Below is a (rough) chronological guide of the highlights from each section we have covered.

Some definitions will be given. In other places it may say "Define:" and then list a series of terms that were given with definitions in the section, usually in bold face in the text. You should try and first define these from memory, then later go back through the section(s) and check to make sure your definition didn't omit anything relevant.

Chapter 5 Review

Section 5.1

Define: eigenvalue, eigenvector, eigenspace. When $Ax = \lambda x$, with $x \neq 0$...we have $\lambda, x, \mathrm{Nul}\,(\mathrm{A} - \lambda \mathrm{I})$.
Triangular matrix has its eigenvalues on the main diagonal.
Theorem: Distinct eigenvalues implies linear independence among the corresponding eigenvectors.

Section 5.2

$\lambda = 0$ eigenvalue if and only if A is not invertible (singular).
Characteristic equation: Characteristic polynomial=det($A - \lambda I$)=0.
Define: Algebraic multiplicity (think factors of the characteristic polynomial).
Similarity: A is similar to B if there exists P invertible such that $A = PBP^{-1}$.

Section 5.3

If A is similar to a diagonal matrix, then A is **diagonalizable**, $A = PDP^{-1}$.
Theorem: diagonalizable if and only if there are n linearly independent eigenvectors.
$P = [v_1 \cdots v_n], D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$, where $v_1, \ldots, v_n$ are linearly independent eigenvectors corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$.
Distinct eigenvalues implies diagonalizable (but not the other way around).
$A^k = PD^kP^{-1}$.

Chapter 6 Review

Section 6.1

The **dot product** (or inner product) of 2 vectors $u, v \in \mathbb{R}^n$ is given by $u \cdot v = u^T v$. What are the properties of the dot product given in theorem 1?
Define: length of a vector in $\mathbb{R}^n$. What is the length squared in terms of the dot product?
The distance between 2 vectors $u, v \in \mathbb{R}^n$ is given by dist(u, v)= $||u - v||$.
$u, v \in \mathbb{R}^n$ are **orthogonal** if their dot product is 0.
What does the pythagorean theorem say in general for orthogonal vectors in $\mathbb{R}^n$?
A vector is orthogonal to a subspace if it is orthogonal to every vector in the subspace...this can be shown simply by showing the vector is orthogonal to every vector in a spanning set of the subspace.
Define: The orthogonal complement of a subspace W of $\mathbb{R}^n$.
Theorem: $(\mathrm{Row} A)^{\perp} = \mathrm{Nul}\,\mathrm{A}$ and $(\mathrm{Col}\,\mathrm{A})^{\perp} = \mathrm{Nul}\,\mathrm{A}^{\mathrm{T}}$.

Section 6.2

An **orthogonal set** is a set $S = \{u_1, \ldots, u_p\}$ such that each pair is orthogonal, i.e. $\forall i \neq j, u_i \cdot u_j = 0$.
Theorem: An orthogonal set of non-zero vectors is necessarily linearly independent.
An **orthogonal basis** is an orthogonal set that is also a basis.
Theorem: Given an orthogonal basis $\{u_1, \ldots, u_p\}$ for a subspace W of $\mathbb{R}^n$, any $y \in \mathbb{R}^n$ can be written as a linear combination $y = c_1u_1 + \cdots + c_pu_p$ where the coefficients are $c_j = \frac{y \cdot u_j}{u_j \cdot u_J}$.

The **orthogonal projection** of a vector y onto another vector u is given by $\hat{y} = \frac{y \cdot u}{u \cdot u} u$.
Any vector y can be decomposed into a sum of two vectors, one ($\hat{y}$) on the line spanned by a vector u and the other (z) orthogonal to that line, $y = \hat{y} + z$.
Define: orthonormal set and orthonormal basis.
Theorem: If a matrix U has orthonormal columns, then $U^T U = I$, then $n \times n$ identity matrix.
Any matrix U such that $U^T U = I$ is said to be an **orthogonal matrix**.
What are the properties of orthogonal matrices? (preserves length, dot products, and orthogonality).

Section 6.3

Orthogonal Decomposition Theorem: Given a subspace W of $\mathbb{R}^n, \forall y \in \mathbb{R}^n$ the decomposition $y = \hat{y} + z$ where $\hat{y}$ is the projection of y onto W and z is in $W^{\perp}$, is unique. Further, if $\{u_1, \ldots, u_p\}$ is an orthogonal basis for W then $\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{y \cdot u_p}{u_p \cdot u_p} u_p$ and $z = y - \hat{y}$.
Best Approximation Theorem: The projection of y onto a subspace W, $\text{proj}_W y = \hat{y}$ is the closest point in W to y, that is $||y - \hat{y}|| < ||y - w||, \forall w \in W$ such that $w \neq \hat{y}$.
If U has columns which form an orthonormal basis for a subspace W of $\mathbb{R}^n$ then $\text{proj}_W y = UU^T y$.

Section 6.4

The Gram-Schmidt Process (theorem): Given a basis $\{x_1, \ldots, x_p\}$ for a subspace W of $\mathbb{R}^n$, define $v_1 = x_1, v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1, v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2, \ \ldots \ , v_p = x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}$. The set $\{v_1, \ldots, v_p\}$ is an orthogonal basis for W and $\text{Span}\{x_1, \ldots, x_k\} = \text{Span}\{v_1, \ldots, v_k\}, \forall k = 1, \ldots, p$.
Normalizing the basis resulting from the Gram-Schmidt process will yield an orthonormal basis.
QR-Factorization Theorem: If an $m \times n$ matrix A has linearly independent columns, A can be factored into $A = QR$ where the columns of Q form an orthonormal basis for Col A and R is upper triangular with positive entries along its diagonal.
Note that A has linearly independent columns if and only if $\text{rank} A = n$ (why?). Also, this can only happen if $n \leq m$ (section 1.7).

Section 6.5

A **least-squares solution** is an $\hat{x} \in \mathbb{R}^n$ such that $||b - A\hat{x}|| \leq ||b - Ax||, \forall x \in \mathbb{R}^n$, where A is an $m \times n$ matrix and b is in $\mathbb{R}^n$. This is found by first projecting b into the column space of A, $\hat{b} = \text{proj}_{\text{Col A}} b$, then finding a solution to the consistent system $Ax = \hat{b}$, any such solutions are least-squares solutions.
Theorem: To find a least-squares solution in general, rather than project b into Col A (which is not easy, unless A has orthogonal columns), use the normal equations, $A^T A x = A^T b$ which any least-squares solution must satisfy (found using the orthogonal decomposition theorem).
Theorem: $A^T A$ is invertible if and only if A has linearly independent columns. When this is the case, the normal equations can be easily solved using the matrix inverse, thus $\hat{x} = (A^T A)^{-1} A^T b$.
Define: the least-squares error
When the columns of A are already orthogonal, $\hat{b}$ can be easily found by projecting onto Col A, and the least-squares solution is also easily obtained from the coefficients used in writing $\hat{b}$ as a linear combination of the columns of A, i.e. the j^{th} component of $\hat{x} = \frac{b \cdot a_j}{a_j \cdot a_j}$, where a_j is the j^{th} column of A.
Theorem: Given a QR factorization (from section 6.4) for an $m \times n$ matrix A, for any $b \in \mathbb{R}^n$ the least-squares solution to $Ax = b$ is given by $\hat{x} = R^{-1} Q^T b$. This works because $A\hat{x} = QQ^T b$, and since Q has orthonormal columns forming a basis for $\mathrm{Col\,A}, \mathrm{QQ^T b} = \mathrm{proj_{Col\,A} b} = \hat{\mathrm{b}}$.

Problems

1. Find the characteristic equation for the following matrix. Then solve for its eigenvalues. $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$

2. Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 7 & 0 \\ -2 & 0 & 4 \end{bmatrix}$, find the eigenvalues of A, and one eigenvector for each eigenvalue.

3. Suppose A and B are square matrices of the same size. Show that if x is an eigenvector of AB and $Bx \neq 0$, then Bx is an eigenvector of BA.

4. Let A and B be two $n \times n$ matrices. Given that $x \in \mathbb{R}^n$ is an eigenvector for A with eigenvalue λ_1 *AND* an eigenvector for B with eigenvalue λ_2, prove that x is also an eigenvector for both AB and BA with corresponding eigenvalue $\lambda_1\lambda_2$.

5. Show that $\{u_1, u_2, u_3\}$ is an orthogonal basis for $\mathbb{R}^3$, then express x as a linear combination in that basis, where $u_1 = \begin{bmatrix} 3 \\ -3 \\ 0 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}, \quad u_3 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}, \quad x = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$

6. Use the Gram-Schmidt process to find an orthogonal basis for the column space of the following matrix $\begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{bmatrix}$

7. Normalize the basis from the previous problem, then find a QR factorization of the matrix above. Hint: The columns of Q are the normalized vectors you just found, and $R = Q^TA$ (because $Q^TQ = I$, since Q has orthonormal columns). To simplify the calculation of R you may wish to factor out a $\frac{1}{\sqrt{12}}$ or $\frac{1}{\sqrt{3}}$ from Q^T before multiplying by A.

8. Find the projection of $y = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$ onto the line $x = y = z$. Hint: find a vector that spans this line.

9. Let U be an $n \times n$ orthogonal matrix, show that if $\{v_1, \ldots, v_n\}$ is an orthonormal basis for $\mathbb{R}^n$, then so too is $\{Uv_1, \ldots, Uv_n\}$.

10. Show that if U is an orthogonal matrix, then any real eigenvalue of U must be ± 1.

11. Suppose the columns of A are linearly independent, determine what happens to the least-squares solution $\hat{x}$ of $Ax = b$ if b is replaced by cb for a nonzero scalar c.