Reference Sheet for CO145 Mathematical Methods

Autumn 2016

Part I

Analysis

1 Sets

For a set S of real numbers:

1. u is an upper bound if $u \geq s \, \forall \, s \in S$. If such a u exists, S is bounded above.
2. l is an lower bound if $l \leq s \, \forall \, s \in S$. If such a l exists, S is bounded below.

Every set, S has a supremum and infimum:

1. $\sup(S)$ is the least upper bound of S.
2. $\inf(S)$ is the greatest lower bound of S.

Proving the Convergence of a Bounded Sequence. Every *increasing* sequence of real numbers that is *bounded above* must converge (this is especially useful when combined with proof by induction).

2 Sequences

2.1 Given Results

2.1.1 Absolute Values

1. $|xy| = |x| \times |y|$
2. $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$

2.1.2 Converging Sequences

1. $\lim_{n\to\infty} \frac{1}{n^c} = 0$ for all $c > 0$
2. $\lim_{n\to\infty} \frac{1}{c^n} = 0$ for all $|c| > 0$
3. $\lim_{n\to\infty} \frac{1}{n!} = 0$
4. $\lim_{n\to\infty} \frac{1}{\log n} = 0$ for $n > 1$

2.1.3 Combinations of Sequences

1. $\lim_{n\to\infty} (\lambda a_n) = \lambda \lim_{n\to\infty} a_n$
2. $\lim_{n\to\infty} (a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$
3. $\lim_{n\to\infty} (a_n - b_n) = \lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n$
4. $\lim_{n\to\infty} (a_n \times b_n) = \lim_{n\to\infty} a_n \times \lim_{n\to\infty} b_n$
5. $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}$

2.2 Proving Convergence

2.2.1 Direct Proof

1. Apply the definition of convergence:

   Suppose a_n converges to L. Then there must exist an $N(\epsilon) \in \mathbb{R}$ such that for all $n \geq N(\epsilon)$ and for all $\epsilon > 0$, $*$ holds:

   $$|a_n - L| < \epsilon *$$

2. Rearrange $*$ such that n is the subject, ensuring that you show implications in the correct direction.
3. Propose an $N(\epsilon)$ such that $*$ holds (the ceiling function applied to the previous result is sufficient).
4. Check that this result is sensible.

2.2.2 Sandwich Theorem

1. Show that $\lim_{n\to\infty} l_n = L$.

2. Show that $\lim_{n\to\infty} u_n = L$.

3. Show that $l_n \le a_n$ and $u_n \ge a_n$ for all $n \ge N$ for some $N \in \mathbb{R}$.

4. Apply the sandwich theorem to show that $\lim_{n\to\infty} a_n = L$.

2.2.3 Ratio Test

1. Determine the value of $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$.

2. Conclude appropriately.

   (a) If $r < 1$ then a_n converges to 0.

   (b) If $r > 1$ then a_n diverges.

Proving the convergence of a sequence b_n to $L \neq 0$ Consider the ratio of the modified sequence $a_n = b_n - L$.

3 Series

3.1 Given Results

3.1.1 Converging Series

1. $\sum_{n=1}^{\infty} x^n$ converges to $\frac{x}{1-x}$ for $|x| < 1$. [Geometric Series]

2. $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges to $\frac{\pi}{6}$. [Inverse Square Series]

3. $\sum_{n=1}^{\infty} \frac{1}{n^c}$ converges for $c > 1$.

3.1.2 Diverging Series

1. $\sum_{n=1}^{\infty} x^n$ diverges for $|x| > 1$. [Geometric Series]

2. $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. [Harmonic Series]

3. $\sum_{p:\text{prime}} \frac{1}{p}$ diverges. [Harmonic Primes]

3.2 Proving Convergence

3.2.1 Using Partial Sums

1. Construct the partial sum, $S_n = \sum_{i=1}^{n} a_i$.

2. Prove the convergence or divergence of the partial sum.

3.2.2 Comparison Test

Standard Comparison Test

1. Show that $\sum_{i=1}^{\infty} c_i$ converges or that $\sum_{i=1}^{\infty} d_i$ diverges.

2. Make a comparison and conclude appropriately.

   (a) Show that $a_i \le \lambda c_i$ for some $\lambda > 0$ and for all $i > N \in \mathbb{R}$ and conclude $\sum_{n=1}^{\infty} a_i$ converges.

   (b) Show that $a_i \ge \lambda d_i$ for some $\lambda > 0$ and for all $i > N \in \mathbb{R}$ and conclude $\sum_{n=1}^{\infty} a_i$ diverges.

Limit Comparison Test

1. Show that $\sum_{i=1}^{\infty} c_i$ converges or that $\sum_{i=1}^{\infty} d_i$ diverges.

2. State the appropriate limit and conclusion.

   (a) Show that $\lim_{i\to\infty} \frac{a_i}{c_i}$ exists and conclude $\sum_{n=1}^{\infty} a_i$ converges.

   (b) Show that $\lim_{i\to\infty} \frac{d_i}{a_i}$ exists and conclude $\sum_{n=1}^{\infty} a_i$ diverges.

3.2.3 Ratio Test

1. Determine $\lim_{n\to\infty} \frac{a_{i+1}}{a_i}$.

2. Conclude appropriately.

   (a) If $\lim_{n\to\infty} \frac{a_{i+1}}{a_i} < 1$ then $\sum_{i=1}^{\infty} a_i$ converges.

   (b) If $\lim_{n\to\infty} \frac{a_{i+1}}{a_i} > 1$ then $\sum_{i=1}^{\infty} a_i$ diverges.

3.2.4 Integral Test

Where $f(n) = a_n$ is a continuous, positive, and decreasing function:

1. Determine $\int_N^\infty f(x)\,\mathrm{d}x$.
2. Conclude appropriately.
   (a) If $\int_N^\infty f(x)\,\mathrm{d}x$ converges, so does $\sum_{n=N}^\infty a_n$.
   (b) If $\int_N^\infty f(x)\,\mathrm{d}x$ diverges, so does $\sum_{n=N}^\infty a_n$.

4 Power Series

4.1 Given Results

1. $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^r}{r!} + \ldots$ for all x
2. $\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{r+1}\frac{x^r}{r} + \ldots$ for $-1 < x \leq 1$
3. $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^r \frac{x^{2r+1}}{(2r+1)!} + \ldots$ for all x
4. $\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^r \frac{x^{2r}}{(2r)!} + \ldots$ for all x

Combinations of Functions Note that you can add, subtract, multiply, divide, differentiate and integrate power series term-wise.

4.2 Determining a Maclaurin Series

1. Repeatedly differentiate $f(x)$ and set x to 0. Use these results to propose a value for $f^{(n)}0$.
2. The Maclaurin series expansion for $f(x)$ is then given by $\sum_{n=0}^\infty f^{(n)}(0)\frac{x^n}{n!}$.

4.3 Determining a Taylor Series

1. Repeatedly differentiate $f(x)$ and set x to a. Use these results to propose a value for $f^{(n)}a$.
2. The Taylor series expansion at a for $f(x)$ is then given by $\sum_{n=0}^\infty \frac{f^{(n)}(a)}{n!}(x-a)^n$.

4.4 Finding the Radius of Convergence

Radius of Convergence: Size of x set for which a power series converges.

1. Apply a ratio test using absolute values.
2. Set $|r| < 1$ to show the region of convergence.
3. If the series converges for $|x-a| < R$, the series has a radius of convergence of R (about a).

4.5 Error Terms

Where c is a constant that lies between x and a:

4.5.1 Lagrange Error Term

$$\begin{aligned} f(x) &= \sum_{n=0}^\infty \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &= \sum_{n=0}^k \frac{f^{(n)}(a)}{n!}(x-a)^n + \underbrace{\frac{f^{(k+1)}(c)}{(k+1)!}(x-a)^{k+1}}_{\text{Lagrange Error Term}} \end{aligned}$$

This result follows from the mean value theorem.

4.5.2 Cauchy Error Term

$$f(x) = \sum_{n=0}^k \frac{f^{(n)}(a)}{n!}(x-a)^n + \underbrace{\frac{f^{(k+1)}(c)}{k!}(x-c)^k(x-a)}_{\text{Cauchy Error Term}}$$

4.6 Using Power Series to Solve ODEs

Example: Given the differential equation, $\frac{\mathrm{d}y}{\mathrm{d}x} = ky$:

1. Express y as the power series, $\sum_{i=0}^\infty a_i x^i$.
2. Differentiate y to give $\frac{\mathrm{d}y}{\mathrm{d}x} = \sum_{i=0}^\infty i a_i x^{i-1}$.
3. Compare coefficients to give $a_i = \frac{k}{i}a_{i-1}$.
4. Deduce that $a_i = \frac{k^i}{i!}a_0$. Hence $y = a_0 e^{kx}$ (where a_0 is the value of x at $y = 0$).

Part II

Linear Algebra

5 Introduction

1. *Vectors* can be added together or multiplied by scalars to produce another vector.

2. *Examples of vector objects*: Geometric vectors, polynomials, $\mathbb{R}^n$, audio signals.

6 Groups

For $(G, \cdot)$ to be called a group, it must have the properties:

1. **Closure:** $\forall\, x, y \in G,\ x \cdot y \in G$

2. **Associativity:** $\forall\, x, y, z \in G,\ (x \cdot y) \cdot z = x \cdot (y \cdot z)$

3. **Identity:** $\exists\, e \in G$ s.t. $\forall\, x \in G,\ x \cdot e = e \cdot x = x$

4. **Inverse:** $\forall\, x \in G\, \exists\, x^{-1} \in G$ s.t. $x \cdot x^{-1} = x^{x-1} \cdot x = e$

For the group to be *abelian* it must also have the additional property:

1. **Commutativity:** $\forall\, x, y \in G,\ x \cdot y = y \cdot x$

7 Methods on Matrices

7.1 Multiplication by a Scalar

Has the properties:

1. Associativity

2. Distributivity

3. $(\lambda \mathbf{C})^\top = \lambda \mathbf{C}^\top$

7.2 Matrix Multiplication

Matrix multiplication is defined for $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{B} \in \mathbb{R}^{n \times p}$, giving the product, $\mathbf{C} = \mathbf{AB} \in \mathbb{R}^{m \times p}$. It has the properties:

1. Associativity

2. Distributivity

3. $\forall\, \mathbf{A} \in \mathbb{R}^{m \times n},\ \mathbf{I}_m \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}$.

4. Note that $\mathbf{AB} \neq \mathbf{BA}$.

7.3 Inverse and Transpose

7.3.1 Further Properties of Matrix Arithmetic

1. $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$

2. Note $(\mathbf{A} + \mathbf{B})^{-1} \neq \mathbf{A}^{-1} + \mathbf{B}^{-1}$

3. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$

4. $\left(\mathbf{A}^\top\right)^\top = \mathbf{A}$

5. $(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top$

6. $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$

7. $\left(\mathbf{A}^{-1}\right)^\top = \left(\mathbf{A}^\top\right)^{-1}$

7.3.2 Determining the Inverse of a Matrix

1. Recall that $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. We therefore wish to solve $\mathbf{AX} = \mathbf{I}$.

2. Starting with $[\mathbf{A}|\mathbf{I}_n]$, use Gaussian elimination to reach reduced row echelon form.

3. $\mathbf{A}^{-1}$ can be read off from the result $\left[\mathbf{I}_n|\mathbf{A}^{-1}\right]$.

7.4 Gaussian Elimination

7.4.1 Elementary Transformations

We are allowed to:

1. Swap two rows.
2. Multiply a row by a constant $\lambda \neq 0$.
3. Add a row to another row.

We use these transformations to reach (reduced) row echelon form.

7.4.2 Row Echelon Form

1. The pivot of a non-zero row is strictly to the right of a pivot of the row above it.
2. Any rows containing only zeros are at the bottom of the matrix.

Columns with pivots define *basic variables*, other columns give *free variables*.

7.4.3 Reduced Row Echelon Form

1. The system is in row echelon form.
2. Every pivot is 1.
3. The pivot is the only non-zero entry in its column.

7.5 Rank and Determinant

7.5.1 Finding the Rank of a Matrix

Rank: number of linearly independant columns of a matrix.

1. Apply Gaussian elimination to reach RREF.
2. The rank is given by the number of pivots (linearly independent cols / rows).

Points to Note

1. Column rank is equal to row rank.
2. $\mathbf{A}$ is regular (invertible) $\iff$ $\mathrm{rk}(\mathbf{A}) = n$.
3. A matrix has full rank if its rank is equal to the lesser of the number of rows and columns, or is rank-deficient otherwise.

7.5.2 Finding the Determinant of a Matrix

1. By the given transformations, bring the matrix into triangular form (all 0 above or below the leading diagonal) or into a 2×2 form.
2. The determinant is given the product of the leading diagonal of a triangular matrix, or the product of the leading diagonal minus the product of the other elements in a 2×2 matrix.

Transformations

1. Adding and subtracting multiples of cols or rows does not change the determinant.
2. Multiplying a col or row by by a constant scales the determinant by that constant.
3. Swapping rows or cols changes the sign.
4. **Laplace expansion** can be used to reduce no. of rows and cols by one. Best used when there is a row or col with only one non-zero entry.

Points to Note

1. $\det(\mathbf{AB}) = \det(\mathbf{A}) \times \det(\mathbf{B})$
2. $\det(\mathbf{A}) = 0 \iff \mathbf{A}$ is singular
3. $\det(\mathbf{A}) = \det\left(\mathbf{A}^\top\right)$
4. $\det\left(\mathbf{A}^{-1}\right) = 1/\det(\mathbf{A})$
5. Similar matrices have the same determinant.

8 Linear Equation Systems

8.1 Solving Inhomogeneous Linear Equation Systems

For $\mathbf{Ax} = \mathbf{b}$:

1. Build an augmented matrix from the system of equations.
2. Use elementary transformations to reach row echelon form (you *must* justify every step).

3. By reading from the resulting rows and setting free variables = 0, find a particular solution.
4. Solve the homogeneous linear equation system $\mathbf{Ax} = \mathbf{0}$, using the row echelon form derived in 2.
5. Combine the solutions from 3 and 4 to form the general solution.

8.2 Solving Homogeneous Linear Equation Systems

For $\mathbf{Ax} = \mathbf{0}$:

1. Use elementary transformations to reach row echelon form as before.
2. For each free variable (non-pivot col), equate it to a sum of basic variables (pivot cols).
3. Rearrange the equations formed in 1 so that they are of the form something = 0.
4. The solutions are given by the column vector of the coefficients of the variables, multiplied by any real scalar value.

8.2.1 The Minus-1 Trick

1. Use Guassian elimination to reach reduced row echelon form.
2. Extend the matrix from 1 by adding rows of the form $[0 \ldots 0\ 1\ 0 \ldots 0]$ such that the leading diagonal is made up entirely of 1 (pivots) or -1 (from the introduced rows).
3. The columns containing -1 in the diagonal form the solutions.

9 Vector Spaces

9.1 Defining Vector Spaces

9.1.1 Vector Spaces

A vector space is a set V with two operations:

1. $V + V \to V$ (inner operation)
2. $\mathbb{R} \cdot V \to V$ (outer operation)

where:

1. $(V, +)$ is an Abelian group and has the distributivity property.
2. The outer operation has distributivity and associativity properties and has a neutral element of 1.

9.1.2 Vector Subspaces and Generating Sets

For a vector space, V:

1. If every vector in V can be expressed as a linear combination of $A\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, then A is a *generating set* for V.
2. For $U \subset V$ and $U \neq \emptyset$, U is a *vector subspace* of V if U is a vector space.

Proving a Set is a Vector Subspace For a vector subspace U, we need to show:

1. $U \neq \emptyset$. Equivalently, $0 \in U$.
2. Closure of U with respect to the inner operation: $\forall \mathbf{x}, \mathbf{y} \in U\, (\mathbf{x} + \mathbf{y} \in U)$.
3. Closure of U with respect to the outer operation: $\forall \lambda \in \mathbb{R} \forall \mathbf{x} \in U\, (\lambda \mathbf{x} \in U)$.

9.2 Determining Linear Dependence or Independence

Linear Dependence: For a vector space $\mathbf{x}_1, \ldots, \dot{\mathbf{x}_k}$ here is a non-trivial linear combination such that $\mathbf{0} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i$.
To prove $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly independent:

1. Write the vectors as columns of a matrix.
2. Apply Gaussian elimination to reach RREF.
3. Conclude appropriately:
   (a) Pivot columns are linearly independent of the previous vectors.
   (b) Non-pivot columns can be expressed as linear combinations of previous pivot columns.

Points to Note

1. If at least one of the vectors is $\mathbf{0}$ or at least two of the vectors are identical then they are linearly dependent.
2. The set of vectors, $\mathbf{x}_1 = \sum_{i=1}^{k} \lambda_{i1} \mathbf{b}_i, \ldots, \mathbf{x}_m = \sum_{i=1}^{k} \lambda_{im} \mathbf{b}_i$ (where $\mathbf{b}_1, \ldots, \mathbf{b}_k$ are linearly independent) are linearly independent if and only if $\lambda_1, \ldots, \lambda_m$ are linearly independent.

9.3 Determining Bases and Dimensions

9.3.1 Determining a Basis and Dimension

Basis: Minimal (linearly independent) generating set for V. It can be determined as follows:

1. Write the spanning vectors as columns of a matrix.
2. Apply Gaussian elimination to reach RREF.
3. The original values of the pivot columns form a basis.
4. The *dimension* is the number of basis vectors.

9.3.2 Determining a Simple Basis

1. Write the spanning vectors as rows of a matrix.
2. Apply Gaussian elimination to reach RREF.
3. The rows with leading ones form a simple basis.

9.3.3 Determining a Basis of the Intersection of Subspaces

For $U_1 = [\mathbf{b}_1, \ldots, \mathbf{b}_k]$ and $U_2 = [\mathbf{c}_1, \ldots, \mathbf{c}_l]$:

1. Find the respective bases of U_1 and U_2.
2. We want to solve $\sum_{i=1}^{k} \lambda_i \mathbf{b}_i = \sum_{j=1}^{l} \mu_j \mathbf{c}_j$, i.e. $\sum_{i=1}^{k} \lambda_i \mathbf{b}_i - \sum_{j=1}^{l} \mu_j \mathbf{c}_j = \mathbf{0}$.
3. This can be solved by the method for homogeneous linear equations, where $\mathbf{b}_1, \ldots, \mathbf{b}_k, -\mathbf{c}_1, \ldots, -\mathbf{c}_l$ form the columns of the augmented matrix.
4. Solve for either $\lambda_1, \ldots, \lambda_k$ or $\mu_1, \ldots, \mu_l$ and determine the basis accordingly.

The *linear hull* is the intersection of a set of subspaces.

9.4 Affine Spaces

9.4.1 Defining Affine Spaces

Affine spaces can be defined as:

1. $L = \mathbf{x}_0 + U$
2. Parametric Equation: $\exists \lambda_1 \ldots \lambda_k \forall \mathbf{x} \in L$ such that $\mathbf{x} = \mathbf{x}_0 + \lambda_1 \mathbf{b}_1 + \cdots + \lambda_k \mathbf{b}_k$

9.4.2 Finding the Intersection of Affine Spaces

1. Recall that for $\mathbf{x} \in L_1$ and $\mathbf{x} \in L_2$, $\mathbf{x}_1 + \sum_{i=1}^{k} \lambda_i \mathbf{b}_i = \mathbf{x} = \mathbf{x}_2 + \sum_{j=1}^{l} \mu_j \mathbf{c}_j$. Hence $\sum_{i=1}^{k} \lambda_i \mathbf{b}_i - \sum_{j=1}^{l} \mu_j \mathbf{c}_j = \mathbf{x}_2 - \mathbf{x}_1$. This can be solved by the method for inhomogeneous linear equations, where the basis vectors $\mathbf{b}_1, \ldots, \mathbf{b}_k, -\mathbf{c}_1, \ldots, -\mathbf{c}_l$ form the columns of the augmented matrix.
2. Determine the basis vectors.
3. Solve the resulting inhomogeneous LEQS.
4. Use the solution to determine a value for $\mathbf{x}$ using one of the original equations.
5. You can check your answer with the other equation.

9.4.3 Determining Parallelism

For $L_1 = \mathbf{x}_1 + U_1$ and $L_2 = \mathbf{x}_2 + U_2$, $L_1 || L_2$ if $U_1 \subseteq U_2$ or $U_2 \subseteq U_1$.

10 Linear Mappings

10.1 Defining Linear Mappings

To prove a mapping Φ is linear (a homomorphism), we must show that:

1. $\Phi(\mathbf{x} + \mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y})$
2. $\Phi(\lambda \mathbf{y}) = \lambda \Phi(\mathbf{x})$

Special Cases

1. **Isomorphism:** also bijective.
2. **Endomorphism:** also maps from V to V.
3. **Automorphism:** also maps from V to V and bijective.

Points to Note

1. For linear mappings $\Phi : V \to W$ and $\Psi : W \to X$, the mapping $\Psi \circ \Phi : V \to X$ is also linear.
2. If Φ is an isomorphism, then so is Φ^{-1}.
3. If Φ and Ψ are linear, then so are $\Phi + \Psi$ and $\lambda \Phi$.

10.2 Image and Kernel (Null Space)

For a mapping $\Phi : \mathbf{x} \in V \to \mathbf{A}\mathbf{x} \in W$:

Determining the Image *Image*: $\{\mathbf{w} \in W | \exists \mathbf{v} \in V : \Phi(\mathbf{v}) = \mathbf{w}\}$: Set of vectors in W that can be reached by Φ from any vector in V.

Return the column space of $\mathbf{A}$ (find its basis).

Determining the Kernel *Kernel*: $\{\mathbf{v} \in V | \Phi(\mathbf{v}) = \mathbf{0}_W\}$: Set of vectors in V that Φ maps onto the neutral element in W. *Note*: If kernel is $\{\mathbf{0}\}$, Φ is injective.

Return the solution to the LEQS $\mathbf{A}\mathbf{x} = \mathbf{0}$.

Rank-Nullity Theorem For $\Phi : V \to W$, $\dim(\mathrm{Im}(\Phi)) = \dim(V) - \dim(\ker(\Phi))$.

10.3 Matrix Representation

For the mapping $\Phi : V \to W$, and ordered bases $B \subseteq V$ and $C \subseteq W$, the *transformation matrix* $\mathbf{A}_\Phi$ is defined such that for the coordinates $\hat{\mathbf{x}}$ of $\mathbf{x} \in V$ with respect to B and $\hat{\mathbf{y}}$ of $\Phi(\mathbf{x}) \in W$ with respect to C: $\hat{\mathbf{y}} = \mathbf{A}_\Phi \hat{\mathbf{x}}$.

10.4 Basis Change

Given an $\mathbf{A}_\Phi$ with respect to bases B and C, we want an $\tilde{\mathbf{A}}_{\boldsymbol{\Phi}}$ w.r.t. $\tilde{B}$ and $\tilde{C}$:

1. Write the vectors of $\tilde{B}$ as a linear combination of the vectors of B. These form the columns of the matrix $\mathbf{S}$.

2. Write the vectors of $\tilde{C}$ as a linear combination of the vectors of C. These form the columns of the matrix $\mathbf{T}$.

3. $\tilde{\mathbf{A}}_{\boldsymbol{\Phi}}$ can be calculated by $\tilde{\mathbf{A}}_{\boldsymbol{\Phi}} = \mathbf{T}^{-1}\mathbf{A}_\Phi\mathbf{S}$.

This can be derived by considering the composition of the required linear mappings.

Points to Note

1. $\mathbf{A}$ and $\tilde{\mathbf{A}}$ are *equivalent* if $\tilde{\mathbf{A}}$ can be expressed as $\tilde{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{S}$.

2. $\mathbf{A}$ and $\tilde{\mathbf{A}}$ are *similar* if $\tilde{\mathbf{A}}$ can be expressed as $\tilde{\mathbf{A}} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$.

10.5 Eigenvalues

For an endomorphism $\Phi : V \to V$, λ is an *eigenvalue* if there exists an $\mathbf{x} \in V \setminus \{\mathbf{0}\}$ s.t. $\Phi(\mathbf{x}) = \lambda\mathbf{x}$. $\mathbf{x}$ is the corresponding eigenvector.

For a transformation matrix **A:**

10.5.1 Determining the Spectrum (Eigenvalues)

1. Calculate the determinant, $|\mathbf{A} - \lambda\mathbf{I}|$.

2. Solve (equal to 0) the result (the *characteristic polynomial*) for λ.

3. The eigenvalues of $\mathbf{A}$ are given by the solutions.

10.5.2 Determining the Corresponding Eigenspaces

1. For each eigenvalue λ, find the solutions to the LEQS $(\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}$.

10.5.3 Applications

1. Used in principal component analysis (principle components have largest eigenvalues) for dimensionality reduction in machine learning applications.

2. Used to determine the theoretical limit to how much information can be transferred through a communication medium by calculating eigenvalues and eigenvectors of communication channel.

3. Used in the PageRank algorithm to determine the rank of a page for a search (based on maximal eigenvalue).

4. Determine numerical stability, e.g. when inverting matrices, by looking at condition numbers (ratio of biggest to smallest eigenvalue).

10.6 Diagonalization

10.6.1 Determining Diagonalizability

1. The characteristic polynomial must decompose into linear factors (and the sign must be correct).

2. The dimension of each eigenspace must be equal to the power (algebraic multiplicity) of its respective factor in the characteristic polynomial.

10.6.2 Diagonalization

1. Determine the eigenspaces of the given matrix $\mathbf{A}$.
2. Collect the eigenvectors in a single matrix $\mathbf{S}$.
3. The diagonalization of $\mathbf{A}$ is given by $\mathbf{D} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$. This corresponds to a matrix with the eigenvalues of $\mathbf{A}$ along the diagonal.

10.6.3 Applications

1. Diagonal matrices are easily raised to a power.
2. Decouple variables - uses in probability theory to interpret random variables.
3. Easier to analyse properties of differential equations.

10.7 Projections

A linear mapping π is a projection if $\pi^2 = \pi \circ \pi = \pi$.

10.7.1 Projection onto a Line

An arbitrary point $\mathbf{x}$ onto a line with basis vector $\mathbf{b}$ is projected onto the point $\mathbf{p}$:

1. Find a λ such that $\mathbf{p} = \lambda\mathbf{b}$ and $\mathbf{x} - \mathbf{p} \perp \mathbf{b}$. Hence $(\mathbf{x} - \lambda\mathbf{b}) \cdot \mathbf{b} = 0 \iff \mathbf{x} \cdot \mathbf{b} - \lambda\mathbf{b} \cdot \mathbf{b} = 0 \iff \boxed{\lambda = \dfrac{\mathbf{x} \cdot \mathbf{b}}{\mathbf{b} \cdot \mathbf{b}}}$.
2. Find the projection point, $\boxed{\mathbf{p} = \lambda\mathbf{b}} = \mathbf{b}\frac{\mathbf{x}\cdot\mathbf{b}}{\mathbf{b}\cdot\mathbf{b}} = \mathbf{b}\frac{\mathbf{b}^\top\mathbf{x}}{\mathbf{b}^\top\mathbf{b}}$.
3. Conclude that the projection matrix $\boxed{\mathbf{P}_\pi = \dfrac{\mathbf{b}\mathbf{b}^\top}{\mathbf{b}^\top\mathbf{b}}}$.

10.7.2 Projection onto a Subspace

An arbitrary point $\mathbf{x}$ onto a subspace with ordered basis $\mathbf{B} = (\mathbf{b}_1 | \dots | \mathbf{b}_n)$ is projected onto the point $\mathbf{p}$:

1. Find a λ such that $\mathbf{p} = \sum_{i=1}^{n} \lambda_i \mathbf{b}_i = \mathbf{B}\lambda$ and $\mathbf{x} - \mathbf{p} \perp \mathbf{b}_i$. Hence $(\mathbf{x} - \mathbf{B}\lambda) \cdot \mathbf{b}_i = 0 \iff \mathbf{B}^\top (\mathbf{x} - \mathbf{B}\lambda) = \mathbf{0} \iff \mathbf{B}^\top\mathbf{B}\lambda = \mathbf{B}^\top\mathbf{x}$. Hence λ can be found by $\boxed{\lambda = \left(\mathbf{B}^\top\mathbf{B}\right)^{-1} \mathbf{B}^\top\mathbf{x}}$.
2. Find the projection point, $\boxed{\mathbf{p} = \mathbf{B}\lambda} = \mathbf{B}\left(\mathbf{B}^\top\mathbf{B}\right)^{-1} \mathbf{B}^\top\mathbf{x}$.
3. Conclude that the projection matrix $\boxed{\mathbf{P}_\pi = \mathbf{B}\left(\mathbf{B}^\top\mathbf{B}\right)^{-1} \mathbf{B}^\top}$.
4. To check your answer, ensure that:
   (a) $\mathbf{p} - \mathbf{x} \perp \mathbf{b}$ for each $\mathbf{b}$ in $\mathbf{B}$.
   (b) $\mathbf{P}_\pi = \mathbf{P}_\pi^2$.

10.7.3 Projection Error

$d(\mathbf{x}, U)$ for a point $\mathbf{x}$ projected onto the basis U.

1. Calculate the value $\|\mathbf{x} - \mathbf{p}\|$.

10.7.4 Applications

1. In graphics, e.g. to generate shadows.
2. Optimisation: orthogonal projections used to iteratively minimise residual errors.
3. Project high dimensional data into a lower dimensional feature space, e.g. for ML.

10.8 Rotations

10.8.1 In Two Dimensions

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

(for a counter-clockwise rotation by θ)

10.8.2 In Three Dimensions

We keep one axis constant and view this axis "from the end towards the origin". A counter-clockwise rotation is then given by:

$$\mathbf{R}_1(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

$$
\mathbf{R}_2(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 1 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}
$$

$$
\mathbf{R}_3(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

10.8.3 In n Dimensions

We keep all but two axes the same:

$$
\mathbf{R}_{ij}(\theta) = \begin{bmatrix} \mathbf{I}_{i-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \cos\theta & \mathbf{0} & -\sin\theta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{j-i} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sin\theta & \mathbf{0} & \cos\theta & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_{n-j} \end{bmatrix}
$$

10.8.4 Properties

1. Composition of rotations is $\mathbf{R}(\phi)\mathbf{R}(\theta) = \mathbf{R}(\phi+\theta)$
2. Preserves lengths and distances, i.e. $\|\mathbf{x}\| = \|\mathbf{R}(\theta)\mathbf{x}\|$ and $\|\mathbf{x}-\mathbf{y}\| = \|\mathbf{R}(\theta)\mathbf{x} - \mathbf{R}(\theta)\mathbf{y}\|$
3. Not commutative (except in two dimensions)

10.9 Cayley-Hamilton Theorem

Useful for "Find an expression for $\mathbf{A}^m$ in terms of $\mathbf{I}, \mathbf{A}, \mathbf{A}^2, \ldots$":

For an endomorphism with transformation matrix $\mathbf{A}_\Phi$ with characteristic polynomial p:

$$
p(\mathbf{A}_\Phi) = \mathbf{0}
$$

10.10 Affine Mappings

10.10.1 Definition

An affine mapping is defined as $x \to a + \Phi(x)$ where Φ is a linear mapping.

Points to Note

1. The composition of affine mappings is an affine mapping (same as for linear mappings).
2. Affine mappings preserve distances and parallelism.

11 Scalar Products

11.1 Proving a Mapping is a Scalar Product

For a mapping $\langle \mathbf{x}, \mathbf{y} \rangle : V \times V \to \mathbb{R}$:

1. Prove the mapping is linear in both arguments.
2. Prove the mapping is symmetric: $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ for all $\mathbf{x}, \mathbf{y} \in V$.
3. Prove the mapping is positive definite: $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ except for $\langle \mathbf{0}, \mathbf{0} \rangle = 0$.

The standard scalar product is $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\top \mathbf{y}$.

11.1.1 Applications

1. Compute angles between vectors or distances, determine whether orthogonal.
2. Allows us to determine specific bases where each vector is orthogonal to others - important for optimistations of numerical algortihms for solving LEQs.
3. Kernel methods in machine learning. Allows for non-linearisation of many ML algorithms.

11.2 Useful Geometric Properties

11.2.1 Finding the Length of a Vector

$$
\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}
$$

A *norm* has the properties:

1. $\|\mathbf{x}\| > 0$ except for $\|\mathbf{0}\| = 0$
2. $\|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

11.2.2 Finding the Distance between two Vectors

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

A *metric* has the properties:

1. Symmetric
2. Positive Definite
3. Obeys triangle inequality: $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$

11.2.3 Showing Orthogonality

$$\mathbf{x} \perp \mathbf{y} \iff \langle \mathbf{x}, \mathbf{y} \rangle = 0$$

11.2.4 Finding the Angle between Vectors

$$\cos \omega = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

11.2.5 Properties of Euclidean Vector Spaces

1. Cauchy-Schwarz Inequality: $\langle \mathbf{x}, \mathbf{y} \rangle \leq \|\mathbf{x}\| \|\mathbf{y}\|$
2. Minkowski Inequality: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
3. Triangle Inequality: $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$
4. Parallelogram Law: $\|\mathbf{x} + \mathbf{y}\| + \|\mathbf{x} - \mathbf{y}\| = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$
5. $4\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2$
6. $\mathbf{x} \perp \mathbf{y} \iff \|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$