Study Guide for the MATH 2331 Final

All content contained within is originally from Linear Algebra with Applications, 5th Ed. by Otto Brescher

Typed up by Michael Wheeler (CCIS '21)

April 10, 2019

3 Subspaces of $\mathbb{R}^n$ and Their Dimensions

3.1 Image and Kernel of a Linear Transformation

The **image** of a function is the set of values that a function takes in its target space.

$$f : X \to Y \implies image(f) = \{f(x) \ \forall x \in X\}$$

The **span** of one or more vectors is the set of all linear combinations of those vectors.

$$\vec{v}_1 \dots \vec{v}_n \implies span(\{\vec{v}_1 \dots \vec{v}_n\}) = \{c_1 v_1 + \dots + c_n v_n \ \forall c_1 ... c_n \in \mathbb{R}\}$$

The image of a linear transformation is equal to the span of its columns.

$$T \text{ is linear}, \ T(\vec{x}) = A\vec{x} \implies image(T) = span(A)$$

The image of a linear transformation from m to n dimensions contains the zero vector in n, and is closed under addition and scalar multiplication.

The **kernel** or **nullspace** of a linear transformation is the set of all vectors that are "zeroes" of the transformation.

$$T : \mathbb{R}^m \to \mathbb{R}^n, \ T \text{ is linear} \implies kernel(T) = \{\vec{x} \in \mathbb{R}^m \mid T(\vec{x}) = \vec{0}_n\}$$

The kernel of a linear transformation from m to n dimensions contains the zero vector in m, and is closed under addition and scalar multiplication.

The kernel of a linear transformation from m to n dimensions contains only the zero vector if and only if the rank of its matrix is equal to m.

$$T : \mathbb{R}^m \to \mathbb{R}^n, \ T(\vec{x}) = A\vec{x} \implies kernel(T) = \{\vec{0}_m\} \iff rank(A) = m$$

If the kernel of a linear transformation contains only the zero vector, then the matrix has at most as many columns as rows.

$$A : \mathbb{R}^m \to \mathbb{R}^n,\;\; kernel(A) = \{\vec{0}_m\} \implies m \leq n$$

If an $n \times m$ matrix has more columns than rows, the kernel of the linear transformation contains nonzero vectors in $\mathbb{R}^m$.

$$A : \mathbb{R}^m \to \mathbb{R}^n,\;\; m > n \implies kernel(A) \neq \{\vec{0}_m\}$$

The kernel of a square matrix contains only the zero vector if and only if the matrix is invertible.

3.2 Subspaces of $\mathbb{R}^n$; Bases and Linear Independence

A subset of $\mathbb{R}^n$ is a (linear) **subspace** of $\mathbb{R}^n$ if it contains the zero vector in $\mathbb{R}^n$ and is closed under addition and multiplication.

For a linear transformation T from $\mathbb{R}^m$ to $\mathbb{R}^n$, $kernel(T)$ is a subspace of $\mathbb{R}^m$ and $image(T)$ is a subspace of $\mathbb{R}^n$.

A set of vectors $\vec{v}_1 \ldots \vec{v}_n$ is **linearly independent** if none of them can be expressed as a linear combination of the others.

To determine that a vector $\vec{b}$ is linearly independent of some other vectors $\left[\vec{v}_1 \ldots \vec{v}_n\right] = A$, use Gauss-Jordanian Elimination to show that the system $A\vec{x} = \vec{b}$ is inconsistent.

A set of vectors $\vec{v}_1 \ldots \vec{v}_n \in V$ forms a **basis** of the subspace V if the vectors are linearly independent and $span(\{\vec{v}_1 \ldots \vec{v}_n\}) = V$.

To construct the basis of $image(A)$, start with the column vectors of A and remove all the linearly dependent ("redundant") vectors.

A linear relation between the column vectors of an $n \times m$ matrix $A = \left[\vec{v}_1 \ldots \vec{v}_m\right]$ corresponds to an entry $kernel(A)$. This gives the following result:

$$\vec{v}_1 \ldots \vec{v}_m \text{ are linearly independent } \iff (kernel(A) = \vec{0}_m \iff rank(A) = m)$$

Based on a theorem from 3.1 this means that there are only n linearly independent vectors in $\mathbb{R}^n$.

A set of vectors $\vec{v}_1 \ldots \vec{v}_n \in V$ forms a **basis** of the subspace V iff all vectors in V can be expressed uniquely as a linear combination of those vectors:

$$\vec{v} = c_1\vec{v}_1 + \ldots + c_n\vec{v}_n \;\; \forall \vec{v} \in V$$

The coefficients $c_1 \ldots c_n$ are known as the **coordinates** of $\vec{v}$ with respect to the basis.

3.3 The Dimension of a Subspace of $\mathbb{R}^n$

All bases of a subspace V of $\mathbb{R}^n$ have the same number of vectors, known as the **dimension** $\boldsymbol{dim(V)}$ of V.

For any subspace V of $\mathbb{R}^n$ with $dim(V) = m$ the following is true:

- There are at most m linearly independent vectors in V.
- At least m vectors are required to span V.
- $\{\vec{v}_1 \dots \vec{v}_m\}$ are linearly independent $\iff$ $\{\vec{v}_1 \dots \vec{v}_m\}$ form a basis of V
- $span(\{\vec{v}_1 \dots \vec{v}_m\}) = V$ $\iff$ $\{\vec{v}_1 \dots \vec{v}_m\}$ form a basis of V

To find a basis of $\boldsymbol{kernel(A)}$ (and thus its dimension):

- Calculate $rref(A)$ and determine which columns don't have leading 1's.
- Parameterize those columns and solve for each entry in $\vec{x}$.
- Lastly, factor out the parameters. The resulting vectors form a basis of $kernel(A)$.

To find a basis of $\boldsymbol{image(A)}$ (and thus its dimension):

- Calculate $rref(A)$ and determine which columns *do* have leading 1's.
- Find the corresponding columns in A. Those column vectors form a basis of $image(A)$.

Theorem: For any matrix A, $dim(image(A)) = rank(A)$.

This all leads to the **Rank-Nullity Theorem**: for an $n \times m$ matrix A:

$$dim(image(A)) + dim(kernel(A)) = m$$

Theorem: The vectors $\vec{v}_1 \dots \vec{v}_n$ form a basis of $\mathbb{R}^n$ iff the matrix $\begin{bmatrix}\vec{v}_1 \dots \vec{v}_n\end{bmatrix}$ is invertible.

3.4 Coordinates

If we have a basis $\mathfrak{B} = (\vec{v}_1 \dots \vec{v}_m)$ of a subspace V of $\mathbb{R}^n$, then any vector in V can be written uniquely as $\vec{v} = c_1\vec{v}_1 + \dots + c_m\vec{v}_m$.

The vector $[\vec{x}]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ \dots \\ c_m \end{bmatrix}$ is then the $\mathfrak{B}$-coordinate vector of $\vec{x}$.

Then $\vec{x} = S[\vec{x}]_{\mathfrak{B}}$, where $S = \begin{bmatrix}\vec{v}_1 \dots \vec{v}_m\end{bmatrix}$ with dimensions $n \times m$.

For any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ and a basis $\mathfrak{B} = (\vec{v}_1 \dots \vec{v}_n)$ there is a matrix B that represents the transformation T in $\mathfrak{B}$:

$$\exists B : n \times n \text{ s.t. } [T(\vec{x})]_{\mathfrak{B}} = B[\vec{x}]_{\mathfrak{B}}$$

B can be constructed column-by-column:

$$B = \left[[T(\vec{v}_1)]_{\mathfrak{B}} \dots [T(\vec{v}_n)]_{\mathfrak{B}} \right]$$

Or B can also be constructed in terms of A (where $T(\vec{x}) = A\vec{x}$):

$$AS = BS, \;\; B = S^{-1}AS, \;\; A = SBS^{-1}$$

The middle formula $\boldsymbol{B = S^{-1}AS}$ is the form most commonly seen, and can be interpreted best when right-to-left:

- take the vector out of the basis $\mathfrak{B}$ (S)
- apply the transformation (A)
- and bring the vector back into $\mathfrak{B}$ (S^{-1})

If the above relationship holds, we say that A and B are **similar matrices**.

5 Orthogonality and Least Squares

5.1 Orthogonal Projections and Orthonormal Bases

Two vectors $\vec{u}, \vec{v}$ are **orthogonal** (perpendicular) if their dot product is equal to zero: $\vec{u} \cdot \vec{v} = 0$.

The **length** or **norm** of a vector $\vec{v}$ is the square root of the dot product with itself: $|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}$.

A **unit vector** $\vec{u}$ has length equal to 1. A unit vector can be obtained by multiplying by the reciprocal of the norm:

$$\vec{u} = \frac{1}{|\vec{v}|}\vec{v}$$

A set of vectors $\vec{u}_1 \dots \vec{u}_n$ is said to be **orthonormal** if they are all unit vectors and all orthogonal to one another. Put another way: $\vec{u}_i \cdot \vec{u}_j = 0 \;\; \forall i \neq j$

Orthonormal vectors are independent, and n orthonormal vectors will form an **orthonormal basis** of $\mathbb{R}^n$.

If V is a subspace of $\mathbb{R}^n$ with orthonormal basis $\vec{u}_1 \dots \vec{u}_m$, the orthogonal projection onto V is equivalent to projecting onto each component of V and adding the results:

$$proj_V(\vec{x}) = x^{\|} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \dots + (\vec{u}_m \cdot \vec{x})\vec{u}_m$$

A special case of this occurs when projecting onto $\mathbb{R}^n$ itself:

$$proj_{\mathbb{R}^n}(\vec{x}) = \vec{x} = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \dots + (\vec{u}_n \cdot \vec{x})\vec{u}_n$$

This tells us that the coordinates of $\vec{x}$ in an orthonormal basis can be found through the dot product: $c_i = \vec{u}_i \cdot \vec{x}$

The **orthogonal complement** of a subspace V is the set of vectors in $\mathbb{R}^n$ perpendicular to all vectors in V. It can be thought of as $V^\perp = kernel(proj_V(\vec{x}))$, and from the rank-nullity theorem we get $dim(V^\parallel) + dim(V^\perp) = n$.

5.2 Gram-Schmidt Process and QR Factorization

The **Gram-Schmidt Process** is an algorithm for taking any basis of V and transforming it into an orthonormal basis for the same subspace V.

We start with the first vector $\vec{v}_1$:

$$\vec{u}_1 = \frac{1}{|\vec{v}_1|}\vec{v}_1$$

Then for each vector following $(\vec{v}_2 \ldots \vec{v}_n)$, we find the component of that vector perpendicular to all unit vectors found so far:

$$\vec{v}_n^\perp = \vec{v}_n - (\vec{u}_1 \cdot \vec{v}_n)\vec{u}_1 - \ldots - (\vec{u}_{n-1} \cdot \vec{v}_n)\vec{u}_{n-1}$$

And then reduce that component to a unit vector by dividing out its magnitude:

$$\vec{u}_n = \frac{1}{|\vec{v}_n^\perp|}\vec{v}_n^\perp$$

This is known as **QR Factorization** because the matrix $\begin{bmatrix}\vec{v}_1 \ldots \vec{v}_n\end{bmatrix}$ is factored into two matrices Q and R, where $Q = \begin{bmatrix}\vec{u}_1 \ldots \vec{u}_n\end{bmatrix}$ and $R_{i,j} = \begin{cases} (\vec{u}_i \cdot \vec{v}_j), & i \neq j \\ |\vec{v}_j^\perp|, & i = j \end{cases}$.

This allows the first column of R to be computed, then the first column of Q, followed by the second column of R, then the second column of Q, ... Note that R is upper triangular.

5.3 Orthogonal Transformations and Orthogonal Matrices

A linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ is **orthogonal** if it preserves the length of its inputs:

$$|T(\vec{x})| = |\vec{x}| \;\; \forall \vec{x} \in \mathbb{R}^n$$

The matrix of an orthogonal linear transformation is called an **orthogonal matrix**.

An orthogonal transformation T preserves orthogonality: if two vectors $\vec{v}, \vec{w}$ are orthogonal, so are $T(\vec{v}), T(\vec{w})$.

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal iff the column vectors of A form an orthonormal basis of $\mathbb{R}^n$.

The product of two orthogonal matrices is also orthogonal. The inverse of an orthogonal matrix is also orthogonal.

A matrix is **symmetric** if $A^T = A$, and **skew-symmetric** if $A^T = -A$

The dot product can be expressed as a matrix product using the transpose: $\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w}$

A matrix A is orthogonal iff $A^T A = I_n$ or $A^T = A^{-1}$.

If V is a subspace of $\mathbb{R}^n$ with orthonormal basis $\vec{u}_1 \dots \vec{u}_n$, the matrix for the orthogonal projection onto V is $P = QQ^T$ where $Q = \begin{bmatrix} \vec{u}_1 \dots \vec{u}_n \end{bmatrix}$

5.4 Least Squares and Data Fitting

TODO – write me!

6 Determinants

6.1 Introduction to Determinants

The **determinant** of a matrix tells if that matrix is invertible: for a matrix A, A^{-1} exists iff $det(A) \neq 0$.

The determinant of a triangular matrix (upper or lower) is the product of its diagonal entries.

Recall that the determinant of a 2×2 matrix is $ad - bc$, and its inverse is $\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

6.2 Properties of the Determinant

If A is a square matrix then $det(A^T) = det(A)$

The following properties of the determinant are useful when doing Gauss-Jordanian Elimination:

- If B comes from dividing a row of A by a scalar k then $det(B) = \frac{1}{k} det(A)$
- If B comes from a row swap in A then $det(B) = -det(A)$
- If B comes from adding a multiple of a row to another in A then $det(B) = det(A)$

To calculate the determinant of an $n \times n$ matrix A:

- Perform row reductions on A until you reach $rref(A)$, keeping a careful tally of which operations you perform at each step.

- Calculate $det(rref(A))$ by multiplying the diagonal entries of $rref(A)$
- Using the rules from above, work backwards from $det(rref(A))$ to reach $det(A)$.

For two $n \times n$ matrices A, B it is true that $det(AB) = det(A)det(B)$, and for a positive integer m it is true that $det(A^m) = det(A)^m$

The determinant is independent of basis: similar matrices have the same determinant.

If A is invertible then $det(A^{-1}) = \frac{1}{det(A)}$

7 Eigenvalues and Eigenvectors

7.1 Diagonalization

A matrix is **diagonal** if there are nonzero entries only along the diagonal of the matrix.

The matrix A of the linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be **diagonalizable** if the matrix B of T with respect to $\mathfrak{B}$ is diagonal.

To diagonalize a matrix, one finds an invertible matrix S and a diagonal matrix B such that the change-of-bases formula holds: $B = SAS^{-1}$.

A nonzero vector $\vec{v}$ is an **eigenvector** of A if $A\vec{v} = \lambda\vec{v}$. The value λ is the **eigenvalue** of eigenvector $\vec{v}$. One or more eigenvectors $\vec{v}_1 \ldots \vec{v}_n$ of A form an **eigenbasis** of A.

If 0 is an eigenvalue of A then $kernel(A) \neq \{\vec{0}\}$ and A is not invertible.

A **discrete linear dynamical system** is a system where the vector $\vec{x}$ is a function of time t. Specifically $\vec{x}(t)$ is a linear transformation from $t-1$ to t such that:

$$\vec{x}(t) = A\vec{x}(t-1) = A^t\vec{x}_0$$

.
When dealing with discrete linear dynamical systems the goal is often to find a nonrecursive or **closed formula** for $\vec{x}(t)$: that is, a formula for $\vec{x}(t)$ in terms of t alone.

To solve for the closed formula of a discrete linear dynamical system:

- Find an eigenbasis $\vec{v}_1 \ldots \vec{v}_n$ of the transformation matrix A.
- Express $\vec{x}_0$ as coordinates in that eigenbasis: $c_1 \ldots c_n$.
- $\vec{x}(t) = c_1\lambda_1^t\vec{v}_1 + \ldots + c_n\lambda_n^t\vec{v}_n$

7.2 Finding the Eigenvalues of a Matrix

To determine the eigenvalues of a matrix, solve the following equation:

$$det(A - \lambda I_n) = 0$$

The above is known as the **characteristic equation** of the matrix A. It gives rise to the **characteristic polynomial** of the matrix, the roots of which are the eigenvalues of the matrix.

The **algebraic multiplicity** of an eigenvalue is equal to the multiplicity of its root of the characteristic polynomial.

The eigenvalues of a triangular matrix are its diagonal entries.

7.3 Finding the Eigenvectors of a Matrix

An **eigenspace** associated with eigenvalue λ is:

$$E_\lambda = kernel(A - \lambda I_n)$$

The eigenvectors with eigenvalue λ are the nonzero vectors in E_λ. Thus finding the eigenvectors of λ is equivalent to finding a basis of the eigenspace E_λ.

To determine the eigenvectors of a matrix:
For each eigenvalue λ:

- Find the eigenspace for λ by solving for the vectors that span $kernel(A - \lambda I_n)$.
- Recall from section 3 that you have to for $rref(A - \lambda I_n)$
- Then you have to and find the relations between the columns by parameterizing the columns without leading 1s.

The **geometric multiplicity** of an eigenvalue λ is the dimension of its associated eigenspace E_λ, and according to the Rank-Nullity theorem is equal to $n - rank(A - \lambda I_n)$.

An $n \times n$ matrix is diagonalizable iff the geometric multiplicities of the eigenvalues add up to n.

If matrix A is similar to matrix B:

- The two matrices have the same characteristic polynomial.
- $rank(A) = rank(B)$ and $nullity(A) = nullity(B)$
- The two matrices have the same eigenvalues, but NOT necessarily the same eigenvectors.

- The two matrices have the same determinant and the same trace.

So to review:

- Solve for eigenvalues: $det(A - \lambda I_n) = 0$
- For each eigenvalue, find a basis of the eigenspace $kernel(A - \lambda I_n)$
- Determine that the dimensions of the bases of all the eigenspaces add up to n. If so, concatenate those bases together to form S.

7.4 Dynamical Systems

According to Prof. Antunes this section is NOT going to be covered on the final.

8 Symmetric Matrices and Quadratic Forms

8.1 Symmetric Matrices

A matrix is **orthogonally diagonalizable** if it has an eigenbasis that is orthonormal.

A matrix is orthogonally diagonalizable if and only if it is symmetric.

$$\exists S \mid S^{-1}AS \text{ is diagonal } \iff A^T = A$$

If A is a symmetric matrix, and A has eigenvectors $\vec{v}_1$ and $\vec{v}_2$ with eigenvalues $\lambda_1 \neq \lambda_2$, then $\vec{v}_2$ is orthogonal to $\vec{v}_1$.

A symmetric $n \times n$ matrix A has n real eigenvalues if they are counted with their algebraic multiplicities.

To orthogonally diagonalize a symmetric matrix A, perform the following steps on A:

- Find the eigenvalues of A.
- For each eigenvalue, find a basis of its eigenspace.
- Use the Gram-Schmidt Process to find an orthonormal basis of each eigenspace.
- Concatenate the orthonormal bases into the matrix to make an orthonormal eigenbasis of A called S: the matrix S will be orthogonal and $S^{-1}AS$ will be diagonal.

8.2 Quadratic Forms

A **quadratic form** is a function $q(x_1 \ldots x_n) : \mathbb{R}^n \to \mathbb{R}$ that is a linear combination of functions of the form $x_i x_j$, where i can be equal to j.

A quadratic form can be expressed as $q(\vec{x}) = \vec{x} \cdot A\vec{x} = \vec{x}^T A\vec{x}$. Note that A is "symmetric by design".

If $\mathfrak{B}$ is an orthonormal eigenbasis for the above symmetric A with eigenvalues $\lambda_1 \ldots \lambda_n$, and if $c_1 \ldots c_n$ are the coordinates of $\vec{x}$ with respect to $\mathfrak{B}$, then:

$$q(\vec{x}) = \lambda_1 c_1^2 + \ldots + \lambda_n c_n^2$$

This is referred to as **diagonalizing the quadratic form**.

For a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, we say that A is:

- **positive definite** if $q(\vec{x}) > 0 \;\; \forall \vec{x} \in \mathbb{R}, \vec{x} \neq \vec{0}$
- **positive semidefinite** if $q(\vec{x}) \geq 0 \;\; \forall \vec{x} \in \mathbb{R}$
- **negative definite** if $q(\vec{x}) < 0 \;\; \forall \vec{x} \in \mathbb{R}, \vec{x} \neq \vec{0}$
- **negative semidefinite** if $q(\vec{x}) \leq 0 \;\; \forall \vec{x} \in \mathbb{R}$
- **indefinite** if $q(\vec{x})$ can take both positive and negative values.

A symmetric matrix A is positive definite iff all of its eigenvalues are positive, and positive semidefinite iff all of its eigenvalues are positive or zero.

8.3 Singular Values

TODO – write me!