Homework #2. Solutions to selected problems

Problems:

For problems (or their parts) marked with a *, a hint is given later in the assignment. Do not to look at the hint(s) until you seriously tried to solve the problem without it.

1. Let V and H be as in Problem 3 of Homework 1.

(a) Prove that H is positive definite directly from definition. You will need some basic facts from real analysis to make the argument rigorous.

(b) Now use the "modified Gram-Schmidt process" (that is, the algorithm from the proof of Theorem 3.4 from class) to find a basis β such that $[H]_\beta$ is the identity matrix.

Solution to 1(b) One possible answer is

$$\{1, \sqrt{3}(2x-1), \sqrt{5}(6x^2-6x+1)\}.$$

While this is not the only answer, even up to permutation (in fact, there are infinitely many possible bases), there are several standard algorithms that produce the above basis.

One possibility is to use the standard Gram-Schmidt process (see § 6.2 of Friedberg-Insel-Spence). Note that it is applicable since we know from (a) that the form is positive definite.

Another possibility is to use the algorithm applying simultaneous row and column operations to the matrix $[H]_{std}$ (where $std = \{1, x, x^2\}$ is the standard basis of $Pol_2(\mathbb{R})$). We briefly discussed this algorithm at the end of the problem session on Fri, Sep 8; the algorithm is described in Friedberg-Insel-Spence on pp. 431-433.

Below we will discuss yet another algorithm which is fairly similar to Gram-Schmidt and follows the proof of the diagonalization Theorem 3.4 given in class. So, suppose H is a symmetric bilinear form on a finite-dimensional vector space V over a field F with $\mathrm{char}(F) \neq 2$. Let $\{v_1, \ldots, v_n\}$ be any basis of V with $H(v_1, v_1) \neq 0$ (we know that such basis exists). Our goal is to find another basis γ such that $[H]_\gamma$ is diagonal (if H is positive definite, once we

found such γ, constructing a basis β with $[H]_\beta = I_n$ is easy – we just need to rescale elements of γ).

The vector v_1 will be the first element of γ. If V_1 is the span of v_1, then $H_{|V_1}$ is non-degenerate, whence $V = V_1 \oplus V_1^\perp$; in particular $\dim V_1^\perp = n-1$. For $2 \leq i \leq n$ define $w_i = v_i - \frac{H(v_i,v_1)}{H(v_1,v_1)} v_1$. Then by direct computation $H(w_i, v_1) = 0$ for each i, so $w_i \in V_1^\perp$. Also it is easy to see that $w_2, \ldots, w_n$ are linearly independent and hence must form a basis of $V_1^\perp$.

Then we apply the same procedure to $V_1^\perp$ etc. Note that if $H(w_i, w_i) \neq 0$ for some $2 \leq i \leq n$, we can use that w_i as the second vector of the basis γ; otherwise, we will need to change the basis of $V_1^\perp$. Fortunately, if H is positive definite, this issue will never arise.

Applying the above algorithm in our case starting with $v_1 = 1, v_2 = x, v_3 = x^2$, we first get $w_2 = x - \frac{1}{2}$ and $w_3 = x^2 - \frac{1}{3}$. We then make w_2 the second basis of γ and finally compute $z_3 = w_3 - \frac{H(w_3,w_2)}{H(w_2,w_2)} w_2 = (x^2 - \frac{1}{3}) - (x - \frac{1}{2}) = (x^2 - x + \frac{1}{6})$ as the third vector of γ. After rescaling v_1, w_2, z_3, we get the answer stated at the beginning of the problem.

2. Let $V = Mat_n(\mathbb{R})$ for some $n \in \mathbb{N}$, and let H be the bilinear form on V given by $H(A, B) = Tr(AB)$. Prove that H is symmetric and compute its signature. It may be a good idea to start with $n = 2$ and $n = 3$.

Solution: Proof of bilinearity is straightforward. Let us check that H is symmetric: if $A = (a_{ij})$ and $B = (b_{ij})$ since

$$Tr(AB) = \sum_{k=1}^{n} \sum_{i=1}^{n} a_{ki} b_{ik} = \sum_{i,k=1}^{n} a_{ki} b_{ik} = \sum_{i=1}^{n} \sum_{k=1}^{n} b_{ik} a_{ki} = Tr(BA).$$

We now compute the signature. Order the standard basis of V (consisting of the matrix units e_{ij}) as follows: diagonal matrix units e_{ii} come first; and then each matrix e_{ij} with $i < j$ is followed by its transpose e_{ji}, e.g. $e_{12}, e_{21}, e_{13}, e_{31}$ etc. If β is the obtained ordered basis, then (by direct computation) $[H]_\beta$ is the block-diagonal matrix which starts with n occurrences of 1 on the diagonal followed by $\frac{n^2-n}{2}$ 2×2 blocks $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

As discussed at the end of Lecture 4, the signature of a block-diagonal matrix is the sum of the signatures of its diagonal blocks. Thus, $sig(H) = n(1,0) + \frac{n^2-n}{2} sig(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix})$. By Homework#3.0 we have $sig(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}) = (1,1)$ and therefore $sig(H) = (n,0) + (\frac{n^2-n}{2}, \frac{n^2-n}{2}) = (\frac{n^2+n}{2}, \frac{n^2-n}{2})$

3. The goal of this problem is to prove the following theorem:

Theorem: *Let F be a finite field with* $\mathrm{char}(F) \neq 2$*, V a finite-dimensional vector space and H a symmetric bilinear form on V. Then there exists a basis β of V such that $[H]_\beta$ is diagonal and at MOST one entry of $[H]_\beta$ is different from 0 or 1 (in particular, if H is non-degenerate, then there exists a basis β such that $[H]_\beta = diag(1, \ldots, 1, \lambda)$ for some $\lambda \in F$).*

If you do not feel comfortable working with arbitrary finite fields, you can assume that $F = \mathbb{Z}_p$ for some $p > 2$ (this does not substantially simplify the problem).

(a) Let Q be the set of squares in F, that is, $Q = \{f \in F : f = x^2 \text{ for some } x \in F\}$. Prove that $|Q| = \frac{|F|+1}{2}$.

(b) Now take any nonzero $a, b \in F$. Use (a) to prove that there exist $x, y \in F$ such that $ax^2 + by^2 = 1$.

(c) Now use (b) to prove the above Theorem.

Solution: (a) Since $\mathrm{char}(F) \neq 2$, the equation $x = -x$ has exactly one solution $x = 0$ (if $\mathrm{char}(F) = 2$, every x is a solution). Thus, we have the partition of $F \setminus \{0\}$ into pairs $\{(a, -a)\}$. Elements in each pair have the same square, and elements coming from different pairs must have different squares (this is because for any field F and for any nonconstant polynomial $f(x) \in F[x]$ the number of solutions to the equation $f(x) = 0$ does not exceed the degree of f). Thus, the total number of distinct squares in F is $\frac{|F|-1}{2} + 1 = \frac{|F|+1}{2}$ (where $+1$ comes from counting 0).

(b) The equation can be rewritten as $ax^2 = 1 - by^2$. Since a and b are nonzero and F is a field, the functions $t \mapsto at$ and $t \mapsto 1 - bt$ from F to F are injective. Thus by (a) there are $\frac{|F|+1}{2} = |Q|$ possible values for the expression ax^2 and $\frac{|F|+1}{2} = |Q|$ possible values for the expression $1-by^2$. Since $2\frac{|F|+1}{2} = |F| + 1 > |F|$, there must exist at least one $f \in F$ representable as ax^2 and representable as $1 - by^2$. The obtained pair (x, y) is then a solution to $ax^2 = 1 - by^2$.

(c) We will give a solution which does not follow the hint directly. We argue by induction on n. In the base case $n = 1$ there is nothing to prove. Assume now that $n > 1$ and we proved the result for smaller values of n.

By Theorem 3.4 we know that there exists a basis $\gamma = \{v_1, \ldots, v_n\}$ such that $[H]_\gamma$ is diagonal. Without loss of generality we can assume that all zeroes on the diagonal of $[H]_\gamma$ (if exist) appear at the end. If $[H]_\gamma$ has at most one nonzero entry, then it already has the required form. So assume

that $[H]_\gamma$ has at least two nonzero entries. By our assumption this means that $a = H(v_1, v_1)$ and $b = H(v_2, v_2)$ are both nonzero.

By (b) there exist $x, y \in F$ such that $ax^2+by^2 = 1$. Let now $w = xv_1+yv_2$. Then $H(w, w) = x^2H(v_1, v_1) + y^2H(v_2, v_2) + 2xyH(v_1, v_2) = ax^2 + by^2 = 1$. From this point on we can simply imitate the induction step in the proof of Theorem 3.4 using w as the first element of our new basis.

More precisely, let $W = Fw$. Since $\dim(W) = 1$ and H is nonzero on W, the restriction $H_{|W}$ is nondegenerate and hence $V = W \oplus W^\perp$ by Lemma 5.3. Since $\dim(W^\perp) = n - 1 < n$, by induction hypothesis there exists a basis β_1 such that $[H_{|W^\perp}]_{\beta_1} = diag(a_2, \ldots, a_n)$ with at most one of the a_i's is different from 0 and 1. Since $H(w, w) = 1$, if we let $\beta = \{w, \beta_1\}$, then $[H]_\beta = diag(1, a_2, \ldots, a_n)$, so β has the required property.

4. Let H be a bilinear form on a finite-dimensional vector space V. In class we proved that for any subspace W of V we have $\dim(W)+\dim(W^\perp) \geq \dim(V)$ (Lemma 3.2) where $W^\perp$ is the orthogonal complement of W with respect to H. Prove that if H is non-degenerate, then $\dim(W)+\dim(W^\perp) = \dim(V)$. One way to prove this is to show that the map ϕ from the proof of Lemma 3.2 is surjective.

Solution: As in Lecture 3, let $m = \dim(W)$, choose a basis $\{w_1, \ldots, w_m\}$ of W, and define a map $\phi : V \to F^m$ by $\phi(v) = (H(w_1, v), \ldots, H(w_m, v))^T$. By the argument from class, equality $\dim(W) + \dim(W^\perp) = \dim(V)$ holds if and only if ϕ is surjective.

We shall assume that ϕ is not surjective and reach a contradiction. So, let $Y = Im(\phi)$, and assume that $Y \neq F^m$. Let $Z = Y^\perp$, the orthogonal complement of Y with respect to the dot product on F^m. Since Y is not the entire space, $Z \neq 0$, so we can choose a nonzero $z \in Z$ and write $z = (\lambda_1, \ldots, \lambda_m)^T$.

By definition of Z, for every $v \in V$ we have $(\lambda_1, \ldots, \lambda_m)\phi(v) = 0$, that is, $\sum_{i=1}^m \lambda_i H(w_i, v) = 0$. Since H is bilinear, we get $H(\sum_{i=1}^m \lambda_i w_i, v) = 0$ for all $v \in V$.

Since H is nondegenerate, we must have $\sum_{i=1}^m \lambda_i w_i = 0$. But w_i are linearly independent, so we must have $\lambda_i = 0$ for each i. This is impossible since by construction $z = (\lambda_1, \ldots, \lambda_m)^T$ is nonzero.

5. In this problem we discuss linear maps and bilinear forms on vector spaces of (infinite) countable dimension over an arbitrary field F. One example of such a space is F_{fin}^∞, the set of (infinite) sequences of elements of F in which only finitely many elements are nonzero. The set $\{e_1, e_2, \ldots\}$ is a basis of F_{fin}^∞ where e_i is the sequence whose i^{th} element is 1 and all other

elements are 0.

Now let V be any countably-dimensional vector space over F and $\beta = \{v_1, v_2, \ldots\}$ a basis of V. Any $v \in V$ is a linear combination of finitely many elements of β, so we can write $v = \sum_{i=1}^{n} \lambda_i v_i$ for some n (if some v_i with $i \leq n$ does not appear in the expansion of v, we simply let $\lambda_i = 0$). Define $[v]_\beta = (\lambda_1, \ldots, \lambda_n, 0, 0, \ldots) \in F_{fin}^{\infty}$.

(a) (practice) Prove that the map $\phi : V \to F_{fin}^{\infty}$ given by $\phi(v) = [v]_\beta$ is an isomorphism of vector spaces.

Denote by $Mat_\infty(F)$ the set of all matrices with countably many rows and columns whose entries are in F. Given a bilinear form H on V, let $[H]_\beta \in Mat_\infty(F)$ be the matrix whose (i,j)-entry is $H(v_i, v_j)$.

(b) Prove that $H(v,w) = [v]_\beta^T [H]_\beta][w]_\beta$ for any $v, w \in V$ (here we consider $[v]_\beta$ and $[w]_\beta$ as columns). In particular, explain why the expression on the right-hand side is well defined even though $[H]_\beta$ is an infinite-size matrix.

(c) Prove that the map $\Phi : Bil(V) \to Mat_\infty(F)$ given by $\Phi(H) = [H]_\beta$ is an isomorphism of vector spaces.

Now let $T \in \mathcal{L}(V)$ be a linear map from V to V. Define $[T]_\beta \in Mat_\infty(F)$ to be the matrix whose i^{th} column is $[Tv_i]_\beta$.

(d) Prove that the map $\Psi : \mathcal{L}(V) \to Mat_\infty(F)$ given by $\Psi(T) = [T]_\beta$ is linear and injective, but not surjective, and explicitly describe its image.

Solution: Some parts of the above statements are quite straightforward and their proofs will be skipped.

(a) Since β spans V, every $v \in V$ can be written as a FINITE sum $v = \sum_{i=1}^{n} \lambda_i v_i$ for some $n \in \mathbb{N}$ and some $\lambda_i \in F$. Equivalently, we can write $v = \sum_{i=1}^{\infty} \lambda_i v_i$ where only finitely many λ_i are nonzero. Such an expression is also unique since β is linearly independent.

By definition of ϕ we have $\phi(\sum_{i=1}^{\infty} \lambda_i v_i) = (\lambda_1, \lambda_2, \ldots)^T$. From this description it is clear that ϕ is bijective and linear.

(b) Set $a_{ij} = H(v_i, v_j)$, so that $[H]_\beta = (a_{ij})$. Take any $v, w \in V$. Then $v = \sum_{i=1}^{\infty} \lambda_i v_i$ and $w = \sum_{i=1}^{\infty} \mu_i v_i$ where only finitely many λ_i and μ_i are nonzero.

We have

$$[v]_\beta[H]_\beta[w]_\beta = (\lambda_1, \lambda_2, \ldots)(a_{ij})(\mu_1, \mu_2, \ldots)^T = \sum_{i,j=1}^{\infty} \lambda_i \mu_j a_{ij}.$$

Since only finitely many λ_i and μ_i are nonzero, the above sum also has only finitely many nonzero terms (and thus is well defined). Also $\sum_{i,j=1}^{\infty} \lambda_i \mu_j a_{ij} = \sum_{i,j=1}^{\infty} \lambda_i \mu_j H(v_i, v_j) = H(\sum_{i=1}^{\infty} \lambda_i v_i, \sum_{i=1}^{\infty} \mu_i v_i) = H(v, w)$, so $H(v, w) = [v]_\beta[H]_\beta[w]_\beta$, as desired.

(c) We will only explain why Φ is surjective. Take any $A \in Mat_\infty(F)$, and define $H : V \times V \to F$ by $H(v, w) = [v]_\beta^T A[w]_\beta$. The expression on the right hand side is well defined by the same computation as in (b). Also H is clearly bilinear and $H(v_i, v_j) = e_i^T A e_j = (i, j)$-entry of A, so $A = [H]_\beta = \Phi(H)$, as desired.

(d) As in (c), we will only address the image/non-surjectivity part. Let $Mat_\infty^0(F)$ be the set of all $\infty \times \infty$ matrices such that every column has only finitely many nonzero entries. We claim that $Im(\Psi) = Mat_\infty^0(F)$ (so in particular, Ψ is not surjective).

The inclusion $Im(\Psi) \subseteq Mat_\infty^0(F)$ holds since each column of $\Psi(T)$ is the vector $[Tv_i]_\beta$ for some i, and by definition such a vector lies in F_∞^0 (that is, has only finitely many nonzero entries).

Conversely, let $A = (a_{ij}) \in Mat_\infty^0(F)$. We need to find $T \in \mathcal{L}(V)$ such that $[T]_\beta = A$. Define $T : V \to V$ by

$$T\left(\sum_{j=1}^{\infty} \lambda_j v_j\right) = \sum_{j=1}^{\infty} \lambda_j \sum_{i=1}^{\infty} a_{ij} v_i \qquad (***)$$

for every $\sum_{j=1}^{\infty} \lambda_j v_j \in V$.

Since $\sum_{j=1}^{\infty} \lambda_j v_j \in V$, only finitely many λ_j are nonzero, so the outer sum in (***) (the sum over j) has only finitely many nonzero terms. For each j, the inner sum in (***) also has only finitely many nonzero terms precisely because $(a_{ij}) \in Mat_\infty^0(F)$. Hence the entire sum in (***) has finitely many nonzero terms and is thus well defined.

It is clear that T defined by (***) is linear, and a straightforward computation shows that $[T]_\beta = A$, as desired.