Solutions for Problem Set 2

1. The Trace of a square matrix is defined to be the sum of its diagonal elements. If PQ and QP are square products of finite-dimensional rectangular (perhaps not square) matrices, prove that Trace(PQ) = Trace(QP) . (Other notations for "Trace" are "tr" and "sp", the latter from the German word "Spur" that means what "spoor" means in English.)

$\mathrm{Trace}(PQ) = \sum_i (\sum_j p_{ij}\, q_{ji})$ and $\mathrm{Trace}(QP) = \sum_j (\sum_i q_{ji}\, p_{ij})$, which differ only in the order of summation. For sums of finitely many terms the order doesn't matter.

2. Can the equation $XY - YX = I$ be solved for matrices X and Y ? Explain.

No, because $\mathrm{Trace}(XY - YX) = \mathrm{Trace}(XY) - \mathrm{Trace}(YX) = 0 < \mathrm{Trace}(I)$.

(If $\mathrm{Trace}(Z) = 0$ then $XY - YX = Z$ has infinitely many solutions X and Y but finding some is not so easy.)

3. For any linear operator **L** that maps a finite-dimensional vector space to itself, show that every matrix L that represents **L** in some basis has the same Trace. (It defines Trace(**L**) .)

For any basis **B** matrix $L := \mathbf{B}^{-1}\mathbf{L}\mathbf{B}$ represents **L** . For any other basis **B**C , where C is an invertible matrix, matrix $C^{-1}LC$ represents **L** . $\mathrm{Trace}(C^{-1}LC) = \mathrm{Trace}(LCC^{-1}) = \mathrm{Trace}(L)$ as claimed.

4. Prove that finite-dimensional matrix multiplication is associative. Exhibit an example of non-associative infinite-dimensional matrix multiplication.

The element in row i and column j of (PQ)R is $\sum_m (\sum_k p_{ik}\, q_{km})\, r_{mj}$; the corresponding element of P(QR) is $\sum_k p_{ik} (\sum_m q_{km}\, r_{mj})$. They differ only in the order of summation, which doesn't matter if all dimensions are finite. However, infinite dimensions are another story; try

$P^T = \begin{bmatrix} 1 & 1 & 1 & \dots \\ -1 & -1 & -1 & \dots \end{bmatrix}$, $Q = \begin{bmatrix} 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & \dots \end{bmatrix}$ and $r^T = \begin{bmatrix} 1 & 1 & 1 & \dots \end{bmatrix}$ so $(PQ)r = Or = o$ but $P(Qr) \longrightarrow P\infty$.

5. An operator **L** that maps one vector space to another or to itself is called " linear " just when $\mathbf{L}(ß\mathbf{x} + \mu\mathbf{y}) = ß\mathbf{L}\mathbf{x} + \mu\mathbf{L}\mathbf{y}$ for all scalars ß and μ and all vectors **x** and **y** in Domain(**L**) . Until now we have taken for granted, as if obvious, that multiplying two linear operators produces another, if it is defined, and that multiplication of linear operators is associative; prove all that.

If Domain($\mathbf{Q}$) includes Range($\mathbf{R}$) for linear operators $\mathbf{Q}$ and $\mathbf{R}$ then operator $\mathbf{QR}$ is defined by the assertion " $(\mathbf{QR})\mathbf{x} := \mathbf{Q}(\mathbf{Rx})$ for every $\mathbf{x}$ in Domain($\mathbf{R}$) ." This definition implies that
$(\mathbf{QR})(ß\mathbf{x} + \mu\mathbf{y}) = \mathbf{Q}(\mathbf{R}(ß\mathbf{x} + \mu\mathbf{y})) = \mathbf{Q}(ß\mathbf{Rx} + \mu\mathbf{Ry}) = ß\mathbf{Q}(\mathbf{Rx}) + \mu\mathbf{Q}(\mathbf{Ry}) = ß(\mathbf{QR})\mathbf{x} + \mu(\mathbf{QR})\mathbf{y}$,
so $\mathbf{QR}$ is linear too. If also Domain($\mathbf{P}$) includes Range($\mathbf{Q}$) then, for every $\mathbf{x}$ in Domain($\mathbf{R}$) ,
$((\mathbf{PQ})\mathbf{R})x = (\mathbf{PQ})(\mathbf{Rx}) = \mathbf{P}(\mathbf{Q}(\mathbf{Rx})) = \mathbf{P}((\mathbf{QR})\mathbf{x}) = (\mathbf{P}(\mathbf{QR}))\mathbf{x}$, so $(\mathbf{PQ})\mathbf{R} = \mathbf{P}(\mathbf{QR})$ as claimed.

The foregoing arguments may seem trivial, and they are if dimensions are finite. But when dimensions are infinite, an assertion like " Domain($\mathbf{Q}$) includes Range($\mathbf{R}$) " can be difficult to confirm; it can be false if Range($\mathbf{R}$) includes a vector $\mathbf{r}$ at which $\mathbf{Qr}$ is undefined because of a failure of convergence, as occurs in the example in Problem **4** above. Moreover, the associativity of operator multiplication cannot be inferred directly from the associativity of matrix multiplication since the representation of operators by matrices assumes the associativity of products of operators like bases (linear maps from column vectors to abstract vectors) and their inverses. It all works the other way around; the formula defining matrix multiplication can be derived from the associativity and distributivity of linear operators' multiplication by a simple but tedious argument. Can you find it? The formula is fore-ordained because linear operators are abstractions inspired by analogies between matrix multiplication and other operations like differentiation, integration and geometrical rotation. Matrix multiplication was used unwittingly by the Chinese to solve linear equation systems as long ago as 300 BC (the date is unsure because a Chinese Emperor burnt all books in 213 BC), and turned up in determinants and the Chain Rule for partial derivatives in eighteenth century Europe. Matrix algebra as we know it now began in mid-nineteenth century England and the U.S.

6. $\mathbf{H}$ is a real symmetric bilinear form just when real scalar $\mathbf{Hxz} = \mathbf{Hzx}$ is linear in $\mathbf{x}$ and $\mathbf{z}$ separately; $\mathbf{Hz}(ß\mathbf{x} + \mu\mathbf{y}) = ß\mathbf{Hzx} + \mu\mathbf{Hzy}$. Any given $\mathbf{H}$ defines a *Quadratic form* thus: $Q(\mathbf{x}) := \mathbf{Hxx}$; verify that this Q satisfies the *Parallelogram Identity*

$$Q(\mathbf{x}+\mathbf{y}) + Q(\mathbf{x}-\mathbf{y}) = 2Q(\mathbf{x}) + 2Q(\mathbf{y}) .$$

Show how to use any procedure that computes Q to compute $\mathbf{H}$ too.

$$\begin{aligned} Q(\mathbf{x}+\mathbf{y}) + Q(\mathbf{x}-\mathbf{y}) &= \mathbf{H}(\mathbf{x}+\mathbf{y})(\mathbf{x}+\mathbf{y}) + \mathbf{H}(\mathbf{x}-\mathbf{y})(\mathbf{x}-\mathbf{y}) \\ &= \mathbf{Hxx} + \mathbf{Hxy} + \mathbf{Hyx} + \mathbf{Hyy} + \mathbf{Hxx} - \mathbf{Hxy} - \mathbf{Hyx} + \mathbf{Hyy} \\ &= 2\mathbf{Hxx} + 2\mathbf{Hyy} = 2Q(\mathbf{x}) + 2Q(\mathbf{y}) \quad \text{as claimed.} \end{aligned}$$

Similarly, $(Q(\mathbf{x}+\mathbf{y}) - Q(\mathbf{x}-\mathbf{y}))/4 = (\mathbf{Hxy} + \mathbf{Hyx})/2 = \mathbf{Hxy}$, so two invocations of Q suffice to compute $\mathbf{Hxy}$ for any given $\mathbf{x}$ and $\mathbf{y}$ without explicit knowledge of $\mathbf{H}$.

NOTE THAT **xy** IN **Hxy** IS NOT A VECTOR. IT IS A PAIR OF VECTORS WRITTEN TO SUGGEST AN OBJECT THAT BEHAVES LIKE A PRODUCT LINEAR IN EACH FACTOR. Strictly speaking, the bilinear form " $\mathbf{Hxz} = \mathbf{Hzx}$ " should be written " $\mathbf{H}(\mathbf{x}, \mathbf{z}) = \mathbf{H}(\mathbf{z}, \mathbf{x})$ " to express its symmetry and " $\mathbf{H}(\mathbf{z}, ß\mathbf{x} + \mu\mathbf{y}) = ß\mathbf{H}(\mathbf{z}, \mathbf{x}) + \mu\mathbf{H}(\mathbf{z}, \mathbf{y})$ " to express its linearity, but the extra commas and parentheses would clutter the page without clarifying anything.

Here is a question you were not asked: Given a "black box" that computes a real function $Q(\mathbf{x})$ but no information about how it's done, what test can be performed to check whether $Q(\mathbf{x})$ is a quadratic form? As it happens, any real continuous $Q(\mathbf{x})$ that satisfies the Parallelogram Identity must be a quadratic form, but the proof of this assertion is not obvious.