UChicago REU 2013 Apprentice Program **Spring 2013**

Lecture 16: July 23nd, 2013

Prof. Babai Scribe: David Kim

Exercise 0.1 *$rk(A^TA) = rk(A)$.*

Proof: Let $A \in \mathbb{R}^{m\times n}$. By the Rank-Nullity Theorem,

$$rk(A) + null(A) = n$$
$$rk(A^TA) + null(A^TA) = n$$

where $null(A) := \dim(\ker A)$, and $\ker A := \{x \in \mathbb{R}^n : Ax = 0\}$.

So $rk(A^TA) = rk(A) \Leftrightarrow null(A^TA) = null(A)$, i.e. $\dim(\ker A^TA) = \dim(\ker A)$.

Claim 0.2 $\ker A^TA \supseteq \ker A$
$\rightarrow$ If $x \in \ker A$ such that $Ax = 0 \Rightarrow A^TAx = 0$, so $x \in \ker A^TA$

Claim 0.3 $\ker A^TA \subseteq \ker A$
$\rightarrow$ This is equivalent to proving if $A^TAx = 0$ then $Ax = 0$. We use: $v = 0 \Leftrightarrow ||v|| = 0$. Since $Ax = 0 \Leftrightarrow ||Ax||^2 = x^TA^TAx = 0$, if $A^TAx = 0$, then $Ax = 0$.

By the previous two claims, $\ker A^TA = \ker A$, and hence $\dim(\ker A^TA) = \dim(\ker A) \Rightarrow rk(A) = rk(A^TA)$. ∎

1 Axioms of "Euclidean Space"

Let V be a vector space over $\mathbb{R}$. An "inner-product" is a function: $V \times V \to \mathbb{R}$, that is $(x, y) \mapsto \langle x, y\rangle \in \mathbb{R}$ for $x, y \in V$, which satisfies the following axioms:

1. $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
2. $\langle \alpha x, y\rangle = \alpha \langle x, y\rangle$
3. $\langle x, y + z\rangle = \langle x, z\rangle + \langle x, y\rangle$
4. $\langle x, \alpha y\rangle = \alpha \langle x, y\rangle$
5. $\langle x, y\rangle = \langle y, x\rangle$ (symmetric)
6. $\langle x, x\rangle > 0$ for $x \neq 0$ (positive definite)

Axioms (1) and (2) state that $\langle x, y\rangle$ is *linear in* x, and (3) and (4) state that $\langle x, y\rangle$ is *linear in* y. Thus $\langle x, y\rangle$ distributes over both the first and second components, and hence is *bilinear*. Note that the symmetric condition (5) makes conditions (3) and (4) superfluous, given (1) and (2) (or vice versa).

Definition 1.1 (Euclidean Space) $V, \langle\cdot,\cdot\rangle$ - *A vector space over* $\mathbb{R}$ *together with a positive definite, symmetric, bilinear inner product.*

Example 1.2 (Examples of Euclidean Spaces)

1. $\mathbb{R}^n$ *with* $\langle x, y\rangle := x^T y$ *(Standard inner product on* $\mathbb{R}^n$*)*
2. $\mathbb{R}^n$ *with* $\langle x, y\rangle := x^T Ay$*, where* $A \in \mathbb{R}^{n\times n}$ *is symmetric and positive definite.*
   - $\forall A \in M_n(\mathbb{R})$*, Axioms (1) - (4) are automatically satisfied (always bilinear).*
   - A *must be symmetric and positive definite.*

Exercise 1.3 *Prove:* $\langle x, y\rangle := x^T Ay$ *is a symmetric function* $(x^T Ay = y^T Ax) \Leftrightarrow A$ *is symmetric,* $(A = A^T)$*.*

Exercise 1.4 *Prove:* $\langle x, y\rangle := x^T Ay$ *is positive definite* $(Axiom\ (6)) \Leftrightarrow A$ *is positive definite.*

Note that given a vector space V, we have many choices of the inner-product to create Euclidean spaces.

Exercise 1.5 *Let* $C[0,1]$ *be the space of continuous functions on* $[0,1]$*. Let* $f, g \in C[0,1]$*. Can you suggest* $\langle f, g\rangle$*?*

Answer: $f(x)g(x)$ *satisfies (1) - (5), but this is not a real value.* $\langle f, g\rangle := \int_0^1 f(x)g(x)dx$ *satisfies all the axioms.*

In fact, we may define $\langle f, g\rangle := \int_0^1 f(x)g(x)\rho(x)dx$, where $\rho \in C[0,1]$ is a "weight function".

Question: under what conditions on $\rho \in C[0,1]$ would the above definition satisfy the inner-product axioms? For any $\rho \in C[0,1]$, axioms (1) - (5) are satisfied automatically.

Exercise 1.6 *if* $\forall x \in [0,1], \rho(x) > 0$*, then (6) holds.*

How can we make this condition weaker and still have it satisfy (6)? If ρ can be 0 at "a few points", and is "mostly" positive.

Exercise 1.7 *Clarify:* $\rho \in C[0,1]$ *is 0 at "a few" points.*

Definition 1.8 (Space of Polynomials) $\mathbb{R}[x]$ *is the set of polynomials, with coefficients in* $\mathbb{R}$*, in a single variable* x*.*

Note that $\mathbb{R}[x]$ has a countable basis - $\{1, x, x^2, \dots\}$.

Definition 1.9 (Space of Rational Functions) *Consider the set of fractions* $\frac{p(x)}{q(x)}$*, where* $p, q \in \mathbb{R}[x], q \neq 0$*, the constant 0 polynomial. As with fractions of integers, two fractions* p_1/q_1 *and* p_2/q_2 *of polynomials are considered "equal" if* $p_1 q_2 = p_2 q_1$*. This is an equivalence relation, and the equivalence classes are called "rational functions." They form the vector space* $\mathbb{R}(x)$*.*

Note that the term "rational function" is a misnomer: these are equivalence classes of formal expressions, not function: we cannpt plug numbers into them (they may turn the denominator into zero).

Definition 1.10 *An infinite set of vectors are linearly independent if every finite subset of the vectors are linearly independent.*

Exercise 1.11 *Prove:* $\dim(\mathbb{R}(x))$ *is uncountably infinite.*
Hint: find for every $\alpha \in \mathbb{R}$*, a rational function* $f_\alpha(x)$ *such that* $\{f_\alpha : \alpha \in \mathbb{R}\}$ *are linearly independent. The expression defining* f_α *should be very simple.*

Returning to $\mathbb{R}[x]$, some important examples of inner-products are:

$$\langle f, g \rangle = \begin{cases} \int_0^1 f(x)g(x)dx \\ \\ \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2}dx \\ \\ \int_{-1}^{1} f(x)g(x)\sqrt{1-x^2}dx \end{cases}$$

Exercise 1.12 *Prove that* $\int_{-\infty}^{\infty} f(x)g(x)e^{-x^2}dx < \infty$*, well-defined, when* f, g *are polynomials.*

For the following definitions, assume that we have $V, \langle \cdot, \cdot \rangle$ a Euclidean space.

Definition 1.13 (Norm) $\forall v \in V$*,* $||v|| = \sqrt{\langle v, v \rangle}$*.*

Note that $||v|| = 0 \Leftrightarrow v = 0$ follows from the condition of positive definiteness of an inner-product.

Definition 1.14 (Orthogonal Vectors) $v \perp w$ *if* $\langle v, w \rangle = 0$*.* v *and* w *are orthogonal, or perpendicular.*

Look up separate chapters of analysis on "Orthogonal functions," with the most famous example coming from Fourier Analysis:

Exercise 1.15 *Prove:* $\{1, \cos x, \sin x, \cos 2x, \sin 2x, \dots\}$ *are orthogonal w.r.t* $\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx$*.*

Thus the Fourier coefficients of a periodic function, when expanded in the above basis, may be obtained by simply taking inner-products with the basis elements.

Theorem 1.16 (Cauchy-Schwarz) *Given* $x, y \in V$, *a Euclidan space,* $|\langle x, y\rangle| \leq ||x|| \cdot ||y||$.

In fact, Cauchy-Schwarz was discovered in the context of function spaces in the following form:

$$\left|\int fg\rho dx\right| \leq \sqrt{\int f^2\rho dx}\sqrt{\int g^2\rho dx}$$

We will see that essentially, for each finite dimension there is only one Euclidean space, as any two of the same dimension are isometric. Thus, given any two functions f, g in an Euclidean space, they could simply be seen as vectors spanning a 2-dimensional plane, and everything we know about vectors in the plane applies exactly to the corresponding functions - the angle between two functions, triangle inequality (giving us Cauchy-Schwarz), law of sines and cosines etc.

Another remarkable subject worth looking into: "Orthogonal polynomials." These result from orthogonalization applied to the basis $1, x, x^2, \ldots$.

2 Gram-Schmidt Orthogonalization

We will give an online algorithm which takes each $\{v_1, v_2, v_3, \ldots\}$ and produces $\{b_1, b_2, b_3, \ldots\}$. Let $U_i = span(v_1, \ldots, v_i)$. (Note: $U_0 = span\{\emptyset\} = \{0\}$ - the empty sum, or the sum of nothing is 0.) The algorithm works under the following axioms:

1. $(\forall i \neq j)(b_i \perp b_j)$
2. $(\forall i \geq 1)(b_i - v_i \in U_{i-1})$

Observation $U_0 \subseteq U_1 \subseteq U_2 \subseteq \ldots$, by definition.

Claim 2.1 $(\forall i)(b_i \in U_i)$

(Proof) For $i = 1$, $b_1 - v_1 \in U_0 = \{0\}$ gives us $b_1 = v_1 \in U_1 = span(v_1)$.
For $i > 1$, $b_i - v_i \in U_{i-1} = span(v_1, \ldots, v_{i-1}) \Rightarrow b_i = v_i + \sum_{j=1}^{i-1} \beta_j v_j \in span(v_1, \ldots, v_i) = U_i$. $\square$

Claim 2.2 $(\forall i)(span(b_1, \ldots, b_i) = U_i)$ *- the first i input vectors and output vectors span the same space.*

(Proof) For $i = 1, b_1 = v_1$ and $span(b_1) = span(v_1) = U_1$. Assume the claim is true for i. Then axiom (2) gives us $b_{i+1} - v_{i+1} \in U_i$, and $U_i = span(b_1, \ldots, b_i)$ by the induction hypothesis. So $v_{i+1} = b_{i+1} + \sum_{j=1}^{i} \beta_j b_j$, and $U_{i+1} \subseteq span(b_1, \ldots, b_{i+1})$. The previous Claim 2.1 and the observation gives us $span(b_1, \ldots, b_{i+1}) \subseteq U_{i+1}$. $\square$

Claim 2.3 *(1) and (2) uniquely determine the output.*

(Proof) For each i, we have $b_i - v_i \in U_{i-1} = span(b_1, \ldots, b_{i-1})$. So $b_i - v_i = \mu_{i_1} b_1 + \cdots + \mu_{i_{i-1}} b_{i-1}$. By axiom (1), $\langle b_i - v_i, b_j\rangle = \langle b_i, b_j\rangle - \langle v_i, b_j\rangle = \mu_{i_j}||b_j||^2 \Rightarrow \mu_{i_j} = -\frac{\langle v_i, b_j\rangle}{||b_j||^2}$ for $j = 1, \ldots, i-1$. So the μ_{i_j} and b_i are uniquely determined by $b_1, \ldots, b_{i-1}$, and we have always have $b_1 = v_1$. (In the case that $b_j = 0$, then we do not need to evaluate μ_{i_j} as it just a 0 term regardless.)

Exercise 2.4 *If $w_1, \ldots, w_k$ satisfy $(\forall i \neq j)(w_i \perp w_j)$ and $(\forall i)(w_i \neq 0) \Rightarrow$ they are linearly independent.*

Proof: assume $\sum_i \alpha_i w_i = 0$. Then $\forall j, \langle w_j, \sum \alpha_i w_i \rangle = \langle w_j, 0 \rangle = \alpha_j \langle w_j, w_j \rangle = 0$. Since $||w_j||^2 \neq 0 \Rightarrow \alpha_j = 0$. $\Box$

Question: When do we have $b_j = 0$?
Answer: $b_j = 0 \Leftrightarrow U_j = U_{j-1} \Leftrightarrow v_j \in U_{j-1}$.
If the input vectors $v_1, v_2, \ldots$ are linearly independent $\Rightarrow$ all $b_1, b_2, \ldots$ would be nonzero.

Geometric speaking, at each stage, the algorithm simply takes the component of v_{i+1} perpendicular to U_i, and "cuts it off" at the furthest tip parallel to the subspace U_i. The resulting orthogonal vector is b_{i+1}.

Definition 2.5 (Volume of Parallelopiped Formed by Vectors) *In this view, we can define the volume of the n-dimensional parallelepiped formed by a set of vectors $\{v_1, \ldots, v_k\}$, $vol(v_1, \ldots, v_k)$, under the following assumptions:*

(1) If $\forall i \neq j, v_i \perp v_j$, then $vol = \prod_i ||v_i||$
(2) the volume is invariant to cut and paste

It is easy to see that in dimension 2, the output vectors b_1, b_2 form an area that is equal to the area of the parallelogram formed by v_1, v_2 (by cut and paste). The same applies to higher dimensions.

Then with this assumption, the determinant gives us the volume of the parallelepiped (parellelogram) formed by a set of vectors. For example in $\mathbb{R}^2$, given v_1, v_2, we would have area$(v_1, v_2) = ||b_1|| \cdot ||b_2||$ by cut and paste. But area$(v_1, v_2) = ||b_1|| \cdot ||b_2|| = \det \begin{bmatrix} v_1^T \\ v_2^T \end{bmatrix}$, since b_1, b_2 is obtained by elementary row operations by the algorithm, and $||b_1|| \cdot ||b_2||$ gives the value of the determiant, as b_1, b_2 are orthogonal ($|\det B| = \det \sqrt{B^T B}$). We would have the same generalization for larger dimensions.

Exercise 2.6 *If $a_1, \ldots, a_n \in \mathbb{Z}^n$ are linearly independent $\Rightarrow$ volume of parallelepiped spanned by $a_1, \ldots, a_n \in \mathbb{Z}$.*

Answer: the determinant of the matrix formed with the a_i as the rows is $\in \mathbb{Z}$, by the definition of the determinant.

Exercise 2.7 *Let v, w be vectors with integer coefficients in a 3-dimensional space. Will the parallelogram spanned by the two vectors have an integer area?*

Let's look at the above exercise in 1-dimensional space. An integer vector on a line would have an integer length. In 2-dimensional space, however, the length of a vector is not necessarily an integer, but the square root of an integer that is the sum of two squares.

Generalizing further, in 3-dimensional space, the length of a vector with integer coordinates is the square root of the sum of three squares.

Question: Is there an integer that is NOT the sum of three squares? 7 is an example.

Theorem 2.8 (Three Squares Theorem - Gauss) *A positive integer is NOT the sum of three squares* $\Leftrightarrow$ *it is of the form* $8k+7$

Note: one direction of the proof is easy. The other is difficult.

Theorem 2.9 (Four Squares Theorem - Legendre) *Every positive integer is the sum of four squares.*

Note: while this follows from the Three Squares Theorem, it is not too difficult to prove it directly. (Hint: quaternions.)

Exercise 2.10 *The volume of the parallelepiped spanned by* $a_1, \ldots, a_k \in \mathbb{Z}^n$ *is the square root of an integer.*

First show this for 2 vectors in a 3 dimensional space (Exercise 2.7).

Theorem 2.11 *Every finite-dimensional Euclidean space has an orthonormal basis. In fact, every orthonormal set of vectors can be extended to an orthnormal basis. (follows from previous algorithm)*

Now given an orthonormal basis, say $e_1, \ldots, e_n$, and $v = \sum_{i=1}^{n} \alpha_i e_i, w = \sum_{i=1}^{n} \beta_i e_i$, what is $\langle v, w \rangle$?

$$\langle v, w \rangle = \left\langle \sum_i \alpha_i e_i, \sum_i \beta_i e_i \right\rangle = \sum_{i,j} \alpha_i \beta_j \langle e_i, e_j \rangle = \sum_i \alpha_i \beta_i \langle e_i, e_i \rangle = \alpha^T \beta$$

This is simply the same standard inner-product, in the coodinates of the orthonormal basis.

Definition 2.12 (Vector Space Isomorphism) *An isomorphism of the vector spaces* V *and* W *is a bijection* $f : V \to W$ *which preserves linear combinations.*

Definition 2.13 (Vector Space Isometry) *An isomorphism* $f : V \to W$ *between two Euclidean spaces is an* isometry *if it preserves* $\langle \cdot, \cdot \rangle$*, i.e.,* $(\forall x, y \in V)(\langle f(x), f(y) \rangle_W = \langle x, y \rangle_V)$.

Exercise 2.14 *If* $(\forall x \in V)(||x||_V = ||f(x)||_W) \Rightarrow \langle \cdot, \cdot \rangle$ *is preserved. (For an isomorphism it suffices that it preserve the norm to be an isometry.)*

When are two finite dimensional Euclidean spaces isometric? The first condition is that they must have the same dimensions. The following theorem states that this is the only condition.

Theorem 2.15 *If V, W are n-dimensional Euclidean spaces $\Rightarrow V, W$ are isometric.*

Proof: Let $e_1, \ldots, e_n$ be an orthonormal basis of V. Let $f_1, \ldots, f_n$ be an orthonormal basis of W. Define $\varphi : V \to W$ as follows

$$v = \sum_i \alpha_i e_i \mapsto \varphi(v) = \sum_i \alpha_i f_i$$

Then φ is a bijection. We need to show that it preserves inner products:

$$\langle v, w \rangle = \alpha^T \beta = \langle \varphi(v), \varphi(w) \rangle$$

■

So every n-dimensional Euclidean space is isometric to the standard Euclidean space of $\mathbb{R}^n$. This even applies to function spaces: a space of functions spanned by a finite number of vectors is isomoteric to $\mathbb{R}^n$ for some n.

Definition 2.16 (Gram matrix) *Let $v_1, \ldots, v_k \in V$. The Gram matrix of $v_1, \ldots, v_k$ is $G = G(v_1, \ldots, v_k) = (g_{ij})_{k \times k}$, where $g_{ij} = \langle v_i, v_j \rangle$.*

Exercise 2.17 *Prove: G is symmetric, positive-semidefinite.*

Question: when is it diagonal? when the v_i are orthgonal.
When is it positive definite? $\Leftrightarrow \det G \neq 0 \Leftrightarrow v_1, \ldots, v_k$ are linearly independent.

Exercise 2.18 ** $vol(v_1, \ldots, v_k) = \sqrt{\det G}$, where G is the Gram-Matrix of $v_1, \ldots, v_k$.*