Exercises supplementing those in Richard O. Hill's
***Elementary Linear Algebra with Applications*, 3rd edition**

This is a collection of exercises that I put on homework sheets for the linear algebra part of Math 54M in Spring 2002, supplementing the exercises in Hill. (Some are adapted from exercises I had similarly given in earlier years when we used a different book.) They follow Hill's notation, and range from exercises that are comparable in difficulty to those in Hill but fill some gap or give some interesting perspective on the material, to exercises that could be ''challenge problems'' in a regular courses and good exercises for an honors courses. Some of those of moderate difficulty I assigned as part of the homework. Others of that sort, and all of the more difficult problems, I recommended (along with some exercises from the text) for ''students interested in further interesting and/or more challenging problems'' to try on their own.

I have given the exercises numbers indicating the relevant section of Hill. If an exercise is closely related to exercise n in a given section, I have generally called it ''Exercise $n+½$'' of that section; otherwise I use numbers continuing where the numbers in that section left off.

Hill has almost nothing on linear algebra over the complex numbers, which we needed for the differential equations part of the course, so I sketched the subject in class and gave some homework problems on it, numbered as exercises from ''section 5.3½''. (I preceded this with an assignment to read http//math.furman.edu/~dcs/book/c7pdf/sec71.pdf for the arithmetic of complex numbers.)

With few exceptions, I only have exercises here to sections that we covered in the course; hence many parts of the text are unrepresented.

At the end of this list of exercises, I give some corrections and additions to the material in the text.

George M. Bergman
gbergman@math.berkeley.edu
11 June, 2002

§1.2, ''Exercise 30'': Suppose we are given two constants p and q, and wish to find a parabola $y = ax^2 + bx + c$ that passes through the two points $(0, p)$ and $(1, q)$. Write down equations saying that the parabola passes through these points. Regard these as linear equations in the unknowns a, b and c, and solve them by Gaussian elimination. How many solutions will there be?

§1.2, ''Exercise 31'': Like the preceding problem, but now assuming we are given three constants p, q and r, and want to make the parabola pass through $(0, p)$, $(1, q)$ and $(2, r)$.

§1.2, ''Exercise 32'': Like the preceding two problems, but assuming we want to make the parabola pass through four points, $(0, p)$, $(1, q)$, $(2, r)$, and $(3, s)$. In this case, the system will be consistent if and only if p, q, r and s satisfy a certain equation. Find this equation.

§3.1, ''Exercise 51'': The four points

$$P = (1, 1, 1), \quad Q = (1, -1, -1), \quad R = (-1, 1, -1), \quad S = (-1, -1, 1)$$

form the vertices of a *tetrahedron* in 3-dimensional space, with the origin $\mathbf{0} = (0, 0, 0)$ as center. (If you have had some organic chemistry, you can think of these points as the locations of the four hydrogen atoms in a methane molecule CH_4, with the carbon atom at the center. If you have had some inorganic chemistry, you can think of an ammonium ion NH_4^+.)

(a) Draw a set of coordinate axes, and show the four points P, Q, R, S.

In the computations asked for in the remaining parts of this question, use the methods of §3.1.

(b) Show that the distances between every two of these points are the same, and that all four points are at the same distance from $\mathbf{0}$. (You can do this either by a long series of calculations, or by a shorter argument.)

(c) Find the angle made by any three of these four points, e.g. $\angle PQR$.

(d) Find the angle between any two of these points and the origin, e.g. $\angle P\mathbf{0}Q$. First express this as an inverse trigonometric function of an exact quantity, then find it to at least the nearest degree, using a trigonometric table or calculator. (This is the angle between the bonds of a carbon atom, important in stereochemistry.)

§3.4, ‘‘Exercise 49’’: Suppose V is a vector space.

(a) Show that if S_1, S_2 are subspaces of V, then their *intersection*, $S_1 \cap S_2$, that is, $\{\mathbf{v}\in V \mid \mathbf{v}\in S_1$ and $\mathbf{v}\in S_2\}$, is also a subspace of V.

(b) Show that if X is any set of subspaces of V, then the intersection of all members of X, i.e., $\{\mathbf{v}\in V \mid \mathbf{v}$ belongs to all $S\in X\}$, is likewise a subspace of V.

(c) Show that if $\mathbf{v}_1,\ldots,\mathbf{v}_r$ are elements of V, then the intersection of *all* the subspaces of V that contain the vectors $\mathbf{v}_1,\ldots,\mathbf{v}_r$ is precisely $\mathrm{Span}\{\mathbf{v}_1,\ldots,\mathbf{v}_r\}$.

§3.4, ‘‘Exercise 50’’: Suppose S and T are subspaces of a vector space V. Show that $\{\mathbf{u}+\mathbf{v} \mid \mathbf{u}\in S, \mathbf{v}\in T\}$ is also a subspace of V.

§3.5, ‘‘Exercise 36’’: If $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$ are nonzero mutually perpendicular vectors in $\mathbb{R}^3$ (i.e., vectors such that each is nonzero, and every two of them are perpendicular), show that the set $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is linearly independent.

§3.5, ‘‘Exercise 37’’: Suppose V is a vector space, and $\mathbf{x}_1,\ldots,\mathbf{x}_n$, $\mathbf{y}$, and $\mathbf{z}$ are elements of V. Show that if $\mathbf{z}\notin \mathrm{Span}\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ but $\mathbf{z}\in \mathrm{Span}\{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{y}\}$ then $\mathbf{y}\in \mathrm{Span}\{\mathbf{x}_1,\ldots,\mathbf{x}_n,\mathbf{z}\}$. (The symbol ‘‘$\notin$’’ means ‘‘is not a member of’’.)

Hint: Use the idea from the proof of Theorem 3.42, of ‘‘solving’’ a linear equation satisfied by a family of elements in a vector space to get an expression for one of these elements as a linear combination of the others.

§3.6, ‘‘Exercise 53’’: (a) Prove that if V is a finite-dimensional vector space and W is a subspace of V, then there exists a basis B of V which contains, as a subset, a basis of W.

(b) If V is a finite-dimensional vector space and W_1 and W_2 are two subspaces of V, must there exist a basis B of V which contains, as subsets, both a basis B_1 of W_1 and a basis B_2 of W_2?

(c) If V is a finite-dimensional vector space and W_1, W_2, W_3 are three subspaces of V, must there exist a basis B of V which contains, as subsets, a basis B_1 of W_1, a basis B_2 of W_2, and a basis B_3 of W_3?

§3.6, ‘‘Exercise 54’’: Show that if V is a vector space (not necessarily finite-dimensional) and n a positive integer, then either V contains n linearly independent elements, or V has a basis of $<n$ elements.

§3.6, ‘‘Exercise 55’’: Suppose that U and W are subspaces of a vector space W. In §3.4 ‘‘Exercise 50’’ you showed that $\{\mathbf{u}+\mathbf{v} \mid \mathbf{u}\in U, \mathbf{v}\in V\}$ is also a subspace of W. Calling the above subspace $U+V$, show that $\dim(U+V) \le \dim(U) + \dim(V)$.

§3.6, ‘‘Exercise 56’’: (a) For n = 1, 2, ..., 10, let $f_i \in C[0,11]$ be the function defined by $f_i(x) = |x-i|$. Show that the vectors $f_1,\ldots,f_{10}\in C[0,11]$ are linearly independent.

(b) Now suppose we regard $f_1,\ldots,f_{10}$ as members of $C[0,1]$. Show that in this space they are linearly dependent.

Explanation: The $f_1, \ldots, f_{10}$ of part (b) are not really the same functions as those of part (a); they are what are called the ''restrictions'' of the functions of part (a) to the interval $[0,1]$. They agree with those functions on $[0,1]$, but unlike those functions, they are undefined on $(1,11]$. We could be more precise, and denote the restrictions by modified symbols, say $\overline{f}_1, \ldots, \overline{f}_{10}$. Certain linear combinations of $f_1, \ldots, f_{10}$ which are nonzero vectors of $C[0,11]$ have restrictions to $[0,1]$ equal to zero. Thus, though $f_1, \ldots, f_{10}$ are linearly independent, the above fact makes their restrictions to $[0,1]$ linearly dependent.

§3.6, ''Exercise 57'': (a) Find a set S of 100 points in $\mathbb{R}^2$ such that every set of two distinct points of S is a basis of $\mathbb{R}^2$.

You should not write out the 100 points in a list – rather, give a formula in n such that substituting $n = 1, \ldots, 100$, gives the coordinates of your 100 points. Likewise, in proving that this set of points has the property claimed, don't give a calculation for every pair of points! Rather, give a general argument that works for all pairs in your set. The analogous instructions apply to:

(b) Find a set S of 100 points in $\mathbb{R}^3$ such that every set of three distinct points of S is a basis of $\mathbb{R}^3$.

§3.6, ''Exercise 58'': Let A be any $n \times n$ matrix.

For each positive integer i, we shall write A^d for the product $AA \ldots A$ with i factors. To also cover the case $i = 0$, let us define $A^0 = I_n$.

(a) Show that there exists a positive integer d such that $\{A^0, A^1, \ldots, A^d\}$ is linearly dependent.

(b) Show that if d is the least positive integer such that $\{A^0, A^1, \ldots, A^d\}$ is linearly dependent, then $A^d \in \mathrm{Span}\{A^0, A^1, \ldots, A^{d-1}\}$.

(c) Show that if d is as in (b) above, then for every nonnegative integer i we have $A^{d+i} \in \mathrm{Span}\{A^i, A^{i+1}, \ldots, A^{i+d-1}\}$. (Hint: multiply a relation expressing A^d in terms of $A^0, A^1, \ldots, A^{d-1}$ by A^i.)

(d) Deduce that if d is as in (b) above, then for every nonnegative integer i we have $A^i \in \mathrm{Span}\{A^0, A^1, \ldots, A^{d-1}\}$. (Suggestion: use mathematical induction.)

(e) Deduce that if d is as in (b) above, then $\mathrm{Span}\{A^0, A^1, \ldots, A^{d-1}\}$ is closed under multiplication of matrices.

(f) Does your proof of (a) give you an explicit integer d such that $\{A^0, A^1, \ldots, A^d\}$ is linearly dependent? If so, try experimenting with various matrices A of small size (say 2×2), and see whether the value of d you got in (a) really seems to be the smallest value that always makes $\{A^0, A^1, \ldots, A^d\}$ linearly dependent.

§3.6, ''Exercise 59'': Suppose V is an n-dimensional vector space, and W_1 and W_2 are subspaces of V, of dimensions m_1 and m_2 respectively. Show that if $m_1 + m_2 > n$, then W_1 and W_2 have nontrivial intersection; i.e., there are nonzero vectors $\mathbf{v}$ that lie in both W_1 and W_2. (For the concept of intersection of subspaces, cf. §3.4, ''Exercise 49''. Suggestion for this proof: let $\{\mathbf{x}_1, \ldots, \mathbf{x}_{m_1}\}$ and $\{\mathbf{y}_1, \ldots, \mathbf{y}_{m_2}\}$ be bases of these subspaces. Note why $\{\mathbf{x}_1, \ldots, \mathbf{x}_{m_1}, \mathbf{y}_1, \ldots, \mathbf{y}_{m_2}\}$ can't be linearly independent. Show that a linear relation among its members yields a nonzero element of the intersection.)

§3.6, ''Exercise 60'': For V, W_1, W_2 as in the preceding exercise, show in fact that $W_1 \cap W_2$ has dimension at least $m_1 + m_2 - n$. (Suggestion: Take a basis of $W_1 \cap W_2$, and observe that by a theorem in the text, it can be extended, on the one hand, to a basis of W_1, and, on the other hand, to a basis of W_2. Show that if the desired inequality did not hold, then on combining these bases and applying the idea of the preceding exercise, one could get a contradiction.)

§3.7, ‘‘Exercise 50’’: (a) Find a 2×2 matrix A whose null space and column space are *both* equal to the space spanned by the vector $\begin{bmatrix}1\\0\end{bmatrix}$.

(b) Prove that any $n\times n$ matrix A whose null space and column space are the same must satisfy $A^2 = \mathbf{0}$. (More generally, prove, if you can, that this is true of any matrix A whose column space is *contained in* its null space.) Verify by computation that the matrix A you gave in (a) does indeed satisfy $A^2 = \mathbf{0}$.

(c) Does there exists a 2×2 matrix A whose column space and null space are both the zero space? One whose column space and null space are both $\mathbb{R}^2$? (As always, you must justify your answers.)

§3.7, ‘‘Exercise 51’’: Let A be an $m\times n$ matrix. The next two results show that $RS(A)$ and $NS(A)$ each determine the other.

(a) Show that $NS(A) = \{\mathbf{v}\in\mathbb{R}^n \mid \text{for all } \mathbf{u}\in RS(A),\ \mathbf{u}\,\mathbf{v} = 0\}$.

(b) Show that $RS(A) = \{\mathbf{u}\in\mathbb{R}^n \mid \text{for all } \mathbf{v}\in NS(A),\ \mathbf{u}\,\mathbf{v} = 0\}$.

Note: In the above statements, we are considering the vectors written $\mathbf{u}$ to be row vectors and the vectors written $\mathbf{v}$ to be column vectors, so that the product $\mathbf{u}\,\mathbf{v}$ makes sense in each case (and is a scalar). A way of making this more explicit would be to write $\mathbf{u}\in M_{1n}$ and $\mathbf{v}\in M_{n1}$ instead of $\mathbf{u}$, $\mathbf{v}\in\mathbb{R}^n$. Alternatively, we could consider both vectors to lie in the same space, $\mathbb{R}^n$, and write $\mathbf{u}\cdot\mathbf{v}$ instead of $\mathbf{u}\,\mathbf{v}$.

§3.7, ‘‘Exercise 52’’: (a) Suppose A is any $m\times n$ matrix and B any $n\times p$ matrix. Show that $\mathrm{rk}(AB)$ is both $\leq \mathrm{rk}(A)$, and $\leq \mathrm{rk}(B)$. (Hint: relate the column space of AB to that of A, and the nullspace of AB to that of B.)

(b) Deduce that if $\mathbf{c}$ is a column vector of height m and $\mathbf{r}$ a row vector of length p, then the $m\times p$ matrix $\mathbf{c}\,\mathbf{r}$ has rank ≤ 1.

§3.7, ‘‘Exercise 53’’: Suppose A and B are $m\times n$ matrices. Show that $\mathrm{rk}(A+B) \leq \mathrm{rk}(A) + \mathrm{rk}(B)$. (Hint: Apply §3.6 ‘‘Exercise 55’’ to column spaces.)

§3.7, ‘‘Exercise 54’’: Let A be an $m\times n$ matrix, and r a positive integer.

(a) Show that A has rank $\leq r$ if and only if there exist an $m\times r$ matrix B and an $r\times n$ matrix C such that $A = BC$.

(b) Deduce that A has rank $\leq r$ if and only if it can be written as a sum of r $m\times n$ ‘‘outer product’’ matrices, where by an $m\times n$ outer product matrix we mean the product of an $m\times 1$ matrix and a $1\times n$ matrix (Hill, p.34. Cf. also §3.7 ‘‘Exercise 52(b)’’, above).

§3.7, ‘‘Exercise 55’’: (a) Show that for any matrix A, $\mathrm{rank}(A^T) = \mathrm{rank}(A)$.

(b) Obtain a formula for $\dim NS(A^T)$ in terms of $\dim NS(A)$, for an $m\times n$ matrix A.

§4.1, ‘‘Exercise 43’’: Show that for A any $m\times n$ matrix, the following conditions are equivalent:

(a) $\mathrm{rank}(A) = m$.

(b) $CS(A) = \mathbb{R}^m$.

(c) The map $f_A : \mathbb{R}^n \to \mathbb{R}^m$ is *onto*; i.e., every vector in $\mathbb{R}^m$ equals $f_A(\mathbf{x})$ for some $\mathbf{x}\in\mathbb{R}^n$.

§4.1, ‘‘Exercise 44’’: Show that for A any $m\times n$ matrix, the following conditions are equivalent:

(a) $\mathrm{rank}(A) = n$.

(b) $RS(A) = \mathbb{R}^n$.

(c) $NS(A) = \{0\}$.

(d) The map $f_A : \mathbb{R}^n \to \mathbb{R}^m$ is *one to one*; i.e., if $\mathbf{x} \neq \mathbf{y}$ in $\mathbb{R}^n$, then $f_A(\mathbf{x}) \neq f_A(\mathbf{y})$ in $\mathbb{R}^m$.

§4.1, ''Exercise 45'': If $f : D \to R$ is any function, one calls f *invertible* if there exists a function $g : R \to D$ which ''inverts'' the action of f; i.e., which has the property that for every $x \in D$, $g(f(x)) = x$ and for every $y \in R$, $f(g(y)) = y$.

(a) Show that if $f : D \to R$ is an invertible function, then f is one-to-one (in the sense of §4.1 ''Exercise 43''(d)) and onto (in the sense of §4.1 ''Exercise 44''(c)).

(b) Suppose $A \in M_{mn}$, and there exists a matrix $B \in M_{nm}$ such that $AB = I_m$ and $BA = I_n$. Show that f_A is an invertible function, and deduce using part (a) above and §4.1 ''Exercises 43 and 44'' that $m = n$.

The above result shows why the concept of the inverse of a matrix was only defined for square matrices!

(c) On the other hand, give an example of non-square matrices $A \in M_{mn}$, $B \in M_{nm}$ such that $AB = I_m$ (without the other equality). In this situation, one calls A a *left inverse* of B, and B a *right inverse* of A.

(d) Prove that for any pair of matrices A, B satisfying $AB = I_m$ as in part (c), f_A must be onto and f_B must be one-to-one.

(e) What can you deduce about the relation between m and n in the above situation?

§4.1, ''Exercise 46'': Suppose V is a vector space with a basis $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$. Define a function $f : V \to \mathbb{R}^n$ by letting $f(\mathbf{x}) = [\mathbf{x}]_S$ for all $\mathbf{x} \in V$. Show that f is a linear transformation.

§4.1, ''Exercise 47'': Suppose V and W are two vector spaces, having bases $S = \{\mathbf{x}_1, \ldots, \mathbf{x}_m\}$ and $T = \{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$ respectively; and suppose $L : V \to W$ is a linear transformation. Show that there exists an $n \times m$ matrix A such that for all $\mathbf{v} \in V$,

$$[L(\mathbf{w})]_T = A\,[\mathbf{w}]_S .$$

(Hill proves this, in slightly different notation, in a section we do not cover in the course. But it is not hard to prove from the material in §4.1.)

§4.1, ''Exercise 48'': Suppose V is an n-dimensional vector space and $f : V \to W$ a linear transformation from V to another vector space W, and suppose f is one-to-one and onto. Show that W is also n-dimensional.

§4.2, ''Exercise 48'': Suppose $\{\mathbf{u}, \mathbf{v}\}$ is a basis of a 2-dimensional vector-space V, and we want to find an inner product on V under which the triangle with vertices 0, $\mathbf{u}$, $\mathbf{v}$ is equilateral, with all sides having length 1 with respect to this inner product.

(a) Write out the three equations saying that all sides of the above triangle have length 1, in terms of the inner product, and solve these to get the values of $\mathbf{u} \cdot \mathbf{u}$, $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{v}$.

(b) Write out a formula for the product of two general vectors in V, $(a\mathbf{u} + b\mathbf{v}) \cdot (c\mathbf{u} + d\mathbf{v})$, under the inner product found in part (a).

(c) Applying this to the case where $V = \mathbb{R}^2$ and $\mathbf{u} = (1, 0)$, $\mathbf{v} = (0, 1)$, write out the formula for $(x, y) \cdot (x', y')$ under the inner product that makes the triangle with vertices $(0,0)$, $(1,0)$, $(0,1)$ equilateral with all sides of length 1.

(Note: I have not asked you to prove that the function you have found really is an inner product. In fact, the first three conditions for an inner product are easy to verify, but positivity is not quite so easy. However, if you did this problem correctly, the function you found satisfies this condition as well.)

§4.2, ''Exercise 49'': Let V be a vector space with a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$.

If ''$\cdot$'' is an inner product on V, then to specify ''$\cdot$'' all we need to give is the values of the products $\mathbf{x}_i \cdot \mathbf{x}_j$ for $i, j \in \{1, \ldots, n\}$. For once we know these values, the dot product of any two vectors, $(a_1\mathbf{x}_1 + \ldots + a_n\mathbf{x}_n) \cdot (b_1\mathbf{x}_1 + \ldots + b_n\mathbf{x}_n)$, can be found by ''multiplying out'', and then substituting in the given values for the dot products of the basis vectors.

This suggests the question: If we want to define an inner product on V by choosing real numbers p_{ij} to be the values of the products $\mathbf{x}_i \cdot \mathbf{x}_j$ and determining ''$\cdot$'' on all elements by the above method, what properties do the p_{ij} have to satisfy to make this work? This exercise will give a partial answer to this question.

(a) Suppose $n = 2$, and we are given four real numbers $p_{11}, p_{12}, p_{21}, p_{22}$. Show that the ''product'' on V defined by

$$(a_1\mathbf{x}_1 + a_2\mathbf{x}_2) \cdot (b_1\mathbf{x}_1 + b_2\mathbf{x}_2) \ = \ a_1 b_1 p_{11} + a_1 b_2 p_{12} + a_2 b_1 p_{21} + a_2 b_2 p_{22}$$

is an inner product if and only if the following four conditions hold:

$$p_{11} > 0, \qquad p_{22} > 0, \qquad p_{12} = p_{21}, \qquad p_{12}^2 \ < \ p_{11}\, p_{22}\,.$$

(Hint: The last condition is related to the Cauchy-Schwartz inequality.)

(b) Now suppose, instead, that $n = 3$, and that we choose nine real numbers $p_{11}, \ldots, p_{33}$, and define a ''product'' on V by the formula analogous to the above, namely $(\Sigma\, a_i \mathbf{x}_i) \cdot (\Sigma\, b_i \mathbf{x}_i) = \Sigma_{i,j}\ a_i\, b_j\, p_{ij}$. Write down conditions analogous to those given in (a), and show that they must hold if the operation ''$\cdot$'' is an inner product on V. (There will be three ''positivity'' conditions, three ''symmetry'' conditions, and three ''Cauchy-Schwartz'' conditions.)

(c) Show, however, that there exist real numbers p_{ij} which satisfy the conditions found in part (b), but such that ''$\cdot$'' is not an inner product.

§4.2, ''Exercise 50'': Show that if $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are vectors in an inner product space, then $\|\mathbf{u}+\mathbf{v}+\mathbf{w}\| \ \leq \ \|\mathbf{u}\| + \|\mathbf{v}\| + \|\mathbf{w}\|$.

§4.2, ''Exercise 51'': Hill's exercises show that the same vector space can have more than one inner product. Suppose $\cdot_1$ and $\cdot_2$ are two inner products on the same nonzero vector space V, and r a real number. Let us define some further functions $\cdot_3$, $\cdot_4$, $\cdot_5$, $\cdot_6$ by the formulas

$$\mathbf{x} \cdot_3 \mathbf{y} = \mathbf{x} \cdot_1 \mathbf{y} + \mathbf{x} \cdot_2 \mathbf{y}, \qquad \mathbf{x} \cdot_4 \mathbf{y} = \mathbf{x} \cdot_1 \mathbf{y} - \mathbf{x} \cdot_2 \mathbf{y}, \qquad \mathbf{x} \cdot_5 \mathbf{y} = r\, \mathbf{x} \cdot_1 \mathbf{y}, \qquad \mathbf{x} \cdot_6 \mathbf{y} = \mathbf{x} \cdot_1 \mathbf{y} + 1, \qquad \mathbf{x} \cdot_7 \mathbf{y} = (\mathbf{x} \cdot_1 \mathbf{y})^2.$$

(a) Show that $\cdot_3$ is an inner product.

(b) Show that $\cdot_4$ is an inner product if and only if for all nonzero $\mathbf{x} \in V$, $\mathbf{x} \cdot_1 \mathbf{x} \ > \ \mathbf{x} \cdot_2 \mathbf{x}$.

(c) Show that $\cdot_5$ is an inner product if and only if $r > 0$.

(d) Show that $\cdot_6$ and $\cdot_7$ are never inner products.

§4.2, ''Exercise 52'': Suppose $T : V \to W$ is a linear transformation of vector spaces, and $\cdot_1$ is an inner product on W. For $\mathbf{x}, \mathbf{y} \in V$, let us define $\mathbf{x} \cdot_2 \mathbf{y} = T(\mathbf{x}) \cdot_1 T(\mathbf{y})$.

(a) Show that if T is one-to-one, then $\cdot_2$ is an inner product on V.

(b) Show that if T is not one-to-one, then $\cdot_2$ is not an inner product.

§4.2, ''Exercise 53'': (a) Suppose V is an inner product space, and $\mathbf{t}$ an element of V. Show that the map $L_{\mathbf{t}} : V \to \mathbb{R}$ defined by $L_{\mathbf{t}}(\mathbf{x}) = \mathbf{t} \cdot \mathbf{x}$ is a linear transformation.

(b) Suppose n is a positive integer and T is a linear transformation $\mathbb{R}^n \to \mathbb{R}$. Show that there exists an element $\mathbf{t} \in \mathbb{R}^n$ such that for all $\mathbf{x} \in R^n$, $T(\mathbf{x}) = \mathbf{t} \cdot \mathbf{x}$. In the notation of part (a), what you will have proved is that every linear transformation $\mathbb{R}^n \to \mathbb{R}$ is of the form $L_{\mathbf{t}}$ for some $\mathbf{t} \in V$.

§4.2, ‘‘Exercise 54’’: This exercise is related to the preceding, but is more difficult.

(a) Suppose V is a finite-dimensional inner product space. Show that (in the notation of part (a) of the preceding exercise) every linear transformation $V \to \mathbb{R}$ is of the form $L_{\mathbf{t}}$ for some $\mathbf{t} \in V$.

(b) Show that the result of part (a) is *not* true for infinite-dimensional inner product spaces.

(Suggestion: Consider either of the following cases: (i) $V = C[0,1]$, with the inner product defined in Hill, and L the function given by $L(f) = f(0)$. (ii) V = the vector space of all polynomials $p(x)$, made an inner product space using the same inner product as for $C[0,1]$, and L the function given by $L(p(x)) = e^x \cdot p(x)$, evaluated in $C[0,1]$.)

§4.4, ‘‘Exercise 33’’: In §3.1 ‘‘Exercise 51’’, you considered the vertices of a tetrahedron:

$$P = (1,\ 1,\ 1),\quad Q = (1,\ -1,\ -1),\quad R = (-1,\ 1,\ -1),\quad S = (-1,\ -1,\ 1),$$

and found that the first three of these formed a basis of $\mathbb{R}^3$.

(a) Apply the Gram-Schmidt process to these vectors P, Q and R to get an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ of $\mathbb{R}^3$.

(b) Find the coordinate matrices of P, Q, R and S (p.217) in terms of that orthonormal basis.

(c) Find the coordinate matrices for the elements of the standard basis of $\mathbb{R}^3$ in terms of that orthonormal basis.

§4.4, ‘‘Exercise 34’’: In $\mathbb{R}^{1,000}$ consider the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{999}$, where $\mathbf{x}_i$ has 1 in the ith and $i+1$st places, and 0 in all other places. (I.e., $\mathbf{x}_i = \boldsymbol{\varepsilon}_i + \boldsymbol{\varepsilon}_{i+1}$.)

(a) Show that these vectors are linearly independent.

(b) Find an orthonormal basis for the 999-dimensional subspace that they span. (Suggestion: Apply the Gram-Schmidt process to get four or five terms, and see whether a pattern is evident. Describe the ith term of the series that this pattern would give, and prove that the vectors you have described will indeed be orthonormal and span the same subspace as $\mathbf{x}_1, \ldots, \mathbf{x}_{999}$.)

§5.1, ‘‘Exercise 40’’: Suppose A is an $m \times n$ matrix and B is an $n \times m$ matrix, where $m > n$. Let A' be the $m \times m$ matrix that one gets by attaching $m-n$ columns of zeroes to the right side of A, and let B' be the $m \times m$ matrix that one gets by attaching $m-n$ rows of zeroes to the bottom of B.

(a) Show that $A'B' = AB$.

(b) Deduce from this that $\det(AB) = 0$. (Hint: What can you say about $\det(A')$ and $\det(B')$?)

(c) In the above situation, must $\det(BA) = 0$? (Suggestion: experiment.)

(d) Find another proof of the result of (b), using Lemma 3.47 and Theorem 5.1. (Hint: B has more columns than rows.)

(e) In the case $m = 2$, $n = 1$, re-prove the result of (b) by direct computation. I.e., write out matrices of the appropriate sizes, with letters for their entries, multiply them together, and compute the determinant of the product.

§5.1, ‘‘Exercise 41’’: Let

$$A = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ 0 & 0 & i & j \\ 0 & 0 & k & l \end{bmatrix}.$$

Find $\det(A)$ by cofactor expansion with respect to one column. Factor your answer if possible.

§5.2, ‘‘Exercise 38’’: Suppose A and B are $m\times n$ matrices. Call a real number λ a ‘‘relative eigenvalue’’ of A and B if there exists a nonzero vector $\mathbf{x}\in\mathbb{R}^n$ such that $A\mathbf{x} = \lambda B\mathbf{x}$.

(a) In the case $m = n$, show that *either* every real number is a relative eigenvalue of A and B *or* that the number of relative eigenvalues of A and B is $\leq n$. Give examples showing that both cases occur.

(b) In the case $m < n$, show that every real number is a relative eigenvalue of A and B.

(For $m > n$, the same result stated for $m = n$ holds, but it takes a bit more ingenuity to prove.)

§5.3, ‘‘Exercise 38’’: Let A be a square matrix, and x an eigenvector of the matrix A^2 whose eigenvalue λ is ≥ 0. Show that x is either an eigenvector of A with eigenvalue $\sqrt{\lambda}$, or an eigenvector of A with eigenvalue $-\sqrt{\lambda}$, or the sum of two eigenvectors, one with each of these eigenvalues. (Hint: Obtain the desired eigenvectors as certain linear combinations of $\mathbf{x}$ and $A\mathbf{x}$.)

‘‘§5.3½, Exercises 1-3’’: Solve the following systems of linear equations over the complex numbers (i.e., give formulas describing all n-tuples of complex numbers $(z_1, \ldots, z_n)$ etc. which satisfy the equations. If there are more than one solutions, these formulas should involve independent variables s, t etc. which take as values arbitrary complex numbers.)

1. $\begin{aligned}(1+i)\,z_1 + 2\,z_2 &= 3 \\ 2\,z_1 - 3\,z_2 &= 1\end{aligned}$

2. $\begin{aligned}z_1 - i\,z_2 - 2\,z_3 &= -3 - 2\,i \\ z_1 + i\,z_2 &= 1 + 2\,i\end{aligned}$

3. $\begin{aligned}(1+i)\,w + (2+i)\,z &= 1 \\ (3+i)\,w + 5\,z &= 2+i.\end{aligned}$

‘‘§5.3½, Exercise 4’’: Find complex numbers a and b such that $a\begin{bmatrix}3+i \\ 2\end{bmatrix} + b\begin{bmatrix}4 \\ 3-i\end{bmatrix} = \begin{bmatrix}1 \\ 0\end{bmatrix}$.

‘‘§5.3½, Exercise 5’’: (a) Find the eigenvalues in $\mathbb{C}$ of the matrix $A = \begin{bmatrix}1 & -1 \\ 1 & 1\end{bmatrix}$.

(b) Find an eigenvector of A in $\mathbb{C}^2$ for each eigenvalue.

(c) Find a matrix P with entries in $\mathbb{C}$ such that $\Lambda = P^{-1}AP$ is diagonal, and compute Λ.

‘‘§5.3½, Exercise 6’’: (a) $\begin{bmatrix}1 \\ 1 \\ 1\end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix}1 & 1 & 2 \\ 2 & 1 & 1 \\ 1 & 2 & 1\end{bmatrix}$; find its eigenvalue.

(b) Find all the eigenvalues of the above matrix A, and an eigenvector corresponding to each eigenvalue.

‘‘§5.3½, Exercise 7’’: (a) $\begin{bmatrix}2 \\ -3+i \\ -2 \\ 3-i\end{bmatrix}$ is an eigenvector of $A = \begin{bmatrix}4 & 3 & -2 & -1 \\ -6 & -2 & 4 & 4 \\ -2 & -1 & 4 & 3 \\ 4 & 4 & -6 & -2\end{bmatrix}$; find its eigenvalue.

(b) Find all the eigenvalues of the above matrix A, and an eigenvector corresponding to each eigenvalue.

‘‘§5.3½, Exercise 8’’: Suppose z is a complex number that is not real. Show that there is one and only one complex number w such that the coefficients of the quadratic polynomial $(t - z)(t - w)$ are all real; namely $w = \bar{z}$.

§5.4, ‘‘Exercise 26’’: Suppose A is a square matrix of real numbers, $\mathbf{x} \in \mathbb{R}^n$ is a real eigenvector of A with (real) eigenvalue λ, and $\mathbf{y} \in \mathbb{R}^n$ is a real eigenvector of A^T with (real) eigenvalue $\mu \neq \lambda$. Show that $\mathbf{x}$ and $\mathbf{y}$ are orthogonal. (Hint: The idea is the same as that of the proof of Theorem 5.43.)

§5.4, ‘‘Exercise 27’’: (a) Suppose Q is an orthogonal matrix of real numbers. Show that every eigenvalue λ of Q in $\mathbb{C}$ satisfies $|\lambda| = 1$.

(Suggestion: The method of proof of Theorem 5.39(a) combines the general fact $A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A^T\mathbf{y}$ for A a real matrix and $\mathbf{x}$, $\mathbf{y}$ complex vectors with the fact that a symmetric matrix satisfies $A = A^T$. The proof of this exercise should similarly combine the general fact noted with the property relating an orthogonal matrix with its transpose.)

(b) Find the eigenvalues of the orthogonal matrix of Example 2, p.288 (where θ is an arbitrary angle), and verify computationally that they satisfy the condition of part (a).

§5.4, ‘‘Exercise 28’’: A square matrix A of real numbers is called *skew symmetric* if $A^T = -A$. (Such matrices are mentioned by Hill in §1.6, Exercises 19-21, but this exercise does not depend on those.) The first part of this exercise shows you what such matrices ‘‘look like’’:

(a) Write down a skew symmetric 3×3 matrix A which satisfies $a_{12} = 1$, $a_{13} = 2$, $a_{32} = 3$. (There is one and only one such matrix.)

(b) Use the idea of §5.4 ‘‘Exercise 27’’ to find a condition which is satisfied by the eigenvalues of any skew symmetric matrix.

(c) Find the eigenvalues of the matrix of part (a), and verify that they satisfy the condition found in part (b).

§5.4, ‘‘Exercise 29’’: The results on symmetric matrices in §5.4 of Hill, on orthogonal matrices on pp.286-288 and in ‘‘Exercise 27’’ above, and on skew symmetric matrices in ‘‘Exercise 28’’ above all concerned matrices with *real* entries, and were based on the formula $A\mathbf{x} \cdot \mathbf{y} = \mathbf{x} \cdot A^T\mathbf{y}$, which is valid for such matrices A, even if the vectors $\mathbf{x}$ and $\mathbf{y}$ have entries in $\mathbb{C}$.

But that formula is *not* valid for a general matrix A with entries in $\mathbb{C}$. The correct formula is then

$$A\mathbf{x} \cdot \mathbf{y} \;=\; \mathbf{x} \cdot \bar{A}^T\mathbf{y},$$

where $\bar{A}$ denotes the matrix obtained from A by replacing each entry by its complex conjugate.

(a) Prove the above formula.

(b) Show that if A is a matrix with complex entries which satisfies $A = \bar{A}^T$, then all its eigenvalues are real, and any two eigenvectors with distinct eigenvalues are orthogonal.

(c) Show that if A is a matrix with complex entries which satisfies $A\bar{A}^T = I$, then A preserves inner products (as defined in §5.5, Exercise 24), all its eigenvalues satisfy $|\lambda| = 1$, and eigenvectors with distinct eigenvalues are orthogonal.

(d) Show that if A is a matrix with complex entries which satisfies $A = -\bar{A}^T$, then its eigenvalues have the same property proved for real skew symmetric matrices in §5.4 ‘‘Exercise 28’’, and any two eigenvectors with distinct eigenvalues are orthogonal.

Note: If A is a matrix over the complex numbers, the matrix $\bar{A}^T$ is commonly written A^*. The three kinds of matrices described in parts (b), (c) and (d) above are called *Hermitian* matrices, *unitary* matrices, and *skew Hermitian* matrices, respectively.

(e) For which complex numbers z will the 1×1 matrix (z) be Hermitian? unitary? skew Hermitian?

§5.4, ‘‘Exercise 30’’: Show that a matrix A with real entries is skew symmetric (as defined in ‘‘Exercise 27’’ above) if and only if for every real vector $\mathbf{x}$, the matrix $A\mathbf{x}$ is orthogonal to $\mathbf{x}$.

(If you have done §5.4 ‘‘Exercise 29’’ above, you can prove the same thing for matrices and vectors with complex entries, with ‘‘skew Hermitian’’ in place of ‘‘skew symmetric’’.)

Suggestion: It is straightforward to prove by the kinds of calculation used so far that if A is skew symmetric, it has the above orthogonality property. To show the converse, let $\mathbf{x}$ and $\mathbf{y}$ be any two vectors, write down the assumed orthogonality conditions for $\mathbf{x}$, for $\mathbf{y}$, and for $\mathbf{x}+\mathbf{y}$, obtain from the resulting three equations a relation between $\mathbf{x}\cdot A\mathbf{y}$ and $\mathbf{y}\cdot A\mathbf{x}$, and deduce skew symmetry from that relation.

Remark: If a rigid body is rotating, then at every point p of the body, the radius vector from the center of rotation to p is perpendicular to the velocity vector at p. Such rotation may in fact be described by a skew symmetric matrix, which, applied to each radius vector, gives the velocity vector at the endpoint.

Some corrections and clarifications to the text

P.1, middle line paragraph before §1.1: The first word of the Chinese title referred to should be ‘‘Chiu’’. (Alternatively, in the system of transliteration currently used, the whole title should be *Jiǔ Zhāng Suàn Shū*.)

P.28: Between (1.21) and (1.22) there should be another law, ‘‘(1.21½) $A(rB) = (rA)B = r(AB)$ for any real number r’’. (This is used, for instance, on p.170, third-from-last line.)

P.45, last line: In the first matrix, the first entry in the second row should not be 3 but 0. (With ‘‘3’’ there, it is not diagonal.)

P.46, Theorem 1.47(a): The second sentence should read, ‘‘The ith row of the product DA is obtained by multiplying the ith row of A by d_i; the ith column of the product BD is obtained by multiplying the ith column of B by d_i.’’

P.47, first sentence: To make what is meant more explicit, one should start a new sentence where the author has ‘‘and we say’’, changing this final statement to ‘‘One says that A is **row equivalent** to B if B can be written $e_k \ldots e_2 e_1 A$ for some elementary row operations $e_1, \ldots, e_k$. From the formula just given in parenthesis, we see that if A is row equivalent to B, then B is also row equivalent to A.’’

P.49, bottom of page, last matrix: In the next-to-last column, the second subscript of each entry is $n-1$, and the first subscript is whatever comes before that; e.g., ‘‘1’’ in the top entry.

P.108, Theorem 2.27: A more precise statement would be that if det is any determinant function, then $\det(A)$ it is given by the sum of all signed elementary products from A, and conversely that if $\det(A)$ is defined to be the sum of all signed elementary products from A, then det is a determinant function.

P.132, second line above Figure 3.3: The word ‘‘identify’’ should be explained to the class. (If in doubt, ask your students what they think the assertion beginning ‘‘we often identify the vector ...’’ means.)

P.149, instructions for Exercises 37-42: ‘‘perpendicular to $\mathbf{u}$’’ should be ‘‘perpendicular to $\mathbf{v}$’’.

P.150, last sentence of longest paragraph, ‘‘In this way we can ...’’: This takes for granted that any 2-dimensional subspace of $\mathbb{R}^n$ ‘‘looks just like’’ (i.e., is isometric to) Euclidean plane geometry. (This could be proved using the sketch of the Gram-Schmidt process at the end of the preceding section, though it would be easier after one had the concept of orthonormal basis, in Chapter 4. In either case, it would take some work to state and to prove.) Hill uses the above assertion in Theorem 3.17 on p.152. He also uses it on p.268, top line, where he takes for granted that the Pythagorean theorem holds in $\mathbb{R}^n$; however, this can be proved very quickly.

P.158, third line of Example 3: ''is equal to all'' should be ''is equal to the set of all''.

P.166, Example 1(b) should read ''T = the set composed of the first and third quadrants''. (For those students who have seen some elementary set theory, one can give the more precise statement that T is the union of the first and third quadrants.) Likewise, on p.168, Example 3(a), ''is the x- and y-axis'' should be ''is the set consisting of [or: is the union of] the x- and y-axes''.

P.167, Definition 3.26, end of third line: ''so that'' should be ''and if'', and the sentence should end ''by S''.

P.169 or shortly thereafter: There should be a discussion of the relation between the solutions to the matrix equations $A\mathbf{x} = \mathbf{0}$ and $A\mathbf{x} = \mathbf{b}$, namely that the solution-set of the latter equation, if nonempty, is the translate of the solution-set of the former, i.e., $NS(A)$, by any particular solution to the latter equation.

P.172, display (3.32): ''that combination may not be unique'' should be ''the coefficients in that combination may not be unique''.

P.192, top line: This should be ''Since the $\mathbf{v}_i$'s are linearly independent, the above equations will hold if and only if these coefficients are all zero''.

P.202, first line of text: The entry of the matrix shown as ''4'' should be 6 (as on preceding page).

P.215, Exercises 41 and 42: What one is being asked to prove is a part of Lemma 3.84, which is used in proving Theorem 3.68; hence it should be made clear that students may not call on that lemma or theorem in their proofs.

P.242, Definition 4.1: Change the first sentence to: ''A **function** f consists of two sets, D and R, and a rule f which assigns to each element x of D a unique element $f(x)$ of R.''

P.268, top, just before end-of-proof sign: Add ''Conversely, suppose $\boldsymbol{\varepsilon} = \mathbf{y} - A\mathbf{x}$ is not perpendicular to $CS(A)$. Thus there exists some $A\mathbf{u} \in CS(A)$ having nonzero dot product with $\boldsymbol{\varepsilon}$. Now consider all vectors $\mathbf{x}(t) = \mathbf{x} + t\mathbf{u}$ $(t \in \mathbb{R})$, and the corresponding error vectors $\boldsymbol{\varepsilon}(t) = \mathbf{y} - A\mathbf{x}(t) = \boldsymbol{\varepsilon} - tA\mathbf{u}$. We see that $\|\boldsymbol{\varepsilon}(t)\|^2 = (\boldsymbol{\varepsilon} - tA\mathbf{u}) \cdot (\boldsymbol{\varepsilon} - tA\mathbf{u}) = \|\boldsymbol{\varepsilon}\|^2 - 2t(\boldsymbol{\varepsilon} \cdot A\mathbf{u}) + t^2\|A\mathbf{u}\|^2$, a polynomial in t whose derivative at $t = 0$ is $2(\boldsymbol{\varepsilon} \cdot \mathrm{A}\mathbf{u}) \neq 0$; hence its value is not minimum when $\boldsymbol{\varepsilon}(t) = \boldsymbol{\varepsilon}(0) = \boldsymbol{\varepsilon}$.''

P.276, Exercises 29-30: Ignore these. They make the incorrect assumption that the solution to a least squares problem is invariant under taking the logarithms of the equations.

P.338, statement of Theorem 5.17: After ''with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$'' add ''where any λ's that are *multiple roots* of the characteristic polynomial of A are repeated in this list exactly as many times as they are repeated as roots.''

P.344, second line of Theorem 5.27: ''where Λ is a diagonal matrix'' should be ''where S is an $n \times m$ matrix each of whose columns is nonzero, and Λ is an $m \times m$ diagonal matrix''.

P.346, top line: ''= 2.4'' should be ''= 2''.

P.352, statement of Theorem 5.39(a): The equivalence of these two sentences relies on the Fundamental Theorem of Algebra, which the class should be reminded of before reading this.

P.352, fourth and fifth lines of proof of Theorem 5.39(a): The equation $A\mathbf{v} \cdot \mathbf{w} = \mathbf{v} \cdot A^T\mathbf{w}$, which as the author notes was proved in Theorem 4.15 for the real inner product, is not valid in general for the complex inner product; the result in that case is $A\mathbf{v} \cdot \mathbf{w} = \mathbf{v} \cdot \bar{A}^T\mathbf{w}$. However, because the matrix A considered here has real entries, this equation reduces to the one for real inner products in this case.