A course on . . .

APPLIED LINEAR ALGEBRA

TSE-WEI WANG
The University of Tennessee
Knoxville, TN 37996-2200

IN BOTH INDUSTRY and academics, as the emphasis on multivariable control designs develop, it becomes indispensable that the concept of linear algebra and its geometric and physical interpretations be mastered as background knowledge. As graduate courses introducing recent developments in the theory and design of multivariable process controls emerge in the graduate curriculum, a concomitant background course in applied linear algebra becomes imperative in understanding the new complexity of multivariable control. Three years ago, the chemical engineering department introduced a new course, cross-listed in both the electrical and computer engineering and the mechanical engineering departments, entitled "Application of Numeric Linear Algebra in Systems and Control Engineering." All chemical engineering graduate students in the system modeling and process control areas and all electrical engineering students taking the graduate linear systems theory course are required to take this course. A prerequisite is senior or graduate standing with a prior introductory undergraduate course to vectors and matrices.

The students usually come into the course knowing only how to do matrix addition, subtraction, and multiplication—finding the determinant and inverse of up to 3x3 matrices. Some of them know a little about basis vectors and have some notions about linear independence of vectors. In all three departments, the students can use this course to satisfy one of their math course requirements. All other graduate students are strongly encouraged to take this course.

The goal of the course is to introduce engineering students, especially those majoring in the systems and control area, to the concepts and the physical as well as the geometric interpretations of some key linear algebra topics and their associated numerical consideration. Examples from system modeling and control areas are used extensively in order to lend a sense of reality to the rather abstract mathematical concepts. In this article, we describe the course teaching philosophy, the computer projects assignments, and the student feedback. We have received such favorable comments and support from the faculty and students that we plan to offer it annually in the fall semester. It will also serve as a corequisite for the 500-level course on linear systems theory offered by the electrical and computer engineering department.

The goal of the course is to introduce engineering students, especially those majoring in the systems and control area, to the concepts and the physical as well as the geometric interpretations of some key linear algebra topics and their associated numerical considerations.

Tse-Wei Wang is an assistant professor of chemical engineering at the University of Tennessee. She received a PhD in biophysics from M.I.T. in 1977, concentrating in the study of human platelet physiology. She obtained a MS is chemical engineering from the University of Tennessee in 1986 and joined the faculty there soon afterwards. Her areas of interest are biotechnology and process control of chemical and biochemical processes.

In a previous article [1] published in the fall, 1984, issue of *Chemical Engineering Education*, entitled "Linear Algebra for Chemical Engineers," K. Zygourakis (Rice University) describes the linear algebra course as the first semester of a two-semester sequence applied math course. Our course at the University of Tennessee differs from that in that we emphasize the *geometric* and *physical* interpretations of the various theorems and decompositions in order to develop, in the students, the ability to answer for themselves questions such as, how do I go about computing the controllable or observable subspace of a dynamic system; how do I use the concept of rank and linear independence to analyze a set of input and output data of a given process; how do I use the concept of orthogonality in analyzing a system matrix; how can I tell if a particular algorithm for system analysis is prone to numerical instability; what is the role of positive-definite matrices in an optimization problem; what does it mean for two physical system matrices to be connected by a similarity transformation; what is the danger of a pole-zero cancellation of a transfer function?

We hope that students will be able to start developing an intuitive understanding of the relationship and interactions among the several system variables by analyzing the matrices that connect between them, thereby guiding them in choosing the most appropriate design and analysis methods. We do not emphasize the writing of computer codes to implement the various numerical algorithms because we recognize that reliable numerical software exists (such as MATLAB [2, 3] that is mainly based on stable routines contained in such packages as LINPACK [4] and EISPACK [5] for various computer models). Rather, we use the software to solve some physical problems or to implement a certain algorithm in order to study the numerical aspects of it. We are trying to impart to the students the intuitive ability to examine a system and by using fundamental linear algebra concepts, to extract physical information from it. For instance, in considering the placement of temperature sensors along a distillation column, how does one decide where to place them in order to extract the most useful information about the behavior of the column from their measurements? Or, in mechanical engineering, where should the accelerometers be placed along a beam in order to detect the first N modes of vibration due to a set of inputs? It can be shown [6] that the choice is the sites where the gain matrix between the control inputs, *e.g.*, reflux ratio, and bottom heat duty, and the system outputs, *e.g.*, the temperature measurements, that yields the smallest condition number and that has the largest sensitivity in the gains, or a compromise of the two, because this arrangement implies a more balanced distribution of energy involved in each of the control input variables.

As the description of a system changes from single variable to multivariable, very often the single-variable concepts, such as size and interaction, do not carry straight forward into the multivariable case. In the latter, the concept of directionality as exhibited by the various variables and their interactions with each other necessitates the use of a set of coordinate systems to describe the dynamics. The motto "happiness is finding things are linear" extends into the realm of linear algebra in that "happiness is finding that coordinates are perpendicular"; therefore, the various decompositions (such as QR, SVD, Householder) emerge so that a system can be transformed into a new representation with mutually orthogonal basis vectors.

True, all these theories and algorithms involved are normally covered in upper-level mathematical courses offered by a math department. One asks, legitimately, why is the engineering department

As the description of a system changes from single variable to multivariable, very often the single-variable concepts, such as size and interaction, do not carry straight forward into the multivariable case.

bothering to cover the same materials? Why not just send the graduate students over to the math department? The answer is that unless the math department maintains a constant liaison with the various engineering departments in order to monitor their need in higher level mathematics, the courses they offer will usually not serve the needs of the engineering students who want to use the mathematics as tools in solving practical problems.

Take linear algebra as an example. At the University of Tennessee, three undergraduate courses (semester) exist in the theory and numerical aspects of linear algebra; at least four graduate courses exist that deal with the theory and algorithms of various topics of linear algebra, such as solving the least square problem and the various decompositions. Engineering students who take them come away knowing how to perform a certain decomposition or how to calculate the eigenvalues and eigenvectors, and have learned the numerical aspects of the various algorithms. But they have not acquired the intuition relating knowledge of the mathematics to selection of the methods for analysis, design, and control of physical systems. Study of the properties of linear vector spaces should be linked to the notion that the state space of a dynamical system constitutes a linear vector space and that the controllable and/or observable space constitutes a subspace of the original state space. Then all the manipulations, such as change of basis, orthogonalization, QR, and SVD, can be viewed as a way to view the system states in a more facilitating coordinate frame (orthogonal), and the system matrices or transformation matrices can be viewed with respect to these new coordinate frames. As a result, the properties associated with these special matrices, such as unitarity, orthonormality, and triangularity, can be used to view the transformation as represented by these matrices in a more intuitive and simplified manner. An area where a variety of physical problems can be used to illustrate the math principles is that of using SVD and pseudoinverse in solving least-square problems. In the long run, we hope that the experience gained in teaching both the engineering and mathematical version of the materials can lead to a single course meeting the goals of both groups.

The textbook used is *Linear Algebra and Its Applications* [7], by Gilbert Strang. Table 1 lists the

TABLE 1
Course Materials

Course Textbook

Strang, G., *Linear Algebra and Its Applications,* 3rd ed., Harcourt Brace Jovanovich, Inc., (1988)

Additional Course References

1. Stewart, G.W., *Introduction to Matrix Computations,* Academic Press (1973)
2. Golub, G., and C. Van Loan, *Matrix Computations,* Johns Hopkins Press (1983)

textbook along with the supplemental reference books, and Table 2 shows the topics covered in the course. From time to time, details of some topics are also presented from references listed [8] and [9]. Strang presents the materials as a systematic development of observations on a set of linear algebraic equations (later on, on a set of linear ordinary differential equations). His presentation elicits enthusiasm from the readers until the mystery of observations is solved, seemingly intuitively. Then, *voilà,* he formally states the deductions in theorems. He leads one from the beginning to the end of the development of a concept in such a manner that one cannot help following him in order to see the interpretation of the observations! Most students in the class also appreciate Strang's style of presentation.

Over half of the class time is devoted to the first three chapters, involving analysis of solving the problem of $Ax = b$, the over- and under-determined, and the inconsistent cases. After the mechanism of Gaussian eliminations with pivoting is presented, the concept of the four fundamental subspaces is introduced. Geometric visualization of the orthogonal complementary subspaces, *e.g.*, the row and null spaces, is stressed. The roles of the four subspaces with respect to linear transformations are, in turn, explained and visualized in detail. The decomposition of any vector into its orthogonal components is emphasized. In geometric visualization, a three-dimensional space is always used because of its familiarity. Then, the visualization of the vectors b and x, as in $Ax = b$, in the recipient and domain space, respectively, of the linear transformation represented by the matrix A, is made. Figures 1 and 2 (from Strang) are used very often to depict the actions of A and the Moore-Penrose pseudo-inverse, A^+, with respect to the four subspaces. The role of each of the four fundamental subspaces with respect to the under- or over-determined and inconsistent cases is analyzed in detail. At this point, an example is given concerning the underdetermined case. The problem is presented of a physical process with more inputs than outputs, and they are related at steady state, by A, as in $y = Au$. The dimension of A is therefore rectangular, mxn, with $m < n$, more variables than constraints. Therefore, from linear algebra theory, many solutions exist. One can pose an optimization problem where one wants to find the solution, x_{op}, from the set of all possible solutions, such that some function of x_{op} is minimized (or maximized).

A physical example where an inconsistent case of $Ax = b$ may arise is offered at this point. Cases involving multiple measurement data points are the most common. A specific example, mentioned earlier, is one of temperature sensor measurements along the many trays of a distillation column. Usually, two control inputs are considered. Yet there may be five or more

TABLE 2
Topical Outline, Applied Linear Algebra Course

Matrices and Gaussian Elimination
- Introduction
- The geometry of linear equations
- An example of Gaussian elimination
- Matrix notation and matrix multiplication
- Triangular factors and row exchanges
- Inverses and transposes

Vector Spaces and Linear Equations
- Vector spaces and subspaces
- Solution of m equations in n unknowns
- Linear independence, basis, and dimension
- The four fundamental subspaces
- Linear transformations

Orthogonality
- Perpendicular vectors and orthogonal subspaces
- Inner products and projections onto lines
- Orthogonal bases, orthogonal matrices, and Gram-Schmidt orthogonalization
- The fast Fourier transform

Determinants
- The properties of the determinant
- Formulas for the determinant
- Applications of determinants

Eigenvalues and Eigenvectors
- The diagonal form of a matrix
- Difference equations and the powers A^k
- Differential equations and exponential e^{At}
- Complex matrices: Symmetric vs. hermitian and orthogonal vs. unitary
- Similarity transformations

Positive Definite Matrices
- Minima, maxima, and saddle points
- Tests for positive definiteness
- Semidefinite and indefinite matrices: $Ax = \lambda Mx$
- Minimum principles and the Rayleigh quotient
- The finite element method

Computations with Matrices
- The norm and condition number of a matrix
- Householder transformation
- Hessenberg form
- Gaussian elimination with pivoting

Linear Programming
- Linear inequalities
- The simplex method
- The theory of duality

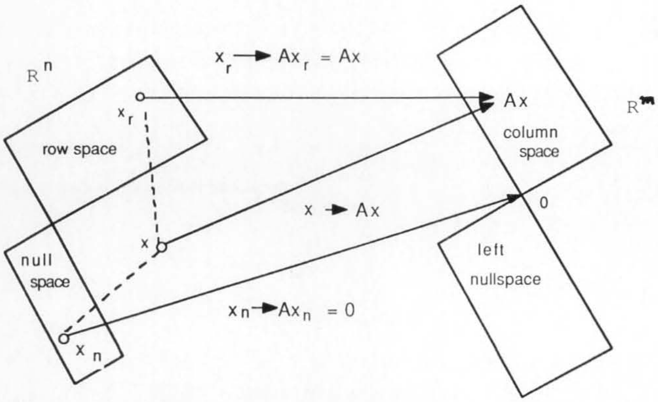

FIGURE 1. The action of a matrix A (from Strang, 1988)

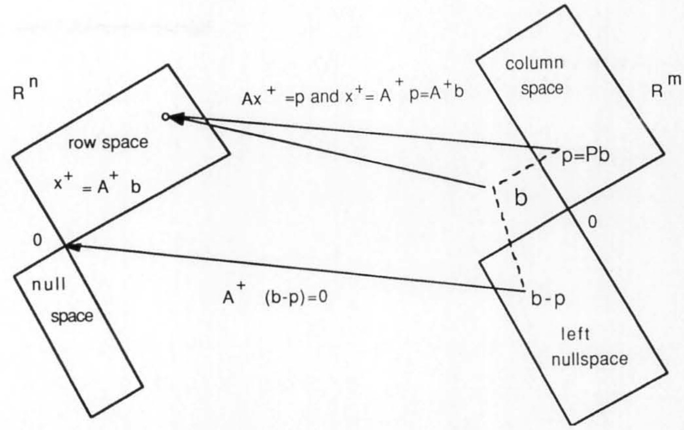

FIGURE 2. The action of the Moore-Penrose pseudoinverse of A (from Strang, 1988)

temperature measurements along the tower. The matrix that relates the inputs and outputs would be of dimension 5x2. Because of noise or biases, the temperature measurements would usually be inconsistent when compared to that calculated from the physical and thermodynamic data of the components and process involved. The solution to $Ax=b$ in this case represents the input necessary to give a set of measurements "closest" to the desired outputs as measured by the sensors. These presentations on the four fundamental subspaces pave the way for introduction of the singular value decomposition (SVD), the pseudoinverse, and application of SVD in solving the least-square problems. SVD has proven to have many applications in system analysis and plays a major role in the implementation of many stable numerical algorithms. See Klema and Laub [10], for example, for more detailed discussion concerning the numerical aspect of SVD.

Let $A=U\Sigma V'$ be the SVD of A. We present the notion that transformation of a vector x by A can be viewed as series of transformations: first a rotation by V', a unitary matrix, followed by a decoupled transformation represented by the diagonal Σ, followed by another rotation by the unitary U. The notion that the columns of the matrices U and V in serving as the orthonormal basis vectors of the appropriate subspaces is presented.

The concept of singular values of a matrix is presented as follows (this geometric representation is borrowed from that of Moore [11]). If an r-dimensional sphere of unit radius resides in the row space of matrix A, with the r orthogonal unit vectors given by the first r columns of the matrix V as the coordinate axes (r denotes the rank of A), then the transformation process maps it into an r-dimensional ellipsoid in the column space of A. The nonzero singular values of A represent the magnitudes of the axes of the ellipsoid (the largest singular value gives the length of the major axis, *etc.*). The mutually orthogonal axes of the ellipsoid point in the directions given by the first r columns of the matrix U. In this way, the singular values can be viewed as scaling factors for the unit radii of a sphere in the row space when mapped into an ellipsoid in the coulmn space of A. Again, the students are asked to picture the various manipulations in 3-D space. Finally, the concept of the pseudo-inverse of A is presented. The roles of U' and V in accomplishing projection and change of basis are carefully presented, using Figure 2 as an aid.

At this point, a computer assignment is made for finding the completely controllable, completely observable, and completely controllable and observable subspaces of a linear dynamic system, described by the equation $\dot{x}(t)=Ax+Bu$, where x represents the state vector and u represents the input vector. The idea is that from the controllability and observability grammians of the system, (positive definite solutions, P and M, to the Lyaponov equations, below)

$$AP + PA' = -BB'$$
$$A'M + MA = -C'C$$

one can project the original state space down to the controllable or observable subspace spanned by the columns of P or M, respectively, by doing an equivalence transformation, using a set of orthogonal basis vectors that span the appropriate subspace, for the transformation. The rank of each of the subspaces is the rank of P or M respectively. Stable routines exist for solving equations of the above type. The matrices P and M can also be solved in a stable manner by assuming a QR decomposition of A, and in conjunction with back substitutions, the elements of P and M can be determined in a straightforward manner.

This exercise also illustrates that often a good algorithm can be ruined by bad numerics. Let me explain. The controllability or observability of a system can also be analyzed by examining the rank and the span of the associated controllability or observability matrix U and V as calculated by

$$U = [B|\ AB|\ A^2B|\ \dots\ |\ A^{(n-1)}B]$$
$$V = [C|\ CA|\ CA^2|\ \dots\ |\ CA^{(n-1)}]$$

respectively. In order to calculate U and V, repeated multiplications by A up to (n-1) times are necessary. If n is large and A is poorly conditioned, then it can lead to numerical instability such that rank determination of the resultant U and V may be obscured by their near singularity; the singularity may have been an artifact of the numerics and not necessarily a representation of any physical defect. For the completely controllable and observable system, one finds the intersection of the two respective subspaces by projecting, for example, the controllable subspace down to the observable subspace. A good illustration of applying numerical linear algebra to system concept here is that if one only desires to test the controllability (observability) of a system, one can normally get accurate results by applying a random state feedback (observer) through gain K (F), to form A + BK (A + FC) in the state propagation equation, where K (F) is randomly chosen. Then one computes the eigenvalues of A and A + BK (A + FC) and pair off nearest eigenvalues between the two matrices. The system is completely controllable (completely observable) if, and only if, the two matrices A and A + BK (A + FC) have no common eigenvalues with probability 1.

About two thirds of the course is spent in covering the first three chapters and the appendix on pseudoinverse, which we consider to be the heart of the matter. Each notion is presented geometrically and intuitively as much as the subject matter allows. Sometimes it takes quite a few lectures to get an idea across. But each decomposition and manipulation is accompanied by an explanation of why one wants to do that decomposition and manipulation and what does it get you? As many physical examples are offered as possible. In this respect, Strang's presentation of the material does lend a much more intuitive appeal than some of the other textbooks.

SECOND HALF OF COURSE

The second portion of the course starts with a review of the properties of determinants. This is followed by the next four chapters on eigenvalues and eigenvectors, positive definite matrices, computations with matrices, and linear programming. The book is followed fairly closely except for the chapter on computations with matrices. For this subject matter, Strang is supplemented by materials from Stewart [8] and Golub and Van Loan [9], which both deal with the numerical aspect of matrix computations. The Gaussian elimination with pivoting is presented first, and is followed by the Householder's transformation and upper Hessenberg matrices and their significance in speeding up the computation efficiency. The condition number and the Raleigh's quotient of a matrix are discussed with respect to stability and perturbation.

At this point, physical examples are offered to illustrate the danger of dealing with a matrix with a high condition number. The students are asked to visualize a system with states residing in an ellipsoid with two long major axes and a very short third minor axis. Suppose one wants to find the control input required to produce some desired states. Such system matrix with high condition number would yield a very large control input upon inversion of the matrix. Therefore, the students are asked to ponder if it would not have been more appropriate to lop off one dimension (the one spanned by the short axis) and project the original system down to a subspace with dimension of one less.

A computer project is assigned to consider a 2x2 case where the gain matrix of a system is derived experimentally where the measurements are rather noisy. The students are asked to calculate inputs necessary in order to yield certain output vector values. The condition number of the gain matrix given is rather high due to the fact that the real gain matrix is singular, because only one of the two outputs is independent. But, due to noise, the experimentally derived gain matrix is not singular, but rather is near singular. The students are to compare the sensitivity of the calculated results using the original full matrix with slightly varying entries to reflect the noisy nature of the data. Further, they are asked to offer a plausible explanation for the high sensitivity of the calculated results to the slight perturbations in the system matrix entries and to offer a solution for avoiding this problem. The students are asked how to compute, using SVD, the reduced order models to eliminate modes which have little effect on system response. They find this exercise enlightening.

The presentation of eigenvalues and eigenvectors is straightforward. The intuitive approach has not been used much except where the notions from the first part of the course apply. A note has to be said about the Jordan canonical form of a matrix A. In every linear algebra textbook there is a section devoted to the explanation and calculation of the Jordan canonical form of a matrix A. Some emphasize it more than others. However, when dealing with large systems (as in many practical problems) where computers are employed for matrix manipulations, an approach employing the calculation of the Jordan decomposition, *i.e.*, $X^{-1} AX = \text{diag}(J_1, \ldots, J_t)$, where each J is a Jordan block, is not numerically stable. This comes about because at several steps of calculating the decomposition, rank decisions must be made, and the final computed block structure depends heavily on these blocks, thus on these rank decisions. In practical applications, Golub and Van Loan suggest using the more stable Schur decomposition in eigenvector problems. Therefore, the Jordan canonical approach is not covered in detail in this course.

The course has now been taught twice at our university, and the students have received it with enthusiasm. Many of them have taken courses in linear algebra in the mathematics department prior to taking this course. They comment that the approach taken here is very different and that their intuitive understanding of the key theorems has increased. They further state that this course has helped them to better understand papers involving matrix manipulations.

CONCLUSION

A new applied linear algebra course, cross-listed in three engineering departments, has been created. The emphasis is on intuitive understanding and geometric visualization and interpretation of the key theorems of linear algebra. The students should learn the why's of doing certain matrix decompositions and manipulations and should be able to visualize the algorithms in 3-D space. Numerous physical examples from systems area are offered, tying together the mathematical manipulations and their physical significance. Computer projects are assigned from time to time to illustrate the utility of the various algorithms in solving practical problems. The course has also been made a co-requisite for the linear systems theory course offered by the electrical engineering department, so as to take the pain of teaching simultaneously both the applied linear algebra and linear systems theory out of that course. The students who have taken the course appreciate its approach, and I have found that every time I have taught it, I find more points that I am able to interpret intuitively that I was not able to before. The Chinese have an old proverb that says that new things are learned from reviewing old things. It has proven to be the case with this course.

REFERENCES

1. Zygourakis, K., "Linear Algebra for Chemical Engineers," *Chem. Eng. Ed.*, **18**, 176 (1984)
2. *Pro-MATLAB,* The MathWorks, Inc., South Natick, MA
3. Kantor, J.C., "Matrix Oriented Computation Using Matlab," *CACHE News,* **28**, 27 (1989)
4. LINPACK, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA
5. EISPACK, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA
6. Moore, C.F., "A Reliable Distillation Column Analysis Procedure for Use During Initial Column Design," paper presented at the November meeting of AIChE (1985)
7. Strang, G., *Linear Algebra and Its Applications,* 3rd ed., Harcourt Brace Jovanovich (1988)
8. Stewart, G.W., *Introduction to Matrix Computations,* Academic Press (1973)
9. Golub, G., and C. Van Loan, *Matrix Computations,* Johns Hopkins Press (1983)
10. Klema, V.C., and A.J. Laub, "The Singular Value Decomposition: Its Computation and Some Applications," *IEEE Trans. on Auto. Cont.,* **AC-25**, 164 (1980)
11. Moore, B., internal report ELE-1633-F, System Control Group, Department of Electrical Engineering, University of Toronto, September (1978) ❐

RANDOM THOUGHTS

Continued from page 207

> *commentary. But when we comment on practice tests or revisable papers we are not saying, "Here's why you got this grade." We are saying, "Here's how you can get a better grade."*

Alternating between the roles of student advocate and guardian of standards–good cop and bad cop–enables teachers to serve comfortably in both capacities. It's easier to set high standards if you know you're going to be helping the students attain them, and it's easier to enforce the standards once you've made them quite clear and given the students every opportunity to meet them. In addition, the approach may also provide a significant fringe benefit:

> *In the end, I do not think I am just talking about how to serve students and serve knowledge or society. I am also talking about developing opposite and complementary sides of our character or personality: the supportive and nurturant side and the tough, demanding side. I submit that we all have instincts and needs of both sorts. The gentlest, softest, and most flexible among us really need a chance to stick up for our latent high standards, and the most hawk-eyed, critical-minded bouncers at the bar of civilization among us really need a chance to use our nurturant and supportive muscles instead of always being adversary.*

There's much more. Get the book.

REFERENCES

1. Peter Elbow, *Embracing Contraries: Explorations in Learning and Teaching,* New York, Oxford University Press (1986)