{-1}v_i = PD(e_i) = Pd_ie_i = d_ie_i = d_iv_i.$$ 

Hence $A$ has a basis of eigenvectors, so is diagonalizable.

Now suppose $A$ is diagonalizable. Pick a basis of eigenvectors $\{v_1, \ldots, v_n\}$ for $\mathbb{R}^n$ with eigenvalues $d_1, \ldots, d_n$ and set $P := [v_1 \cdots v_n]$ and $D := \text{diag}(d_1, \ldots, d_n)$. $P$ is evidently invertible and $D$ is diagonal, so we must only show that $A = PDP^{-1}$. The above computation shows that $PDP^{-1}$ has eigenvectors $\{v_1, \ldots, v_n\}$ with eigenvalues $d_1, \ldots, d_n$, so $PDP^{-1}v_i = d_iv_i$. By construction the same holds for $A$, namely $Av_i = d_iv_i$. The linear transformations of $PDP^{-1}$ and $A$ thus agree on a basis, so they agree everywhere, so $PDP^{-1} = A$. Alternatively, $(A - PDP^{-1})v_i = 0$, so the 0-eigenspace of $A - PDP^{-1}$ has dimension at least $n$, so $A - PDP^{-1}$ must be the zero matrix. 

Date: June 1, 2015.
Remark 5. Since $A = PDP^{-1}$, we have $P^{-1}AP = D$. Multiplying by $P^{-1}$ on the left and $P$ on the right is really computing $A$ “in the basis of eigenvectors”, and in that basis $A$ is diagonal. More formally, $P$ is the change of basis matrix from the basis of eigenvectors $B := \{v_1, \ldots, v_n\}$ to the standard basis $S$ and $P^{-1}$ is the change of basis matrix from $S$ to $B$. Hence if $x \in \mathbb{R}^n$ we have

$$(P^{-1}AP)x_B = P^{-1}Ax_S = P^{-1}(Ax) = [Ax]_B.$$ Compare this to $(A)x_S = (Ax)_S$, which says that $A$ takes in coordinate vectors with respect to the standard basis and returns coordinate vectors with respect to the standard basis. On the other hand, $(P^{-1}AP)x_B = [Ax]_B$ says that $P^{-1}AP$ takes in coordinate vectors with respect to basis $B$ and returns coordinate vectors with respect to basis $B$.

Remark 6. One of the main uses of diagonalization is to quickly compute powers of a matrix, which is essentially what happened in the Fibonacci number example from lecture. Indeed, if $A = PDP^{-1}$, then $A^2 = (PDP^{-1})(PDP^{-1}) = PDDP^{-1} = PD^2P^{-1}$, and more generally $A^k = PD^kP^{-1}$ for any $k \geq 0$ (and also any $k < 0$ if $A$ or equivalently $D$ is invertible). Computing powers of a diagonal matrix is easy since $\text{diag}(d_1, \ldots, d_n)^k = \text{diag}(d_1^k, \ldots, d_n^k)$.

Example 7. We can rephrase the Fibonacci computation from class in this language. With $F_0 = 0, F_1 = 1$, and $F_{n+2} = F_{n+1} + F_n$, we showed that

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} F_{n+1} \\ F_n \end{bmatrix}.$$ We found a basis of eigenvectors for the matrix on the left-hand side, which using the theorem shows that

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n = \begin{bmatrix} \phi & \bar{\phi} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \phi^n & 0 \\ 0 & \bar{\phi}^n \end{bmatrix} \begin{bmatrix} \phi & \bar{\phi} \\ 1 & 1 \end{bmatrix}^{-1},$$

where $\phi := (1 + \sqrt{5})/2$ and $\bar{\phi} := (1 - \sqrt{5})/2$. Combining these two observations immediately gives Binet’s formula.