Linear Algebra Notes

Dan Hathaway

January 2010

1 Goal

The purpose of this document is to organize facts about vector spaces and vector spaces with an inner product operation. The Axiom of Choice will be used without mention. Facts that apply only to finite dimensional vector spaces will be prefaced with (Fin) while facts that apply to infinite dimensional spaces will be prefaced with (Inf).

2 Vector Spaces and Linear Transformations

2.1 Characterization of Vector Spaces

(Fin+Inf) Using the Axiom of Choice, every vector space has a basis. Using the boolean prime ideal theorem (which is strictly weaker than AC but still independent of ZF) it follows that all bases of a vector space have the same cardinality.

(Fin) All vector spaces of dimension n with the same scalar field are isomorphic.

2.2 Linear Transformations

Let V be a vector space over a field F. A function $A : V \to V$ is a **linear transformation** iff

(i) $\forall_{\mathbf{x},\mathbf{y} \in V} A(\mathbf{x}+\mathbf{y}) = A(\mathbf{x}) + A(\mathbf{y})$

(ii) $\forall_{\mathbf{x} \in V, \lambda \in F} A(\lambda \mathbf{x}) = \lambda A(\mathbf{x})$.

Let the space of all linear transformations from V to V be denoted $L(V,V)$. We could have defined linear transformations to map from one vector space to another, $L(U,V)$, but we will be concerned here with transformations of a space to itself. $L(V,V)$ is not only a vector space, but it poseses an operation of *multiplication*, which is transformation composition.

Every $L(V,V)$ posesses an identity (linear) transformation I. The transformation I is such that $\forall_{T \in L(V,V)} TI = IT = T$.

(Fin) Given a linear transformation A, if there exists a transformation B s.t. $AB = I$, then B is the unique transformation with this property. Furthermore, $BA = I$. We denote B by A^{-1}. The same applies considering A the unique transformation for which $AB = I$ so $A = B^{-1}$.

(Inf) While $L(V,V)$ still has an identity transformation I, it is possible for there to exist $A, B \in L(V,V)$ s.t. $AB = I$ but there is no $C \in L(V,V)$ s.t. $CA = I$. For example, consiter a countable infinite dimensional vector space with basis $B = \{b_i\}_{i=1}^{\infty}$ and let A be the transformation that changes the i-th coordinate of a vector to the $(i+1)$-th coordinate. While there exists a B with the desired property, there is no C. The same argument applies for right multiplication instead of left.

A transformation $T \in L(V,V)$ is determined completely by its behavior on a basis of V.

(Fin) $T \in L(V,V)$ is injective iff it is surjective. $T \in L(V,V)$ is not injective iff it is **singular**.

(Inf) It is possible for a transformation from an infinite dimensional vector space to itself to be surjective but not injective as well as injective but not surjective.

2.3 (Fin) Matrices

A **matrix** over a field F, called a matrix over F for short, is an n-by-m array of elements in F. Matrices can be:

(i) added to each other,

(ii) multiplied by each other (if they have the appropriate dimensions), and

(iii) multiplied by elements in F

in the usual way. Let $M(n, m, F)$ denote the space of n my m matrices over F. $M(n, m, F)$ is a vector space over F.

Let V be a vector space over F. The space $M(n, n, F)$ with the operations of matrix addition, matrix multiplication, and scalar multiplication is isomorphic to the space $L(V, V)$ with the operations of transformation addition, transformation composition, and scalar multiplication.

2.4 Similarity

Two linear transformations $A, B \in L(V, V)$ are **similar** if there exists a $C \in L(V, V)$ such that $A = C^{-1}BC$. Similarity is an equivalence relation.

(Fin) The usage of the term "similarity" can be applied to matrices without confusion. In the group of non-singular transformations in L(V,V), the similary equivalence classes are precisely the conjugacy classes of the group.

(Fin (+Inf?)) [reword for transformations] Two matrices are similar iff they represent the same linear transformation (in different bases). A property of a linear transform is independent of its matrix representation in any basis. Thus, any property of a linear transformation (a property involving only the basic machinery of vector spaces) is a property that is invariant with respect to similar matrices. When considering inner product spaces later on, "properties" that involve the inner product will be invariant with respect to unitary similarity (verify this!).

(Fin (+Inf?)) If A and B are matrices over F and E is an extension field of F, then A and B are similar when considered as matrices over F iff they are similar when considered as matrices over E. That is, if A and B are similar matrices over any field E, then there is a matrix C such that $A = C^{-1}BC$ whose elements are in the smallest subfield of E that contains the elements of both A and B.

3 Vector Spaces with Inner Products

3.1 Remark

When we talk about an inner product space, we are taking the scalar field to be either the real or the complex numbers. Although we could generalize to fields that satisfy certain special axioms, this is generally not done in practice. We differentiate between real inner product spaces and complex inner product spaces.

(Fin (+Inf?)) Any finite dimensional vector space can be turned into an inner product space by choosing a basis and pretending these vectors are orthogonal and definite the inner product as the sum of the product of coordinates with respect to this basis.

3.2 Real Inner Product Spaces

A **real inner product space** (**Euclidean Space**) is a vector space V over the reals R together with a function $\langle x, y\rangle : V \times V \to R$ that satisfies (for all $x, y \in V$ and $\lambda \in R$):

1. $\langle x, y\rangle = \langle y, x\rangle$
2. $\langle \lambda x, y\rangle = \lambda\langle x, y\rangle$
3. $\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_1, y\rangle$
4. $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0$ iff $x = 0$.

In a real inner product space, we can define angle in the usual way (we do not define angle in a complex inner product space). In both a real and complex inner product space, we have the notion of **orthogonal**. All finite dimensional inner product spaces (both real and complex) have an orthonormal basis.

All real inner product spaces of dimension n are isomorphic.

(Complex also?) Given the linearly independent vectors $e_1, ..., e_n$, the coeficients $c_1, ..., c_n$ such that $c_1e_1 + ... + c_ne_n$ is the closest point in span($\{e_1, ..., e_n\}$) to f are given by the **normal equations** (i = 1, ..., n):

$$\sum_j^n \langle e_j, e_i \rangle c_j = \langle f, e_i \rangle.$$

Viewing the vectors $e_1, ..., e_n$ (written in the coordinates of a basis) as the columns of the matrix A, the normal equations become $A^T Ac = A^T$ and the vector c is the vector that minimizes $\|Ac - f\|^2$.

3.3 Complex Inner Product Spaces

(We all know the definition).

3.4 Inner Product Spaces (Real + Complex)

Given an orthonormal basis b_1, ..., $b_n \in V$, we can consider a vector x as being represented by the (unique) **coordinates** x_1, ..., $x_n \in F$ where $x = x_1b_1 + ... + x_nb_n$. We then have $\langle x, y\rangle = x_1y_1 + ... + x_ny_n$.

Both real and complex inner product spaces satisfy the cauchy-schwartz inequality:

$$\forall_{x,y \in V} |\langle x, y\rangle|^2 \leq \langle x, x\rangle\langle y, y\rangle$$

3.5 Norms

We will be concerned generally with inner products instead of norms, but norms exemplify a particularly important feature of inner products.

A vector space V over S ($= R$ or C) together with a function $\|x\| : V \to R$ that satisfies the following (for all $x, y \in V$ and $\lambda \in S$) is called a **normed linear space**:

1. $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$,
2. $\|\lambda x\| = |\lambda|\|x\|$,
3. $\|x + y\| \leq \|x\| + \|y\|$.

In any (real or complex) inner product space, the function $f(x) = \langle x, x\rangle^{1/2}$ is a norm. In a normed vector space, there exists an inner product operation that yeilds the norm iff the norm satisfies the parallelogram law:

$$2\|x\|^2 + 2\|y\|^2 = \|x + y\|^2 + \|x - y\|$$

in which case the inner product is uniquely specified as:

$$\langle x, y\rangle = \frac{1}{2}(\|x + y\| - \|x\| - \|y\|).$$

3.6 Adjoints (Complex Case)

(Fin) There is a one-to-one correspondence between bilinear forms $B(x, y)$ and linear transformations A s.t. $\forall_{x,y}B(x, y) = \langle Ax, y\rangle$. (Explain how to compute one matrix from the other!). There is an analogous one-to-one correspond between bilinear forms $B(x, y)$ and linear transformations A^* s.t. $\forall_{x,y}B(x, y) = \langle x, A^*y\rangle$. The induced one-to-one correspondence that maps each A to A^* is called the **adjoint operation** and A^* is the **adjoint** of A. The adjoint operation satisfies the following:

1. $A^{**} = A$

2. If A is invertible, then A is invertible and $(A^*)^{-1} = (A^{-1})^*$

3. $(A+B)^* = A^* + B^*$

4. $(\lambda A)^* = \bar{\lambda} A^*$

5. $(AB)^* = B^* A^*$.

Thus, given the linear transformation A and the operations of taking the inverse and adjoint, the only transformations that can be obtained are A, A^*, A^{-1}, and $(A^*)^{-1}$.

3.7 Special Types of Linear Transformations on Complex Inner Product Spaces

A transformation T is **normal** iff $TT^* = T^*T$. A transformation T is **unitary** iff $TT^* = T^*T = I$ $(T^{-1} = T^*)$. A transformation T is **self-adjoint**, also called **Hermitian**, iff $T = T^*$. A transformation T is **idempotent** iff $T^2 = I$ $(T^{-1} = T)$.

All of these classes are closed with respect to unitary similarity.

On a finite dimensional vector space, $TT^* = I \leftrightarrow T^*T = I$ but this is not the case on infinite dimensional vector spaces.

The self-adjoint (unitary, normal) matrices are precicely those matrices that correspond to the self-adjoint (unitary, normal) transformations relative to any (as well as every) orthonormal basis.

(Fin) A matrix is normal iff it is unitarily similar to a diagonal matrix.

(Fin) A matrix is self-adjoint iff it is unitarily similar to a diagonal matrix with real entries.

(Fin) A matrix is unitary iff it is unitarily similar to a diagonal matrix with entries of absolute value one.

(Fin+(Inf?)) A matrix is unitary iff it corresponds to a transformation from one orthonormal basis to another.

(Fin) A transformation is normal iff there is an orthonormal (or just orthogonal) basis relative to which the matrix of the transformation is diagonal.

(Fin) A transformation is self-adjoint iff there is an orthonormal (or just orthogonal) basis relative to which the matrix of the transformation is diagonal with real entries.

(Fin) A transformation is unitary iff there is an orthonormal basis relative to which the matrix of the transformation is diagonal with entires of absolute value one.

(Inf) Although for finite dimensional vector spaces a matrix U is unitary iff $\|x\| = \|Ux\|$ (U is an isometry), in an infinite dimensional vector space an isometry need not be unitary.

3.8 Interesting Conditions Equivalent to Being Normal(Fin)

The following conditions are equivalent (wikipedia)

1. A is normal.

2. A is diagonalizable by a unitary matrix.

3. The entire space is spanned by some orthonormal set of eigenvectors of A.

4. $\forall_x \|Ax\| = \|A^*x\|$.

5. $tr(A^*A) = \sum_i^n |\lambda_i|^2$.

6. The Hermitian part $1/2(A + A^*)$ and skew-Hermitian part $1/2(A - A^*)$ commute.

7. A^* is a polynomial (of degree $\leq n-1$) of A.

8. $A^* = AU$ for some unitary matrix U.

9. U and P commute, where we have the polar decomposition $A = UP$.

10. A commutes with some normal matrix N with distinct eigenvalues.

3.9 Analagous Special Types of Linear Transformations on Real Inner Product Spaces

May want to talk about the representation of all "Orthogonal" transformations from R^n to R^n.

3.10 Closure Properties of Special Transformations

Given any linear transformation A, both AA^* and A^*A are self-adjoint. The product of two self-adjoint transformations is self-adjoint iff they commute. If A and B are self-adjoint then so is $AB + BA$ and $i(AB - BA)$.

3.11 Existence of Eigenvectors

If the chararistic polynomial of A has a root in F, then A has at least one eigenvector. Indeed, there is an eigenvector for every root (in the scalar field) of the characteristic polynomial. Thus, a linear transformation over a real vector space of odd dimension has at least one eigenvector. To show that any self-adjoint matrix has an orthogonal basis of eigenvectors we iterate the argument that every invariant subspace has at least one eigenvector and the fact that the space orthogonal to any set of eigenvectors is invariant.

3.12 Analogy To Complex Numbers

(Fin(+Inf?)) The adjoint operator is analogous to the operation of taking the complex conjugate. Self-adjoint matrices are analogous to real numbers (they can be diagonolized by a unitary matrix to have real entires). Positive definite matrices are analogous to positive real numbers and positive semi-definite matrices are analogous to non-negative real numbers. Unitary matrices are analogous to complex numbers of absolute value one (they can be diagonolized by a unitary matrix to have complex entries of absolute value one). There is no analogue to normal matrices for the complex numbers because complex numbers commute.

(Fin+Inf) Every linear transformation A can be written as a sum $A_1 + iA_2$ where A_1 and A_2 are self-adjoint.

3.13 Polar Decomposition

(Fin) Continuing the analogy between linear transformations complex numbers: Every non-singular linear transformation A can be uniquely written as UP where U is unitary and P is positive definite (self-adjoint were all eigenvalues are positive) (analogous to the unique representation of a non-zero complex number in polar form). Ever linear transformation A can be written (not neccesarily uniquely) as UP where U is unitary and P is positive semi-definite (self-adjoint were all eigenvales are non-negative).

(Fin) A linear transformation $A = UH$ where U is unitary and H is self-adjoint is normal iff U and H commute.

3.14 Diagonalization

(Fin) The following conditions are equivalent for a matrix A:

(i) A is diagonalizable.

(ii) There exists a matrix C such that $C^{-1}AC$ is diagonal.

(iii) There exists a basis B such that the matrix representation of the transformation of A with respect to B is diagonal.

(iv) These set of eigenvectors of A spans the vectorspace.

(Fin) The following conditions are equivalent for a matrix A:

(i) A is normal.

(ii) These exists a unitary matrix U such that $U^{-1}AC$ is diagonal.

(iii) These exists an orthonormal basis B such that the matrix representation of the transformation of A with respect to B is diagonal.

(iv) These exists an orthonormal set of eigenvectors of A.

3.15 Simultaneous Diagonalization

(Fin) A set of matrices is simultaneously diagonalizable iff it consists of pairwise commuting diagonalizable matrices.

(Fin) A set of matrices is simultaneously diagonalizable by a unitary matrix iff it consists of pairwise commuting normal matrices.

3.16 Note About Similarity

Question: If E is an extension field of F and if matrices A and B over a field F are similar to a matrix over the field E, then are A and B similar to each other over F? The answer is yes. If A and B are not similar over E, then they are of course not similar over F.

4 Bilinear and Quadratic Forms

4.1 Real Bilinear Forms

Let V be a vector space over the reals. A **real bilinear form** is a function $A(x, y) : V \times V \to R$ that is linear in each argument when the other argument is kept fixed. A **symmetric** bilear form A is one that satisfies $\forall_{x,y \in V} A(x, y) = A(y, x)$. A **real quadratic form** is a function $B(x) : V \to R$ such that $\forall_x B(x) = A(x, x)$ for some real bilinear form A. Given any quadratic form B, there is a unique *symmetric* bilinear form that yields the quadratic form in this way, called the bilinear form **polar** to the quadratic, and it can be computed as:

$$A(x, y) = \frac{1}{2}[B(x + y) - B(x) - B(y)]$$

analogous to the formula for computing the inner product that yeilds a norm. A quadratic form B is **positive definite** iff $\forall_{x \neq 0} B(x) > 0$ ($B(0) = 0$ is always implied by linearity). A real inner product is a symmetric bilinear form corresponding to a positive definite quadratic form. Conversly, such a bilinear form always defines an inner product.

Given a basis for V and a bilinar form A, the matrix M such that $A(x, y) = y^T Mx$ is called the matrix of the bilinear form and $y^T Mx = \langle Mx, y\rangle$ where the inner product is the induced inner product from the basis. The same holds for complex bilinear forms in a complex inner product space. Given a basis, a bilinear form A is symmetric iff its matrix with respect to this basis is symmetric in the usual sense.

Given any real quadratic form, there exists a basis (and even an orthonormal basis if we are in an inner product space I believe) in which the quadratic form can be written as the sum of squares of coordinates of vectors in the space: $\lambda_1 c_1^2 + ... + \lambda_n c_n^2$. Let A be the matrix of the quadratic form $Q(x) = x^T Ax$ with respect to some basis. By the **method of Jacobi** there exist a basis relative to which the values of λ_i are $\lambda_1 = \frac{\Delta_0}{\Delta_1}, ..., \lambda_n = \frac{\Delta_{n-1}}{\Delta_n}$ where each Δ_i is the determinant of the upper left minor of A of size i. The **Law of Inertia** states that given any two bases such that a quadratic form is expressible as a sum of squares of components, the number of positive, negative, and zero values of λ_i will be the same.

4.2 Complex Bilinear Forms

A **complex bilinear form** is a function $A(x, y) : V \times V \to R$ that is linear in the first argument when the second is kept fixed and the second argument satisfies $A(x, y_1 + y_2) = A(x, y_1) + A(x, y_2)$ and $A(x, \lambda y) = \bar{\lambda} A(x, y)$. A **complex quadratic form** B is any function such that $\forall_x B(x) = A(x, x)$ where A is some complex bilinear form. A complex bilinear form is **Hermitian** if $A(x, y) = \overline{A(y, x)}$. Given a basis, a blinear form is Hermitian iff its matrix with repsect to this basis is Hermitian in the usual sense.

A complex bilinear form A is Hermitian iff $A(x, x)$ is real for all $x \in C$. Hence, a complex quadratic form is Hermitian iff it is real valued.

Given a complex quadratic form B, there is **unique** complex bilinear form A such that $B(x) = A(x,x)$ and it can be computed by:

$$A(x,y) = \frac{1}{2}(B(x+y) + iB(x+iy) - B(x-y) - iB(x-iy).$$

This establishes a 1-1 correspondence between complex bilinear forms and complex quadratic forms, which contrasts with the real case where only *symmetric* real bilinear forms are associated with unique real quadratic forms.

A complex inner product is a Hermitian complex bilienar form corresponding to a positive definite quadratic form. Conversly, every complex bilinear form of this type corresponds to a complex inner product. This is analagous to the real case, with but with "Symmetric" replaced with "Hermitian".

If a complex quadratic form A is Hermitian (contrasting this with the real quadratic form case), then there is a basis of V relative to which the form is given by: $A(x) = \sum_i \lambda_i c_i \bar{c_i}$ where the c_i are the coordinates of x relative to the basis. The *method of Jacobi* applies to Hermetian quadratic forms. The *law of inertia* also holds for (Hermitian) complex quadratic forms: the number of positive, negative, and zero λ_i values is the same in different bases.

4.3 Positive Definiteness

(Real (+Complex?))A quadratic form is positive definite iff its coeficients of a representation as a sum of squares are all positive. A quadratic form is positive semi-definite iff these coeficients are all non-negative etc.

A matrix is positive-semidefinite if and only if it arises as a Gram matrix of some set of vectors (woah!).

A matrix A is positive semi-definite iff there exist a matrix B s.t. $B^2 = A$.

There is much more to say about Positive Definiteness.