Math 25a Solution Set #10 (Part C)

Isidora Milin

December 2001

Problem 1

This was problem 3.5.1 in the textbook.

Let V be a vector space. A *symmetric bilinear function* on V is a mapping $B : V \times V \to \mathbb{R}$ such that:

(1) $B(av_1+bv_2, w) = aB(v_1, w)+bB(v_2, w)$ for all $v_1, v_2, w \in V$ and $a, b \in \mathbb{R}$;

(2) $B(v, w) = B(w, v)$ for all $v, w \in V$.

Part (a). You were asked to show the following implication:

$$A \in Mat(n, n) \text{ symmetric } \Rightarrow B_A(v, w) = v^T Aw \text{ symmetric bilinear form.}$$

All that needs to be done is to check for bilinearity and simmetry:

(1) **Bilinearity.**

$$B_A(av_1+bv_2, w) = (av_1+bv_2)^T Aw = av_1^T Aw+bv_2^T w = aB_A(v_1, w)+bB_A(v_2, w)$$

Above we have used the fact that the transpose map is a linear map from $Mat(n, n)$ to itself.

(2) **Symmetry.** We know A is symmetric iff $A = A^T$. Also, since $v_1^T Av_2 \in \mathbb{R}$ is a scalar for any $v_1, v_2 \in V$, it equals its own transpose. Using this, we get:

$$B_A(v, w) = v^T Aw = (v^T Aw)^T = w^T A^T (v^T)^T = w^T Av = B_A(w, v)$$

Above we have used that for any two matrices A, B which can be multiplied, we have $(AB)^T = B^T A^T$.

Note that because of symmetry, B has to be bilinear in both coordinates, namely:

$$B(v, aw_1+bw_2) = B(aw_1+bw_2, v) = aB(w_1, v)+bB(w_2, v) = aB(v, w_1)+bB(v, w_2)$$

Part (b). Here you were asked to show that every symmetric bilinear function on $\mathbb{R}^n$ is of the form B_A for a unique symmetric matrix A.

The word *unique* confused a lot of people here. What you needed to do was to show that $\forall B$, some A would work, but also that two different matrices A

cannot work for the same B. Common mistake was to show how, given some symm. bilin. form B, matrix A could be constructed and then from the fact that the described construction yields a unique result conclude A is unique. What is neglected here is the question of the uniqueness of your construction; You had to show there was no other way to derive A that works but the one you used! Thus, when you have to prove unique existence of something, a possible strategy is to prove what it has to be given the conditions, rather than exhibit one and then still rake your brain as to why there aren't any others.

Let B be our SBF on $\mathbb{R}^n$ and suppose $A = [a_{ij}] \in Mat(n,n)$ is such that $B = B_A$. Let e_i, $i \in \{1,\dots,n\}$ be the standard basis on $\mathbb{R}^n$. Then we must have:

$$a_{ij} = e_i^T A e_j = B(e_i, e_j)$$

which completely determines matrix A, thus excluding the possibility of multiple solutions.

Now consider the matrix $A = [a_{ij}]$ given by the formula above. It is symmetric since B is, and the only thing left to check is that it actually works, i.e. that $B(v,w) = v^T A w \ \forall v, w \in \mathbb{R}^n$. So take any two vectors in $\mathbb{R}^n$, write them as $v = v_1 e_1 + \ldots + v_n e_n$ and $w = w_1 e_1 + \ldots + w_n e_n$, and use bilinearity of B:

$$\begin{aligned}
B(v,w) &= B(v_1 e_1 + \ldots + v_n e_n, w_1 e_1 + \ldots + w_n e_n) \\
&= \sum_{i=1}^{n} (\sum_{j=1}^{n} v_i w_j B(e_i, e_j)) \\
&= \sum_{i=1}^{n} (\sum_{j=1}^{n} v_i w_j a_{ij}) \\
v^T A w &= \begin{pmatrix} v_1 & \ldots & v_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \\
&= \begin{pmatrix} \sum_{i=1}^n a_{i1} v_i & \sum_{i=1}^n a_{i2} v_i & \ldots & \sum_{i=1}^n a_{in} v_i \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \\
&= \sum_{j=1}^{n} (\sum_{i=1}^{n} a_{ij} v_i w_j)
\end{aligned}$$

We conclude $\forall v, w \ B(v,w) = v^T A w$, so $B = B_A$ and we're done.

Part (c). For P_k the space of polynomials of degree at most k, we check that $B : P_k \times P_k \to \mathbb{R}$ defined by $B(p,q) = \int_0^1 p(t)q(t)dt$ is a symmetric bilinear function. This involves just the basic properties of the integral of a single variable:

Bilinearity.

$$\begin{aligned} B(ap_1+bp_2,q) &= \int_0^1 (ap_1(t)+bp_2(t))q(t)dt \\ &= a\int_0^1 p_1(t)q(t) + b\int_0^1 p_2(t)q(t)dt = aB(p_1,q)+bB(p_2,q) \end{aligned}$$

Symmetry.

$$B(p,q) = \int_0^1 p(t)q(t)dt = \int_0^1 q(t)p(t)dt = B(q,p)$$

Part (d). We let $p_1(t) = 1, p_2(t) = t, \ldots, p_{k+1}(t) = t^k$ be the usual basis of P_k. The "concrete to abstract" linear transformation this problem mentions is just a fancy way of saying that P_k is isomorphic to $\mathbb{R}^{k+1}$.

$$\begin{aligned} \Phi_p :\ \mathbb{R}^{k+1} &\to P_k \\ (a_1, a_2, \ldots, a_{k+1}) &\mapsto a_1 + a_2 t + \ldots + a_{k+1} t^k \end{aligned}$$

Notice that Φ_p is a bijective linear transformation that sends the standard basis of $\mathbb{R}^{k+1}$ to the usual basis of P_k. Using the fact that B (from part (c)) is a SBF, $B_0(a,b) = B(\Phi_p(a), \Phi_p(b))$ is now easily verified to be bilinear symmetric:

$$\begin{aligned} B_0(\alpha a_1 + \beta a_2, b) &= B(\Phi_p(\alpha a_1 + \beta a_2), \Phi_p(b)) = B(\alpha \Phi_p(a_1) + \beta \Phi_p(a_2), \Phi_p(b)) \\ &= \alpha B(\Phi_p(a_1), \Phi_p(b)) + \beta(\Phi_p(a_2), \Phi_p(b)) = \alpha B_0(a_1, b) + \beta B_0(a_2, b) \end{aligned}$$

$$B_0(a,b) = B(\Phi_p(a), \Phi_p(b)) = B(\Phi_p(b), \Phi_p(a)) = B_0(b,a)$$

To calculate the matrix A_0 of B_0 we use the result of part (b) which gives us an explicit formula for $a_{ij} = [A_0]_{ij}$:

$$\begin{aligned} a_{ij} &= B_0(e_i, e_j) = B(\Phi_p(e_i), \Phi_p(e_j)) = B(p_i, p_j) \\ &= \int_0^1 t^{i-1} t^{j-1} dt = \int_0^1 t^{i+j-2} dt = \frac{1}{i+j-1} \end{aligned}$$

where $i, j \in \{1, \ldots, k+1\}$.

Problem 2

This problem describes the connection between bilinear symmetric functions (BSFs) and quadratic forms. If B is a BSF, we can associate to it a function $Q_B : V \to \mathbb{R}$, $v \mapsto B(v,v)$. We show here that every quadratic form on $\mathbb{R}^n$ is of the form Q_B for some symmetric bilinear function B.

Let B be any BSF. We begin by exploring the connection between the matrix $A = [a_{ij}]$ of B and the quadratic form Q_B. We have that $\forall x = (x_1, \dots, x_n) \in \mathbb{R}^n$:

$$Q_B(x) = B(x,x) = x^T A x = \sum_{i=1}^{n}(\sum_{j=1}^{n} a_{ij}x_i x_j) = \sum_{i<j} 2a_{ij}x_i x_j + \sum_{i} a_{ii}x_i^2$$

with the last equality following from the fact that A is symmetric, so that $a_{ij} = a_{ji}$ for all i, j. We group the coefficients of $x_i x_j$ and $x_j x_i$ together because the quadratic form is usually written in such a way.

Now suppose that we are given a quadratic form $Q(x_1, \dots, x_n) = \sum_{i=1}^{n} q_{ii}x_i^2 + \sum_{i<j} q_{ij}x_i x_j$ on $\mathbb{R}^n$. Notice that q_{ij}'s do not correspond completely to a_{ij}'s. In particular, quadratic form is given by $\frac{n(n+1)}{2}$ coefficients. But it is easy to reorganize q_{ij}'s so that they fit in with the previous paragraph. We define a matrix $A = [a_{ij}]$ as follows:

$$a_{ij} = \begin{cases} \frac{q_{ij}}{2}, & \text{for } i < j \\ \frac{q_{ji}}{2}, & \text{for } i > j \\ q_{ii}, & \text{for } i = j \end{cases}$$

After substituting these values in the formula in the previous paragraph, we conclude $Q = Q_{B_A}$.

Problem 3

Part (a). We show that $A^T A$ is symmetric by showing it equals its own transpose. Note that here A is any matrix, not necessarily square.

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

This has been done in a previous homework, although perhaps only for square matrices, but that does not change the proof. Note again how brute-force approach of writing out entries as sums can be avoided.

Part (b). This part makes two claims:
(1) All eigenvalues λ of $A^T A$ are ≥ 0.
(2) All eigenvalues λ of $A^T A$ are $> 0 \Leftrightarrow ker(A) = 0$

The crucial fact you needed to use for this problem was that any symmetric matrix has an orthonormal basis consisting of eigenvectors. So let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of $A^T A$ (not necessarily all distinct) and $\{v_1, \dots, v_n\}$ the orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A.

The proof of (1) is fairly straightforward. If λ is any eigenvalue of $A^T A$, we have that $A^T A v = \lambda v$ for some $v \in \mathbb{R}^n$ such that $|v| = 1$, so that:

$$|Av|^2 = (Av)^T(Av) = (v^T A^T)Av = v^T(A^T A v) = v^T(\lambda v) = \lambda(v^T v) = \lambda |v|^2$$

We conclude that $\lambda = \frac{|Av|^2}{|v|^2} \geq 0$ and we are done.

Second statement is shown separately for each direction:

$\Rightarrow$ We assume that $\lambda_i \neq 0 \ \forall i$, so that all eigenvalues of $A^T A$ are positive, and want to show A is injective. Notice that $v \in ker(A) \Rightarrow v \in ker(A^T A)$ so that $ker(A) \subset ker(A^T A)$ and it is enough to show $A^T A$ is injective. Suppose $v = a_1 v_1 + \ldots + a_n v_n \in ker(A^T A)$.Then:

$$0 = A^T A v = A^T A(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i A^T A v_i = \sum_{i=1}^{n} a_i \lambda_i v_i$$

Since $\{v_1, \ldots, v_n\}$ is a basis for $\mathbb{R}^n$, we must have $a_i \lambda_i = 0 \ \forall i$. But we know that all $\lambda_i \neq 0$, so that $\forall i \in \{1, \ldots, n\}$ $a_i = 0$ and hence $v = 0$. Thus, $ker(A^T A) = 0$, and $ker(A) = 0$ as well.

$\Leftarrow$ Now, assume that $ker(A) = 0$. If some eigenvalue $\lambda_i = 0$, we have (where v_i is its eigenvector):

$$A^T A v_i = \lambda_i v_i = 0 \Rightarrow 0 = v_i^T A^T A v_i = (Av_i)^T (Av_i) = |Av_i|^2 \Rightarrow |Av_i| = 0 \Rightarrow Av_i = 0$$

so that we get a contradiction with $ker(A) = 0$. Thus, all eigenvalues are strictly positive in this case.

Part (c). We show that $\|A\| = \sup_{\lambda \text{ eigenvalue of } A^T A} \sqrt{\lambda}$.

Let us denote the largest eigenvalue of $A^T A$ by λ_{max}. We know that $\|A\| = \sup_{v \neq 0} \frac{|Av|}{|v|}$, so we show that $\forall v \in \mathbb{R}^n$ $\frac{|Av|}{|v|} \leq \sqrt{\lambda_{max}}$ and that $\exists v \in \mathbb{R}^n$ $\frac{|Av|}{|v|} = \sqrt{\lambda_{max}}$ and that suffices.

Any $v \in \mathbb{R}^n$ can be written as $v = \sum_{i=1}^n a_i v_i$, where v_i's are the orthonormal basis of eigenvectors.

$$\begin{aligned}
|Av|^2 &= (Av)^T Av = v^T (A^T A v) \\
&= v^T (A^T A (a_1 v_1 + \ldots + a_n v_n)) \\
&= (\sum_{i=1}^{n} a_i v_i)(\sum_{i=1}^{n} a_i \lambda_i v_i) \\
&= \sum_{i=1}^{n} (\sum_{j=1}^{n} \lambda_i a_i a_j (v_i \cdot v_j)) \\
&= \sum_{i=1}^{n} a_i^2 \lambda_i \\
&\leq \lambda_{max} (\sum_{i=1}^{n} a_i^2) \\
&= \lambda_{max} |v|
\end{aligned}$$

We conclude that $\|A\| \leq \lambda_{max}$. But if v_{max} is the unit eigenvector corresponding to λ_{max}, we get that

$$|Av_{max}|^2 = v_{max}^T A^T A v_{max} = \lambda_{max} v_{max}^T v_{max} = \lambda_{max}$$

so that $\sqrt{\lambda_{max}} = \frac{|Av_{max}|}{|v_{max}|}$ and hence $\|A\| = \sqrt{\lambda_{max}}$.