MATH 23A
SOLUTION SET 5

NILS R. BARTH
November 5, 1999

Explanation: Rather than a conventional solution set, I begin with an exposition on change of basis, then discuss bilinear forms, their representation by matrices and the effect of change of basis on these.

You may want to use this as a reference on these subjects.

Change of Basis: Given $\{e_1, \dots, e_n\}, \{v_1, \dots, v_n\}$ two bases for V, we can therefore express $v_j = \sum a_{ij} e_i$, and $e_j = \sum b_{ij} v_i$. Now the matrix $A = (a_{ij})$ will take an expression for the vector $v = \sum c_i v_i = (c_1, \dots, c_n)^T$ in terms of the v_i and return $(d_1, \dots, d_n)^T$, where $v = \sum d_i e_i$, that is, it *changes the coordinates* from being in the basis $\{v_i\}$ to being in the basis $\{e_i\}$. Clearly the matrix B does the exact opposite, and thus $B = A^{-1}$, so

5.1. Proposition. *Change of basis matrices are invertible.*

Here's another way to see this: a choice of basis for V is a choice of isomorphism[1] $\Phi\colon \mathbf{R}^n \to V$. Thus, given two bases $\{e_i\}, \{v_i\}$, we have the following diagram:

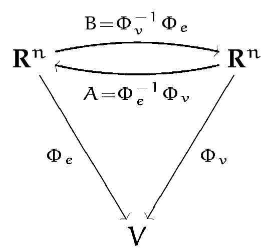

FIGURE 1. Change of basis

from which it is clear that A, B are inverses.

Note that a change of basis is a map $\mathbf{R}^n \to \mathbf{R}^n$, *not* a map $V \to V$. That is, a change of basis matrix is always in the standard basis for $\mathbf{R}^n$, not in the basis $\{e_i\}$ or $\{v_i\}$.

A very common computation is to be given a matrix M for a linear operator[2] with respect to one basis $\{e_i\}$, and want to find its matrix with respect to a different basis $\{v_i\}$. Here's how you do it: given coordinates with respect to the v_i, first convert them into coordinates with respect to the e_i, then feed them through M, then take the output (which is in e_i coordinates) and convert it back to v_i coordinates. Thus, in the above notation, the new matrix is

$$BMA = A^{-1} MA. \tag{5.2}$$

This is how you transform a matrix for a *linear operator*.

Bilinear Forms as Matrices: Note that if you have a basis $\{v_1, \dots, v_n\}$ for V, then any bilinear form on V is determined by the values $\langle v_i, v_j \rangle$, for all i, j, since given $v, w \in V$, we can express $v = \sum a_i v_i, w = \sum b_j v_j$, so by bilinearity of bilinear forms, $\langle v, w \rangle = \langle \sum a_i v_i, \sum b_j w_j \rangle = \sum_{i,j} a_i b_j \langle v_i, v_j \rangle$. If you rewrite this in matrix form, you get that

$$\langle v, w \rangle = v^T M w \tag{5.3}$$

where $M_{ij} = \langle v_i, v_j \rangle$. Since any M gives us a formula for a bilinear form, we see that two bilinear forms are represented by the same matrix iff they are equal; and any matrix gives rise to a bilinear form. Thus, a choice of basis gives an isomorphism between the space of bilinear forms and the space of $n \times n$ matrices[3].

[1] Called by Hubbard & Hubbard a "concrete to abstract map".
[2] A linear operator is a linear transform from a vector space to itself.
[3] Actually, we haven't checked that this bijection is linear, but that is easily verified.

Note that the dot product (in a given basis) corresponds to the identity matrix. Every statement that one can make about a bilinear form can be rephrased as a statement about a matrix representing it.

(Skew)symmetric Bilinear Forms: A bilinear form is called symmetric (resp., skewsymmetric or antisymmetric) iff $\langle v, w\rangle = \langle w, v\rangle \, \forall v, w$ (resp., $\langle v, w\rangle = -\langle w, v\rangle \, \forall v, w$). Now a (skew)symmetric form clearly gives rise to a (skew)symmetric matrix. Conversely, given a symmetric matrix M, we see that the bilinear form that it determines is symmetric, since

$$\langle v, w\rangle = v^T M w = w^T M^T v = w^T M v = \langle w, v\rangle . \tag{5.4}$$

Note that we used the fact that every 1×1 matrix is symmetric. An analogous computation shows that a skewsymmetric matrix gives rise to a skewsymmetric form.

One reason that we care about (skew)symmetric bilinear forms is that there is a direct sum decomposition

$$\text{Bilinear}(V) = \text{Symmetric}(V) \oplus \text{SkewSym}(V); \tag{5.5}$$

this is similar to the fact that every function of a real (or complex!) variable is the sum of an odd function and an even function in a unique way. That is, let $f_o(x) = \frac{f(x)-f(-x)}{2}$, $f_e(x) = \frac{f(x)+f(-x)}{2}$; then $f = f_o + f_e$; this decomposition is unique because odd and even functions both form vector subspaces of the space of all real (or complex) functions on the real (complex) numbers, and their intersection is zero, since this is the only function such that $-f(x) = f(-x) = f(x)$. Exactly the same reasoning holds for bilinear forms: given a bilinear form $\langle -, -\rangle$, consider

$$\langle v, w\rangle_s = \frac{\langle v, w\rangle + \langle w, v\rangle}{2} \quad \langle v, w\rangle_{an} = \frac{\langle v, w\rangle - \langle w, v\rangle}{2}, \tag{5.6}$$

and apply the above reasoning.

Incidentally, these constructions are not just formally similar; they are true because of general results in *representation theory*, which is a field that allows you to study lots of stuff (much of math, physics and chemistry) using linear algebra. If you go further in any of these fields, you will see and use representation theory.

Change of Basis for Bilinear Forms: A very common computation is to be given a matrix M for a bilinear form with respect to one basis $\{e_i\}$, and want to find its matrix with respect to a different basis $\{v_i\}$. Here's how you do it:

Given coordinates for v, w with respect to the v_i (call these vectors of coordinates c, d), convert them both into coordinates with respect to the e_i, then feed them into M. That is,

$$\langle c, d\rangle = (Ac)^T M(Ad) = c^T(A^T M A)d. \tag{5.7}$$

Thus, the matrix with respect to the $\{v_i\}$ basis is $A^T M A$. This is how you transform a matrix for a *bilinear form*; this differs from linear operators in that we use the transpose instead of the inverse.

1: (a) With respect to the basis $\{v_1, v_2\}$, the matrix for f is $M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. The change of basis matrix from $\{v_1, v_2\}$ to $\{e_1, e_2\}$ is $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. We actually want the inverse of this, so the matrix of f with respect to $\{e_1, e_2\}$ is ${A^{-1}}^T M A^{-1}$, which you can compute by hand or using your favorite computer math program (mine is Dr. Genius!), and find $M' = \begin{pmatrix} 3/4 & 1/4 \\ -1/4 & 1/4 \end{pmatrix}$. In particular, $M'_{1,2} = f(e_1, e_2) = 1/4$.

(b) We showed this above (more generally) in (5.5).

(c) By our above construction, this is $\frac{f(e_1,e_2)+f(e_2,e_1)}{2} = \frac{1/4-1/4}{2} = 0$.

2: We showed parts (a) and (b) above.

For part (c), that B is a basis is equivalent to the statement that $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$ is invertible, which one can show by whatever method one likes, say row-reduction (yup, it works).

By the above,

$$M^B(f) = A^T I A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & -1 \\ 1 & -1 & 3 \end{pmatrix}. \tag{5.8}$$

(computed again using Dr. Genius)

3: (a) To show that $(P_3, (-,-))$ is a euclidean space, we need to show that the form is bilinear, symmetric, and positive.
Bilinearity follows from distributivity of multiplication across addition and linearity of integration, and symmetry is by symmetry of multiplication of real numbers/real-valued functions.
Positivity is harder; it's clear that $p^2(x) \geq 0$ for all x, and if $p \not\equiv 0$, then $p^2(x)$ is only zero at finitely many points. However, concluding from this that $\int_{-1}^{1} p^2(x)\,dx > 0$ takes a bit more work. You can do this for the special case of degree three polynomials by a direct calculation, but we will show a more general result which is in fact easier to prove.
A problem like this was asked on one of the first two math 25/55 sets my freshman year. Only two people got this part.

5.9. Lemma. $\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,dx$ *is an inner product on the space of continuous real-valued functions on* $[-1,1]$.

Proof. The only part we show is the positivity; the rest follows as above. Suppose $f \not\equiv 0$; then $\exists x_0$ such that $f(x_0) \neq 0$. Then $f^2(x_0) = \epsilon > 0$. By continuity of f^2, there exists some $\delta > 0$ such that for all x in a δ-ball (which we assume is in $(-1,1)$) of x_0, $f(x) > \epsilon/2$ (by δ-ϵ definition of continuity). So $\int_{x_0-\delta}^{x_0+\delta} f^2(x)\,dx \geq 2\delta\frac{\epsilon}{2} = \delta\epsilon > 0$, and thus $\langle f, f\rangle = \int_{-1}^{1} f^2 \geq \int_{x_0-\delta}^{x_0+\delta} f^2 \geq \delta\epsilon > 0$, which is what we needed to show. ☐

(b) Solving lots of linear equations or using gram-schmit yields:

$$a = 0 \quad b = 0 \quad c = \frac{-1}{3} \quad d = 0 \quad e = \frac{-3}{5} \quad f = 0, \tag{5.10}$$

which in turn yields

$$w_0 = 1 \quad w_1 = x \quad w_2 = x^2 - \frac{1}{3} \quad w_3 = x^3 - \frac{3}{5}x. \tag{5.11}$$

(c) Setting $r_i = \frac{1}{\|w_i\|}$ yields

$$r_0 = \sqrt{\frac{1}{2}} \quad r_1 = \sqrt{\frac{3}{2}} \quad r_2 = \sqrt{\frac{45}{8}} \quad r_3 = \sqrt{\frac{175}{8}}. \tag{5.12}$$

(d) I'm not going to prove this; the proof is formal and uninteresting.
More interesting is what you can do with these; that is, we call a set of polynomials that are orthogonal with respect to some inner product *orthogonal polynomials*, and you can use these to approximate functions. That is, given a function, project the function onto the subspace generated by each polynomial, and take the sum of these projections. This gives you a best fit polynomial (of a given degree) for the function, and this technique is used extensively in modeling situations.

4: The first part follows by (5.1).
For the second part, it asks for the (i,j)-th entry of the matrix $A^T I A = A^T A$, which is $\sum_{k=1}^{d} a_i^k a_j^k$.