ADVANCED LINEAR ALGEBRA
CLASS DIARY FOR MATH 4377, FALL 2003

Time and location: 10am–11am MWF, in 347 PGH
Instructor: J. Hausen, 666 PGH, 713–743–3479, hausen@math.uh.edu
Office Hours: 1pm–2pm MWF, and by appointment
Text: *Linear Algebra* (Second Edition), K. Hoffman and R. Kunze, Prentice Hall, Englewood Cliffs, NJ, 1971 (ISBN 0-13-536797-2).

These class notes are intended as a compact diary of the course. References to our text are made in the form [HK, §1.1], etc.

1. Chapter 1. Linear Equations

Session 1, Monday, August 25. Distribute Course Info sheet. Lecture on **§1.1, Fields:** Definition. Examples and nonexamples. Subfields. The finite fields $\mathbf{Z_p}$ with p a prime. A field F has characteristic zero if the element 1 added to itself any finite number of times is **non**zero: $1 + 1 + \cdots + 1 \neq 0$; if it is possible to obtain the zero element of F as the sum of finitely many 1s, then the characteristic of F is the least positive integer n such that $1 + 1 + \cdots + 1 = 0$ (n terms). The field $\mathbf{C}$ of complex numbers has characteristic zero (and so do all subfields of $\mathbf{C}$). For p a prime, the field $\mathbf{Z_p}$ has characteristic p. For example, in $\mathbf{Z_3} = \{\mathbf{0}, \mathbf{1}, \mathbf{2}\}$, $1 + 1 + 1 = 0$ but $1 \neq 0$ and $1 + 1 = 2 \neq 0$.

Session 2, Wednesday, August 27. Lecture on **§1.2, Systems of Linear Equations:** Solution to a system of m linear equations in n variables over a field F. Homogeneous systems. Linear combination of the equations of a system. Equivalent systems. Equivalent systems have exactly the same solutions.

Session 3, Friday, August 29. Lecture on **§1.3, Matrices and Elementary Row Operations:** Matrix representation $AX = Y$ of a system of linear equations over F. Definition of matrix. The three types of elementary row operations. Row operations are invertible and their inverses are elementary row operations of the same type. Row–equivalent matrices. If A and B are $m \times n$ matrices which are row–equivalent, then the homogeneous systems $AX = 0$ and $BX = 0$ have exactly the same solutions.

Session 4, Wednesday, September 3. Every matrix over a field is row–equivalent to a row–reduced matrix. Lecture on **§1.4. Row–Reduced Echelon Matrices:** Definition. Every matrix over a field is row–equivalent to a row–reduced echelon matrix. Method to solve a homogeneous system. Free (or independent) variables, dependent variables.

Session 5, Friday, September 5. Continue §1.4. Let R be an $m \times n$ row–reduced echelon matrix with r nonzero rows and consider the homogeneous system $RX = 0$ in variables $x_1, \ldots, x_n$. A variable x_j is **free** (or **independent**) if the jth column of R **does not** contain the leading 1 of a nonzero row; if the jth column of R **does** contain the leading 1 of a nonzero row, then x_j is said to be **dependent**. Note that there are r dependent variables and, hence, $n-r$ independent variables (called $u_1, \ldots, u_{n-r}$ in the text). For $i = 1, \ldots, r$, the ith equation of the system $RX = 0$ contains exactly one dependent variable with nonzero coefficient, namely $1 \cdot x_{k_i}$ where k_i is the column number of the column containing the leading 1 of the ith row. Choosing arbitrary values for the free variables, one easily solves the r nonzero equations for the dependent variables; the remaining $n-r$ zero rows correspond to equations of the form $0 \cdot x_1 + \cdots + 0 \cdot x_n = 0$. A solution $x_1, \ldots x_n$ to the homogeneous system is **trivial** if $x_1 = x_2 = \cdots = x_n = 0$. A homogenous system with more variables than equations has (at least one) non–trivial solution. An $n \times n$ matrix A is row equivalent to the $n \times n$ identity matrix if and only if the system $AX = 0$ has only the trivial solution.

Session 6, Monday, September 8. Finish §1.4. Consider systems of m linear equations in n variables $x_1, \ldots, x_n$ which are not linear. If $AX = Y$ is such a system, the $m \times 1$ "right hand side" column vector Y is not the zero vector. Define $A' = [A \,|\, Y]$ to be the augmented matrix of the system. Then A' is an $m \times (n+1)$ matrix. Suppose that e is an elementary row operation. If $e([A \,|\, Y]) = [B \,|\, Z]$ where Z is an $m \times 1$ column vector, then the systems $AX = Y$ and $BX = Z$ are equivalent systems. By [HK, p. 5, Theorem 1], $AX = Y$ and $BX = Z$ have exactly the same solutions. Hence, two systems $AX = Y$ and $BX = Z$ have exactly the same solutions if the augmented matrices $[A \,|\, Y]$ and $[B \,|\, Z]$ are row eqivalent. Suppose $e_1, e_2, \ldots, e_k$ are elementary row operations such that $e_k(e_{k-1}(\ldots(e_2(e_1(A))\ldots)) = R$ is a row–reduced echelon matrix. Then $e_k(e_{k-1}(\ldots(e_2(e_1([A \,|\, Y]))\ldots)) = [R \,|\, Z]$ for some $m \times 1$ column vector Z. Let r be the number of nonzero rows of R. Then $RX = Z$ (and hence $AX = Y$) is solvable iff each entry in rows $r+1, r+2, \ldots, m$ of Z is 0. If this is the case, the set of all solutions to $RX = Z$ (and hence to $AX = Y$) can be found as in the case of homogeneous systems: Suppose for $i = 1, \ldots, r$, the leading 1 in the ith row is located in column k_i, $1 \le i \le r$. Denoting the free variables by $u_1, \ldots, u_{n-r}$ and the coefficient of u_j in the ith equation of $RX = Z$ by C_{ij}, the ith equation of $RX = Z$ is

$$x_{k_i} + \sum_{j=1}^{n-r} C_{ij} u_j = z_i, \; i = 1, \ldots, r\,.$$

The remaining $m-r$ equations are of the form $0 = 0$. Example 9. Suppose F is a subfield of a field E (Example: $F = \mathbb{R}$ and $E = \mathbb{C}$) and $AX = Y$ is a system of linear equations with all scalars including the right hand sides y_i, $i = 1, \ldots, m$, belonging to F. If it is known that $AX = Y$ has a solution $x_1, \ldots, x_n$ with each $x_i \in E$, then $AX = Y$ has also a solution $x'_1, \ldots, x'_n$ with each $x'_i \in F$. This follows from the fact that the elementary row operations used to reduce A to a row–reduced echelon matrix, R, can be chosen to involve multiplication by scalars which all belong to F. Thus, the augmented matrix $[R|Z]$ will be a matrix over F, and solvability is determined solely by examining the last $m-r$ entries of Z. Should one of the z_i, $i = r+1, \dots, m$ be nonzero, $RX = Z$ (and hence $AX = Y$) will have no solution, not in F and not in E. Given solvability in E, this cannot happen. The method outlined above provides all the solutions in F.

Session 7, Wednesday, September 10. Work Problem 8, page 16, in class. Lecture on **§1.5, Matrix Multiplication:** Let F be a field, and let A and B be matrices over F of size $m \times n$ and $n \times p$, respectively. Then the product AB of A and B is defined to be the $m \times p$ matrix C with $C_{ij} = A_{i1}B_{1j} + A_{i2}B_{2j} + \cdots + A_{in}B_{nj} = \sum_{r=1}^{n} A_{ir}B_{rj}$, $i = 1, \dots, m$, $j = 1, \dots, p$. If the number of columns of A does **not** equal the number of rows of B, then AB is undefined. Given row vectors $\beta_i = [B_{i1}, \dots, B_{ip}]$, $i = 1, \dots, n$, and scalars $c_1, \dots, c_n$, one defines $c_1\beta_1 + \cdots + c_n\beta_n$ to equal the row vector, call it γ, given by

$$\gamma = \sum_{i=1}^{n} c_i\beta_i = \sum_{i=1}^{n} c_i[B_{i1}, \dots, B_{ip}] = \sum_{i=1}^{n} [c_iB_{i1}, \dots, c_iB_{ip}]$$

and says that γ is a **linear combination** of $\beta_1, \dots, \beta_n$.

Session 8, Friday, September 12. Continue §1.5. Given matrices A of size $m \times n$ and B of size $n \times p$, each row of the matrix product $AB = C$ is a linear combination of the rows of B: If β_i denotes the ith row of B and γ_i the ith row of C, then, for each $i = 1, \dots, m$,

$$\gamma_i = [C_{i1}, \dots, C_{ip}] = [\sum_{r=1}^{n} A_{ir}B_{r1}, \dots, \sum_{r=1}^{n} A_{ir}B_{rp}] = \sum_{r=1}^{n} A_{ir}[B_{r1}, \dots, B_{rp}]$$

which equals $\sum_{r=1}^{n} A_{ir}\beta_r$. In general, $AB \neq BA$ even if both products are defined. If I is the $n \times n$ identity matrix then $AI = A$ and $IB = B$ whenever the matrix products are defined. For the $m \times n$ zero matrix $0^{m,n}$, all matrix products $A \cdot 0^{m,n}$ and $0^{m,n} \cdot B$, assuming they are defined, are zero matrices of the appropriate sizes. Note that a system of m linear equations in n variables $x_1, \dots, x_n$ and right-hand-sides $y_1, \dots, y_m$ (as the system (1-1) on page 3) can be written as a matrix product $A \cdot X = Y$ where A is the $m \times n$ matrix of coefficients, X is the $n \times 1$ matrix containing the variables and Y is the $m \times 1$ matrix containing the right-hand-side constants. Given an $n \times p$ matrix B, let B_j denote the $j$the column of B, i.e. B_j is the $n \times 1$ matrix with B_{ij} in row i. Then $B = [B_1, \dots, B_p]$ is **partitioned** into its columns. Given any $m \times n$ matrix A, one has $AB = [AB_1, \dots, AB_p]$, i.e. the jth column of AB is the matrix product AB_j, $j = 1, \dots, p$. Matrix multiplication (if defined) is associative.

Session 9, Monday, September 15. Continue §1.5. An elementary matrix is a (necessarily square) matrix which is obtained from an identity matrix by a single elementary row operation. Let e be an elementary row operation defined on the set of m–rowed matrices over F, and $E = e(I)$ where I denotes the $m \times m$ identity matrix. Then $e(A) = E \cdot A$ for all $m \times n$ matrices A. Thus, an $m \times n$ matrix B is is row equivalent to A iff $B = PA$ where P is a product of finitely many $m \times m$ elementary matrices.

Session 10, Wednesday, September 17. Questions on Homework Set 3? Lecture on §1.6: Invertible Matrices. Let A be an $n \times n$ matrix over the field F. A **left inverse** of A is an $n \times n$ matrix B over F such that $BA = I$; a **right inverse** of A is an $n \times n$ matrix C over F such that $AC = I$. If $BA = AB = I$, then B is said to be a **two–sided inverse** of A and A is said to be **invertible**. If a square matrix has a left inverse, B, and a right inverse, C, then $B = C$ and A is invertible with unique inverse $A^{-1} = B = C$. The inverse of an invertible matrix is invertible; products of invertible matrices are invertible. Every elementary matrix is invertible and its inverse is an elementary matrix. Thus, products of elementary $m \times m$ matrices are invertible.

Session 11, Friday, September 19. Continue lecture on §1.6. All matrices are matrices over some given field F. **Re homework problems:** Let A be an $m \times n$ matrix, and let R be a row–reduced echelon matrix which is row–equivalent to A. Consider the following problems: (1) Find R. (2) Find an invertible $m \times m$ matrix P such that $R = PA$. (3) Given an $m \times 1$ matrix Y_0 with entries in F, find all solutions to $AX = Y_0$. (4) For the case that $m = n$, determine whether A is invertible and, if it is, find A^{-1}. In order to do a problem of type (1), choose your favorite elementary row operations $e_1, e_2, \dots e_k$ such that $R = e_k(e_{k-1}(\dots(e_2(e_1(A)))\dots))$. Let $E_i = e_i(I)$ where I is the $m \times m$ identity matrix. Let $P = E_k E_{k-1} \dots E_1$. Then, as a product of elementary matices, each of which is invertible, P is invertible (page 22, Corollary and Theorem 11). For every m–rowed matrix M over F, $e_k(e_{k-1}(\dots(e_2(e_1(M)))\dots)) = PM$. For problems of the other types, it is a good approach to consider the augmented matrix $[A|Y]$ with Y the $m \times 1$ matrix containing the variables $y_1, \dots, y_m$. Then $e_k(e_{k-1}(\dots(e_2(e_1([A|Y])))\dots)) = P[A|Y] = [PA|PY] = [R|Z]$ [HK, page 19, top paragraph]. The entries of the $m \times 1$ matrix $PY = Z$ are linear functions of the $y_1, \dots, y_m$. For problems of type (2), note that $P = PI = [PI_1, \dots, PI_m]$, so that the the jth column P_j of P is the product $PI_j = e_k(e_{k-1}(\dots(e_2(e_1(I_j)))\dots))$, $j = 1, \dots, m$. Substituting, in the matix Z, $y_j = 1$ and $y_i = 0$ for all $i \neq j$, one can read off the jth column of P. For problems of type (3), if the entry in row i of Y_0 is the scalar c_i, let Z_0 be the matrix obtained from Z by replacing each y_i with c_i. $AX = Y_0$ is solvable iff $PAX = PY_0$ is solvable iff $RX = Z_0$ is solvable iff the entries of Z_0 in rows $r+1, \dots, m$ are zero should the number of nonzero rows, r, be less than m. Assuming this is the case, one assigns an arbitrary value to every variable x_j with j having the property that the jth column of R does **not** contain a leading 1 of a nonzero row; one then solves the ith of the r nonzero equations of $RX = Z_0$ for x_{k_i} where k_i denotes the column number of the leading 1 of the ith row of R. See [HK, page 12, bottom paragraph]. For problems of type (4), suppose $m = n$, i.e. A has size $n \times n$. As before, $[R|Z] = [PA|PY] = P[A|Y]$. If R has a zero row, then $RX = 0$ has the nontrivial solution $X = I_n$ where I_n denotes the nth column of the $n \times n$ identity matrix. The solution sets of $AX = 0$ and of $RX = 0$ being equal [HK, page 7, Theorem 3], $AX = 0$ has a nontrivial solution and the following Super Theorem (or [HK, page 23, Theorem 13]) implies A is not invertible. Suppose R has no zero rows. One verifies that the only row–reduced echelon matrix of size $n \times n$ with no zero rows is the $n \times n$ identity matrix (induct on n). Thus $I = R = AP$ which implies $P = A^{-1}$, and you determine the columns of P from X as described for type (2) problems: For $j = 1, \ldots, n$, the jth column of $P = A^{-1}$, is obtained from $Z = PY$ by substituting $y_j = 1$ and $y_i = 0$ for all $i \neq j$.

We proved:

Super Theorem. For an $n \times n$ matrix A over a field, the following conditions are equivalent:

(1) A is row–equivalent to the identity matrix.
(2) A is a product of elementary matrices.
(3) A is invertible.
(4) A has a left inverse.
(5) A has a right inverse.
(6) The homogeneous system $AX = 0$ has only the trivial solution $X = 0$.
(7) The linear system $AX = Y$ has a solution X for each $n \times 1$ matrix Y.

Session 12, Monday, September 22. Finish §1.6. Immediate consequences of the Super Theorem are Theorems 12 and 13 and the first corollary on page 24 of [HK]; the second corollary on page 23 follows from the Super Theorem together with the corollary on page 20. For the first corollary on page 23, note that if the e_i are elementary row operations with $e_k(e_{k-1}(\ldots(e_2(e_1(A)))\ldots)) = I$ and $E_i = e_i(I)$, then $P = E_k \ldots E_2E_1$ has the property that $PA = I$, so $P = A^{-1}$ and

$$A^{-1} = P = e_k(e_{k-1}(\ldots(e_2(e_1(I)))\ldots)) .$$

The final result of §1.6 is the second corollary on page 24: If A is a product of $n \times n$ invertible matrices, then A is invertible by [HK, page 22, Corollary]. Conversely if A is invertible and a product of k matrices A_i of size $n \times n$, then each of the A_i is invertible. This is proved by induction on k.

Session 13, Wednesday, September 24. Homework day. Several students requested to see a solution to Exercise 7, §1.6, page 27. Bethany Fields and Melahat Almus volunteered and presented their solutions. Thanks to both. Lecture on Chapter 2: Vector Spaces. Definition and Examples.

Session 14, Friday, September 26. Review for Test 1.

Session 15, Monday, September 29. Test 1.

2. Chapter 2. Vector Spaces

Session 16, Wednesday, October 1. Lecture on §2.1. Geometrical interpretation of vector addition and scalar multiplication for the vector spaces $\mathbf{R^2}$ and $\mathbf{R^3}$. Proof of 2-8, 2-9, and 2-10 on page 31. Linear combinations. Definition of a subspace. Examples for subspaces of $\mathbf{R^2}$ and $\mathbf{R^3}$.

Session 17, Friday, October 3. Continue lecture on §2.2. Let V be a vector space over the field F and let $W \subseteq V$. Write $W \leq V$ to signify that W is a subspace of V. Note that the empty set is never a subspace. If $W \neq \emptyset$, then $W \leq V$ iff, for all $\alpha, \beta \in W$ and all $c \in F$, $c\alpha + \beta \in W$. Every vector space is a subspace of itself, and the subset $\{\mathbf{0}\}$ consisting of only the zero vector is a subspace. The set of all solutions to a homogeneous system $AX = 0$ of m linear equations in n unknowns is a subspace of $F^{n,1}$ (or of F^n if you regard $n \times 1$ matrices over F as n–tuples). The intersection of any nonempty collection of subspaces of V is a subspace of V. Given a set S of vectors in V, the intersection of all subspaces of V containing S is said to be the subspace **spanned** by S. Thus, the subspace spanned by the empty set is $\{\mathbf{0}\}$. If $S \neq \emptyset$, the subspace spanned by S consists of all linear combinations of vectors in S. Given subspaces $W_1, \ldots, W_n$ of V, $W_1 + \cdots + W_n$ is defined to be the set of all vectors α of the form $\alpha = \gamma_1 + \cdots + \gamma_n$ with $\gamma_i \in W_i$ for $i = 1, \ldots, n$.

Session 18, Monday, October 6. Continue lecture on §2.2. Given a subset S of a vector space V, the sub**space** spanned by S will be denoted by $< S >$. Given a field F, the set of all functions $f : F \to F$ which are polynomials (i.e. functions f for which there exist scalars $a_0, a_1, \ldots, a_n$ in F such that $f(x) = a_0 + a_1x + \cdots + a_nx^n$ for every $x \in F$) is a nonempty subset of the vector space of all functions from F to F (Example 3, page 30) which is closed under vector addition and scalar multiplication. By Theorem 1, page 35, $F[x]$ is a subspace and, hence, is a vector space on its own right. Denoting by f_k, for $k = 0, 1, 2, \ldots$, the function defined by $f_k(x) = x^k$ for all $x \in F$, one verifies that every polynomial in $F[x]$ is a linear combination of finitely many of the f_ks; indeed, if $f(x) = a_0 + a_1x + \cdots + a_nx^n$ for every $x \in F$, then $f = a_0f_0 + a_1f_1 + \cdots + a_nf_n$. Thus, the set $S = \{f_k | k = 0, 1, 2, \ldots\}$ spans $F[x]$. The **row space** of an $m \times n$ matrix A over F is the subspace of F^n spanned by the rows of A. Row–equivalent matrices have the same row space. To see this, note that, if C is row–equivalent to the $m \times n$ matrix B, then there exists an invertible $m \times m$ matrix P such that $C = PB$. Denote, for $i = 1, \ldots, m$, the i th row of B by β_i and the i th row of C by γ_i. By (1–4) on page 16 of the text, each γ_i is a linear combination of the β_j s. Theorem 3, page 37, implies that the row space $< \gamma_1, \ldots, \gamma_m >$ of C is contained in $< \beta_1, \ldots, \beta_m >$, the row space of B. Conversely, since $B = P^{-1}C$, the same argument proves that $< \beta_1, \ldots, \beta_m > \subseteq < \gamma_1, \ldots, \gamma_m >$. It follows that $< \beta_1, \ldots, \beta_m > = < \gamma_1, \ldots, \gamma_m >$.

Session 19, Wednesday, October 8. Lecture on §2.3. Bases and Dimension. Let V be a vector space over a field F. A set S of vectors in V is said to be **linearly dependent** if, given any finitely many pairwise distinct vectors $\alpha_1, \ldots, \alpha_n$ in S, there exist scalars $c_1, \ldots, c_n$ in F **not all of which are zero** such that $c_1\alpha_1 + \cdots + c_n\alpha_n = \mathbf{0}$. A **basis** for a vector space V is a linearly independent subset of V which spans V. **Examples:** The set of unit vectors $\{\epsilon_1, \ldots, \epsilon_n\}$ is a basis for $\mathbf{R^n}$. The set of all matrices in $F^{m\times n}$ which have exactly one nonzero entry and this entry is 1_F is a basis for $F^{m\times n}$. The set $\{1, x, x^2, \ldots\}$ is a basis for the space of all polynomial functions in $F[x]$. Given an invertible matrix P of size $n \times n$ over F, its columns are a basis for $F^{n\times 1}$. If a vector space V contains a finite spanning set consisting of m vectors, then every linearly independent set of vectors in V contains at most m elements (Theorem 4).

Session 20, Friday, October 10. Continue lecture on §2.3. Any two bases of a finite dimensional vector space contain the same (finite) number of vectors (Corollary 1, page 44). One defines the **dimension** of a finite dimensional vector space V to be the number of vectors in a (and hence in every) basis for V. Notation: $\dim V$. Thus, $\dim F^n = n$, $\dim F^{m\times n} = m \cdot n$. Note that the empty set is a basis for the zero space $\{\mathbf{0}\}$. Thus, $\dim\{\mathbf{0}\} = \mathbf{0}$. Given a finite dimensional vector space V of dimension n, every subset of V containing more that n vectors is linearly dependent; and every subset of V which spans V must contain at least n vectors (Corollary 2, page 44). If S is a linearly independent subset of a vector space V and β is a vector in V which is not contained in the span of S, then the set $S \cup \{\beta\}$ is linearly independent (Lemma, page 45). Applications of this lemma to $\mathbf{R^3}$.

Session 21, Monday, October 13. Continue lecture on §2.3. We used the Lemma on page 45 to prove the following version of Theorem 5: If V is a finite dimensional vector space over F and W is a subspace of V containing a linearly independent set S_0 of vectors, then (i) S_0 is finite; (ii) S_0 can be extended to a finite basis S for W; and (iii) there exists a basis B for V containing S. Consequences: (1) Subspaces of finite dimensional vector spaces are finite dimensional. (2) A linearly independent subset S_0 of a finite dimensional vector space V can be extended to a basis of V. (3) If W is a proper subspace of a finite dimensional vector space V, then $\dim W < \dim V$. The last property can be seen as follows: If $W \neq V$ is a proper subspace of the finite dimensional vector space V and S is a basis of W, then S is a linearly independent set of vectors in V. By Theorem 4, if $\dim V = n$, then $|S| \leq n$. Since $W =< S >\neq V$, there exists $\alpha \in V$ such that $\alpha \notin W$ and, by the lemma, the set $S \cup \{\alpha\}$ is linearly independent. Again, $|S \cup \{\alpha\}| \leq n$ proving $|S| < n$. If the rows of an $n \times n$ matrix A over F are linearly independent, then A is invertible (Corollary 3). Statement of Theorem 6: If W_1 and W_2 are subspaces of a finite dimensional vector space V, then $W_1 + W_2$ and $W_1 \cap W_2$ are subspaces of V and

$$\dim(W_1 + W_2) = \dim W_1 + \dim W_2 - \dim(W_1 \cap W_2)\,.$$

Example: If W_1 and W_2 are two–dimensional subspaces of $\mathbf{R^3}$, then their intersection cannot equal the zero subspace: $W_1 \cap W_2 \neq \{\mathbf{0}\}$.

Session 22, Wednesday, October 15. Work problems from HW Set 6 with Maureen Royce.

Session 23, Friday, October 17. Postpone proof of the dimension formula, Theorem 6. Remark: Given a (ordered) **sequence** $\alpha_1, \ldots, \alpha_n$ of vectors, define the sequence to be linearly dependent if there exist scalars $c_1, \ldots, c_n$ not all of which are zero such that $c_1\alpha_1 + \cdots + c_n\alpha_n = \mathbf{0}$. If this is not the case, then the sequence of vectors is defined to be linearly independent. Note that $\alpha_1, \ldots, \alpha_n$ is linearly dependent if some vector appears twice in the sequence, i.e. there exist i, j such that $i \neq j$ and $\alpha_i = \alpha_j$. Let V be a vector space over the field F and suppose V has finite dimension n. Let $\mathcal{B}$ denote an ordered basis for V consisting of the sequence $\alpha_1, \ldots, \alpha_n$ of vectors. Let $\beta \in V$. Then there exist unique scalars $b_1, \ldots, b_n$ such that $\beta = b_1\alpha_1 + \cdots + b_n\alpha_n$. The n–tuple $(b_1, \ldots, b_n) \in F^n$ is called the **coordinate vector** of β with respect to the ordered basis $\mathcal{B}$, and the **coordinate matrix of β with respect to the ordered basis** $\mathcal{B}$ is the $n \times 1$ matrix, denoted by $[\beta]_{\mathcal{B}}$, such that

$$[\beta]_{\mathcal{B}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Theorem 7 concerns the relationship between the coordinates of the same vector $\beta \in V$ with respect to two different ordered bases for V: Suppose $\mathcal{B}$ is an ordered basis consisting of vectors $\alpha_1, \ldots, \alpha_n$ and $\mathcal{B}'$ is an ordered basis consisting of vectors $\alpha'_1, \ldots, \alpha'_n$, then there exists a unique $n \times n$ matrix P over F, which is necessarily invertible, such that, for every $\beta \in V$, $[\beta]_{\mathcal{B}} = P[\beta]_{\mathcal{B}'}$ (and hence, $[\beta]_{\mathcal{B}'} = P^{-1}[\beta]_{\mathcal{B}}$). Moreover, the jth column of P is $P_j = [\alpha'_j]_{\mathcal{B}}$.

Session 24, Monday, October 20. An example was given with $V = F^{2\times 2}$ where F is some subspace of the field of complex numbers. Consider the ordered basis $\mathcal{B}$ which consists of the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Note that,

$$\text{if } M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \text{then} \quad [M]_{\mathcal{B}} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}.$$

Consider the following matrices:

$$A' = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, B' = \begin{bmatrix} 0 & 1 \\ -1 & 4 \end{bmatrix}, C' = \begin{bmatrix} 0 & 2 \\ 0 & 5 \end{bmatrix}, D' = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}.$$

If a, b, c, d are scalars such that $aA' + bB' + cC' + dD' = \mathbf{0}$, the 2×2 zero matrix, one verifies that $a = 0 = b = c = d$ which proves that the four primed matrices are linearly independent and, hence, constitute another ordered basis. Denote by $\mathcal{B}'$ the ordered basis $\{A', B', C', D'\}$. Let $P = [\,[A']_{\mathcal{B}}, [B']_{\mathcal{B}}, [C']_{\mathcal{B}}, [C']_{\mathcal{B}}]$. Thus,

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & -1 & 0 & 0 \\ -3 & 4 & 5 & -1 \end{bmatrix}.$$

Row reducing the augmented 4×5 matrix $[P|Y]$, one obtains

$$[P|Y] = \begin{bmatrix} 1 & 0 & 0 & 0 & y_1 \\ 0 & 1 & 2 & 0 & y_2 \\ 2 & -1 & 0 & 0 & y_3 \\ -3 & 4 & 5 & -1 & y_4 \end{bmatrix} \longrightarrow \cdots \longrightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & y_1 \\ 0 & 1 & 0 & 0 & 2y_1 - y_3 \\ 0 & 0 & 1 & 0 & -y_1 + \frac{1}{2}y_2 + \frac{1}{2}y_3 \\ 0 & 0 & 0 & 1 & \frac{5}{2}y_2 - \frac{3}{2}y_3 - y_4 \end{bmatrix}$$

which implies that

$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & -1 & 0 \\ -1 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{5}{2} & -\frac{3}{2} & -1 \end{bmatrix}.$$

Thus, for example,

$$\text{if } M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \text{then} \quad [M]_{\mathcal{B}'} = P^{-1}[M]_{\mathcal{B}} = P^{-1} \cdot \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} a \\ 2a - c \\ -a + \frac{1}{2}b + \frac{1}{2}c \\ \frac{5}{2}b - \frac{3}{2}c - d \end{bmatrix}.$$

Theorem 8 provides a source for finding different ordered bases for a finite dimensional vector space: If V is an n–dimensional vector space with ordered basis $\mathcal{B}$ then, given any invertible $n \times n$ matrix P, there exists a unique ordered basis $\mathcal{B}'$ for V such that, for every vector $\beta \in V$, $[\beta]_{\mathcal{B}} = P[\beta]_{\mathcal{B}'}$ and, consequently, $[\beta]_{\mathcal{B}'} = P^{-1}[\beta]_{\mathcal{B}}$. For example, if $V = \mathbf{R}^2$ and

$$P = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

then P has determinant 1, so P is invertible. It follows that the homogeneous system $PX = \mathbf{0}$ has only the trivial solution $X = \mathbf{0}$. This implies that the columns of P are linearly independent. Hence $\mathcal{B}' = \{(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\}$ is an ordered basis of V. Geometrically, the vectors in $\mathcal{B}'$ are obtained from the unit vectors ϵ_1 and ϵ_2 by a rotation about the origin through the angle θ. Note that

$$P^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

If $\mathcal{B}$ denotes the standard basis $\{\epsilon_1, \epsilon_2\}$ of V, then for every vector $\beta = (x, y) \in V$,

$$[\beta]_{\mathcal{B}'} = P^{-1}[\beta]_{\mathcal{B}} = P^{-1}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta + y\sin\theta \\ -x\sin\theta + y\cos\theta \end{bmatrix}.$$

Session 25, Wednesday, October 22. Lecture on §2.5: Summary of Row Equivalence. The **row space** of an $m \times n$ matrix A over F is the subspace of F^n spanned by the rows of A. The **row rank** of a matrix is the dimension of its row space. Row–equivalent matrices have the same row space (Theorem 9). We proved (1) through (4) of the following

Lemma. Let R be an $m \times n$ row–reduced echelon matrix over F with r nonzero rows, $\rho_1, \ldots, \rho_r$, and let the leading 1 of row j occur in column k_j, $j = 1, \ldots, r$. Then $k_1 < \ldots, k_r$. Define $W = \langle \rho_1, \ldots, \rho_r \rangle$ to be the row space of R and let $\beta = (b_1, \ldots, b_n) \in W$. Then: (1) $\beta = \sum_{j=1}^{r} b_{k_j}\rho_j$. (2) The set $\{\rho_1, \ldots, \rho_r\}$ is a basis for W (i.e. Theorem 10 holds). (3) If $\beta \in W$ is a nonzero vector and the first nonzero coordinate of β is located in the t^{th} column, then $t \in \{k_1, \ldots, k_r\}$. (4) If one defines the set $S(W)$ to consist of all positive integers k for which there exists a nonzero $\beta \in W$ with first nonzero coordinate in the k^{th} column, then $S(W) = \{k_1, \ldots, k_r\}$. (5) For each $t \in \{1, \ldots, r\}$, there is one and only one vector $\beta = (b_1, \ldots, b_n) \in W$ such that $b_{k_t} = 1$ and $b_{k_j} = 0$ for all $j \neq t$, namely $\beta = \rho_t$.

Session 26, Friday, October 24. Condition (5) of the Lemma follows from condition (1) completing the proof of the Lemma. Consequences: Given positive integers m, n and a subspace W of F^n with $\dim W \leq m$, there exists exactly one row–reduced echelon matrix of size $m \times n$ having W for its row space (Theorem 11). Proof: Let $\dim W = r \leq m$ and let $\alpha_1, \ldots, \alpha_r$ be a basis for W. Let A be the $m \times n$ matrix containing $\alpha_1, \ldots, \alpha_r$ in its first r rows followed by $m - r$ zero rows. Then $A \in F^{m\times n}$ and A has row space W. By Theorem 5, page 12, A is row–equivalent to a row–reduced echelon matrix R. By Theorem 9, page 56, W is the row space of R, and by Theorem 10 (which was (2) of the Lemma), R has r nonzero rows, $\rho_1, \ldots, \rho_r$. Let the leading 1 of row j be in column k_j, $j = 1, \ldots, r$. Then $k_1 < \ldots .k_r$. Suppose $R' \in F^{m\times n}$ is another row–reduced echelon matrix having W for its row space. Then R' has r nonzero rows $\rho'_1, \ldots, \rho'_r$ with leading 1s located in columns $k'_1, \ldots, k'_r$, $k'_1 < \cdots < k'_r$. By the Lemma, $S(W) = \{k_1, \ldots, k_r\}$. Since R and R' have the same row space, W, we must have $S(W) = \{k_1, \ldots, k_r\} = \{k'_1, \ldots, k'_r\}$. Hence $k'_j = k_j$ for $j = 1, \ldots, r$. By the Lemma, part (5), for $t = 1, \ldots, r$, there is exactly one vector in W with k_t^{th} coordinate 1 and k_j^{th} coordinate 0 for all $j \neq t$, namely ρ_t, the t^{th} row of R. Note that for $t = 1, \ldots, r$, $\rho'_t \in W$ has k_t^{th} coordinate 1 and k_j^{th} coordinate 0 for all $j \neq t$. Thus, for $t = 1, \ldots, r$, the t^{th} row of R and the t^{th} row of R' are equal; the remaining $m - r$ rows of both R and R' are zero rows. It follows that $R = R'$.

Every matrix is row–equivalent to one and only one row–reduced echelon matrix (first Corollary, page 58). Two $m \times n$ matrices over F are row–equivalent iff they have the same row space (second Corollary, page 58).

Lecture on §2.6: Computations Concerning Subspaces. Tasks: Given $\alpha_1, \ldots, \alpha_m \in F^n$, how does one determine whether the vectors are linearly dependent? How does one determine whether a vector $\beta \in F^n$ belongs to the subspace spanned by $\alpha_1, \ldots, \alpha_m \in F^n$? There are two methods: Method 1 ("the Row Method") and Method 2 ("the Column Method"). See Example 21, page 60, for both. Note homework assignment due Monday.

Session 27, Monday, October 27. Review for Test 2. Graded Homework Set 8 may be picked up at my office after 6pm.

Session 28, Wednesday, October 29. Test 2 (on Chapter 2). See the web for the solutions.

3. Chapter 3. Linear Transformations

Session 29, Friday, October 31. Lecture on §3.1. Let V and W be vector spaces over the field F. A **linear transformation** from V to W is a function $T : V \to W$ such that $T(c\alpha_1 + \alpha_2) = cT(\alpha_1) + T(\alpha_2)$ for all $c \in F$ and all $\alpha_1, \alpha_2 \in V$. Examples: (1) If V denotes the set of all polynomial functions from $\mathbb{R}$ to $\mathbb{R}$, then the derivative $D : V \to V$ is a linear transformation. (2) Let $A \in F^{m\times n}$ be a fixed matrix and define $T_A : F^{n\times 1} \to F^{m\times 1}$ by $T_A(X) = A \cdot X$ for all $X \in F^{n\times 1}$. Then T_A is a linear transformation.

If $T : V \to W$ is a linear transformation, then (i) $T(0_V) = 0_W$; and (2) for all $\alpha_1, \ldots, \alpha_n \in V$ and all $c_1, \ldots, c_n \in F$,

$$T(c_1\alpha_1 + \cdots + c_n\alpha_n) = c_1T(\alpha_1) + \cdots + c_nT(\alpha_n)\,.$$

Let V and W be vector spaces over F. If V is finite dimensional and $\alpha_1, \ldots, \alpha_n$ is a basis for V then, given any vectors $\beta_1, \ldots, \beta_n \in W$, there exists a unique linear transformation $T : V \to W$ such that $T(\alpha_i) = \beta_i$ for $i = 1, \ldots, n$ (Theorem 1). Example: Let $V = \mathbf{R^2}$ and let W be the space of all functions $f : \mathbf{R} \to \mathbf{R}$ (Example 3, page 30). Then ϵ_1, ϵ_2 form a basis for V. Pick functions $\beta_i \in W$, for instance $\beta_1(x) = x^2$ and $\beta_2(x) = \sin x$ for all $x \in \mathbf{R}$, then there is a unique linear transformation $T : \mathbf{R^2} \to \mathbf{W}$ such that $T(\varepsilon_i) = \beta_i$, $i = 1, 2$; indeed, if $\alpha = (a, b) \in \mathbf{R^2}$, then $T(\alpha) = T(a\epsilon_1 + b\epsilon_2) = a\beta_1 + b\beta_2$, so $T(\alpha) : \mathbf{R} \to \mathbf{R}$ is the function $T(\alpha)(x) = (a\beta_1 + b\beta_2)(x) = ax^2 + b \sin x$ for all $x \in \mathbf{R}$.

Given a linear transformation $T : V \to W$, the **range space** of T is defined to be the set R_T of all vectors $\beta \in W$ such that $\beta = T(\alpha)$ for some $\alpha \in V$; the **null space** of T is defined to be the set N_T of all vectors $\alpha \in V$ such that $T(\alpha) = 0_W$. One verifies that R_T is a subspace of W and N_T is a subspace of V.

Session 30, Monday, November 3. Finish §3.1, Linear Transformations. If $T : V \to W$ is a linear transformation and V has finite dimension, then the null space N_T and the range space R_T are finite dimensional and $\dim N_T +$

$\dim R_T = \dim V$. The **nullity** of T is defined to be the dimension of N_T and the **rank** of T is defined to be the dimension of R_T, thus

$$nullity(T) + rank(T) = \dim V$$

(Theorem 2).

Given an $m \times n$ matrix A over F, the **column space** of A is the subspace of $F^{m\times 1}$ spanned by the columns of A. The dimension of the column space of A is called the **column rank** of A. Consider the linear transformation

$$T_A : F^{n\times 1} \to F^{m\times 1}$$

of example (2), Session 29: $T_A(X) = A \cdot X$ for all $X \in F^{n\times 1}$. Then the null space of T_A is the solution space of the homogeneous system $AX = 0$, and the range space of T_A is the set of all $Y \in F^{m\times 1}$ such that $AX = Y$ is solvable. Since any such Y can be written as

$$Y = x_1 A_1 + \cdots + x_n A_n$$

with scalars $x_i \in F$ and A_j being the j^{th} column of A, the column space of A is equal the range of T_A. Thus, the column rank of A equals the rank of the linear transformation T_A. Suppose that E is the row–reduced echelon matrix which is row–equivalent to A. Then $AX = 0$ and $EX = 0$ have precisely the same solution space. If r denotes the number of nonzero rows of E, then the solution space of $EX = 0$ has dimension r (Example 15 on page 42); hence, the linear transformation T_A has nullity $n-r$. By Theorem 2,

$$n = nullity(T_A) + rank(T_A) = (n-r) + (column\ rank(A))$$

which implies that the column rank of A is r, and hence the row rank and the column rank of A are equal (Theorem 3).

Session 31, Wednesday, November 5. Upon request, Exercises 1, 2, 4, and 7 on page 73 were worked. Continued lecture on §3.2. Let V and W be vector spaces over the field F and let $L(V,W)$ denote the set of all linear transformations from V to W. Given $S, T \in L(V,W)$ and $c \in F$, define $S+T : V \to W$ and $cT : V \to W$ by $(S+T)(\alpha) = S(\alpha) + T(\alpha)$ and $(cT)(\alpha) = c \cdot T(\alpha)$ for all $\alpha \in V$. Then $L(V,W)$ is a vector space over F (Theorem 4). If V has finite dimension n and W has finite dimension m, then $L(V,W)$ has dimension mn (Theorem 6); indeed, if $\{\alpha_1, \ldots, \alpha_n\}$ is a basis for V and $\{\beta_1, \ldots, \beta_m\}$ is a basis for W, then the set

$$\overline{\mathcal{B}} = \{E^{p,q} \,|\, p = 1, \ldots, m \text{ and } q = 1, \ldots, n\}$$

is a basis for $L(V,W)$ where, for $p \in \{1, \ldots, m\}$ and $q \in \{1, \ldots, n\}$, $E^{p,q}$ is the (according to Theorem 1) unique linear transformation from V to W such that

$$E^{p,q}(\alpha_j) = \begin{cases} 0 & \text{if} \quad j \neq q \\ \beta_p & \text{if} \quad j = q \end{cases}$$

or, more compactly,

$$E^{p,q}(\alpha_j) = \delta_{jq}\beta_p$$

for each $j = 1, \ldots, n$. Proof next time.

Session 32, Friday, November 7. Continue lecture on §3.2. Proof of Theorem 5. If V, W, Z are vector spaces over F and $T : V \to W$ and $S : W \to Z$ are linear transformations, then the composition $S \circ T : V \to Z$ is a linear transformations (Theorem 6). Thus, the vector space $L(V, V) = \mathcal{L}(\mathcal{V})$ has a multiplication defined, namely composition of functions (the text writes ST for $S \circ T$), which is associative (i.e. $(ST)U = S(TU)$ for all $S, T, U \in \mathcal{L}(\mathcal{V})$), has a multiplicative identity I (i.e. $IT = TI = T$ for all $T \in \mathcal{L}(\mathcal{V})$), namely $I : V \to V$ is the identity function, multiplication is right and left distributive over addition (i.e. $(S + T)U = SU + TU$ and $S(T + U) = ST + SU$ for all $S, T, U \in \mathcal{L}(\mathcal{V})$), and has the property that $c(ST) = (cS)T = S(cT)$ for all $S, T \in \mathcal{L}(\mathcal{V})$ and all $c \in F$. In other words, in addition to being a vector space over F, $\mathcal{L}(\mathcal{V})$ is a ring (noncommutative if V has dimension 2 or larger) with multiplicative identity, and scalars from F in ring products can be associated with any factor. A structure of this type is called an F**–algebra with identity**. Suppose V has finite dimension n and $\mathcal{B}$ is an ordered basis for V consisting of $\alpha_1, \ldots, \alpha_n$, in this order. Let $T \in \mathcal{L}(\mathcal{V})$ and let $A_{pj} \in F$ such that $T(\alpha_j) = \sum_{p=1}^{n} A_{pj}\alpha_p$. From the proof of Theorem 5,

$$T = \sum_{p=1}^{n} \sum_{q=1}^{n} A_{pq} E^{p,q} ,$$

i.e. the $n \times n$ matrix A the j^{th} column of which contains the coordinate matrix of $T(\alpha_j)$ relative to the ordered basis $\mathcal{B}$ has as (p, q) entry the coefficient of $E^{p,q}$ needed to write T as a linear combination of the n^2 elements of the basis $\overline{\mathcal{B}}$. Note Example 10. If $S \in \mathcal{L}(\mathcal{V})$ is another linear transformation and $B_{pj} \in F$ such that $S(\alpha_j) = \sum_{p=1}^{n} B_{pj}\alpha_p$, then

$$S = \sum_{p=1}^{n} \sum_{q=1}^{n} B_{pq} E^{p,q} .$$

One verifies that $E^{p,q}E^{r,s} = 0$ if $r \neq q$ and $E^{p,q}E^{q,s} = E^{p,s}$. From this it follows that

$$TS = \sum_{p=1}^{n} \sum_{s=1}^{n} (AB)_{ps} E^{p,s} .$$

Thus, if A and B are the $n \times n$ matrices with with (p, q) entry the coefficient of $E^{p,q}$ when T and S are written as a linear combination in the basis $\overline{\mathcal{B}}$, respectively, then, writing TS as a linear combination in the basis $\overline{\mathcal{B}}$, the coefficient of $E^{p,q}$ is the (p, q) entry $(AB)_{pq}$ of the product matrix AB.

Session 33, Monday, November 10. Continue lecture on §3.2 with preview of §3.3. A linear transformation $T : V \to W$ is **invertible** if there exists a function $g : W \to V$ such that $g \circ T = I_V$, the identity function on V, and $T \circ g = I_W$, the identity function on W. It is known that a function is invertible iff it is bijective, i.e. both one–to–one and onto, and that the inverse function, g, is unique. For this reason, the inverse g of T

is denoted by $T^{-1} : W \to V$. Note that T^{-1} is invertible (its inverse is T), and T^{-1} is again a linear transformation (Theorem 7). An invertible linear transformation is called an **isomorphism**. Given vector spaces V and W over a field F, one says W is **isomorphic to** V if there exists an isomorphism $T : V \to W$. By Theorem 7, if W is isomorphic to V, then V is isomorphic to W. Also, since the composition of two linear transformations is a linear transformation and since the composition of two bijections is a bijection, the composition of two isomorphisms is an isomorphism. Thus, being isomorphic is an equivalence relation on the class of all vector spaces over F.

A linear transformation $T : V \to W$ is said to be **nonsingular** if $\alpha \in V$ and $T(\alpha) = 0_W$ imply $\alpha = 0_V$, i.e. if the null space $N_T = \{0_V\}$. A linear transformation $T : V \to W$ is nonsingular iff the function T is one–to–one, i.e. $T(\alpha) = T(\alpha')$ implies $\alpha = \alpha'$. Multiplication by x is a nonsingular linear operator on the vector space of all polynomials over the real numbers which is not an isomorphism. This can happen only in the case of infinite dimension: If V and W have the same finite dimension n and $T : V \to W$ is a linear transformation, then T is nonsingular iff T is one–to–one iff T is onto iff T is invertible (Theorem 9). This follows from the formula

$$\dim N_T + \dim R_T = n$$

when $\dim V = n$ is finite. The set G of all invertible linear operators on a vector space V is a **group** under the operation of composition of functions; in fact, G is the group of units of the linear F–algebra $L(V, V)$. If $\dim V \geq 2$, then G is noncommutative, i.e. there exist invertible $T, T' \in G$ such that $TT' \neq T'T$. Indeed, if there exists a basis $\mathcal{B}$ for V with two or more vectors, pick $\alpha_1, \alpha_2 \in \mathcal{B}$. There exist linear operators $T, T' : V \to V$ such that $T(\alpha_1) = \alpha_1 + \alpha_2$ and $T(\alpha) = \alpha$ for all $\alpha \in \mathcal{B}$ with $\alpha \neq \alpha_1$, and $T'(\alpha_2) = \alpha_1 + \alpha_2$ and $T(\alpha) = \alpha$ for all $\alpha \in \mathcal{B}$ with $\alpha \neq \alpha_2$ (Theorem 1). One verifies T, T' are invertible and $TT'(\alpha_1) = \alpha_1 + \alpha_2 \neq 2\alpha_1 + \alpha_2 = T'T(\alpha_1)$

Session 34, Wednesday, November 12. Lecture on §3.3. Every vector space of finite dimension n over the field F is isomorphic to F^n (Theorem 10). Indeed, if the ordered basis $\mathcal{B}$ of V consists of the vectors $\alpha_1, \ldots, \alpha_n$, in that order, then the map $T : V \to F^{n\times1}$ defined by $T(\alpha) = [\alpha]_{\mathcal{B}}$ for all $\alpha \in V$ is an isomorphism. It should be evident that the vector spaces $F^{n\times1}$ and F^n are isomorphic. Thus V is isomorphic to F^n (the composition of two linear transformations is a linear transformation, and the composition of two functions each of which is one–to–one and onto is again one–to–one and onto).

Lecture on §3.4: Representation of Linear Transformations by Matrices. Let V and W be finite dimensional vector spaces over F of dimensions n and m, respectively, let $\mathcal{B}$ be an ordered basis for V consisting of the vectors $\alpha_1, \ldots, \alpha_n$, in that order, let $\mathcal{B}'$ be an ordered basis for W consisting of the vectors $\beta_1, \ldots, \beta_m$, in that order, and let $T : V \to W$ be a linear transformation. Then there exists a unique matrix $A = A_T \in F^{m\times n}$ such that, for all $\alpha \in V$,

$$[T(\alpha)]_{\mathcal{B}'} = A \cdot [\alpha]_{\mathcal{B}}\,.$$

For $j = 1, \ldots, n$, the j^{th} column of A is $[T(\alpha_j)]_{\mathcal{B}'}$; moreover, defining

$$f : L(V, W) \to F^{m\times n}$$

by $f(T) = A_T$ for all $T \in L(V, W)$ is an isomorphism (Theorems 11 and 12). Given finite dimensional vector spaces V, W, Z over F with bases $\mathcal{B}, \mathcal{B}', \mathcal{B}''$, respectively, and linear transformations $T : V \to W$ and $U : W \to Z$, the matrix representing the composition map $UT : V \to Z$ relative to the bases $\mathcal{B}$ and $\mathcal{B}''$ is equal to the product matrix $B_U A_T$ where A_T is the matrix representing T relative to the bases $\mathcal{B}$ and $\mathcal{B}'$ and B_U is the matrix representing U relative to the bases $\mathcal{B}'$ and $\mathcal{B}''$ (Theorem 13).

Session 35, Friday, November 14. Finish §3.4: The special case when $V = W$ and $\mathcal{B} = \mathcal{B}'$. Let V be an n–dimensional vector space over F with ordered basis $\mathcal{B}$ consisting of the vectors $\alpha_1, \ldots, \alpha_n$. Then, for every linear transformation $T : V \to V$, there exists a unique matrix $A \in F^{n\times n}$ such that, for every $\alpha \in V$, $[T(\alpha)]_{\mathcal{B}} = A[\alpha]_{\mathcal{B}}$. The matrix A is denoted by $[T]_{\mathcal{B}}$ and said to be the **matrix representing T relative to the ordered basis $\mathcal{B}$**. See Example 14. From [HK, page 52, Theorem 7], if $\mathcal{B}'$ is another ordered basis for V consisting of vectors $\alpha'_1, \ldots, \alpha'_n$, then there exists a unique matrix $P \in F^{n\times n}$ (which is necessarily invertible and contains in its j^{th} column the coordinate matrix of α'_j relative to $\mathcal{B}$: $P_j = [\alpha'_j]_{\mathcal{B}}$) such that, for every $\alpha \in V$,

$$[\alpha]_{\mathcal{B}} = P[\alpha]_{\mathcal{B}'}\,.$$

Thus, if $T \in L(V, V)$, then, for all $\alpha \in V$,

$$P[T(\alpha)]_{\mathcal{B}'} = [T(\alpha)]_{\mathcal{B}} = [T]_{\mathcal{B}}[\alpha]_{\mathcal{B}} = [T]_{\mathcal{B}}P[\alpha]_{\mathcal{B}'}\,.$$

Hence,

$$[T(\alpha)]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P[\alpha]_{\mathcal{B}'}$$

for all $\alpha \in V$. The uniqueness of the matrix representing T relative to the ordered basis $\mathcal{B}'$ implies that $[T]_{\mathcal{B}'} = P^{-1}[T]_{\mathcal{B}}P$ (Theorem 14). See Example 16. Given two $n \times n$ matrices A over F, B is said to be **similar** to A if there exists an invertible $n \times n$ matrix P over F such that $B = P^{-1}AP$. Every $A \in F^{n\times n}$ is similar to itself ($A = I^{-1}AI$), if B is similar to A, then A is similar to B ($B = P^{-1}AP$ implies $A = PBP^{-1} = (P^{-1})^{-1}B(P^{-1})$), and if B is similar to A and C is similar to B, then C is similar to A ($B = P^{-1}AP$ and $C = Q^{-1}BQ$ imply $C = Q^{-1}P^{-1}APQ = (PQ)^{-1}A(PQ)$). Thus, similarity is an equivalence relation on $F^{n\times n}$.

We shall omit §3.5 and §3.6. The material covered there is used in the proof of Theorem 3 in §4.2; an alternate proof of this result will be provided.

4. Chapter 4. Polynomials

Session 36, Monday, November 17. Lecture on §4.1. Let F be a field. A **linear F–algebra** (or linear algebra over F) is a vector space $\mathcal{A}$ over F with an additional operation called multiplication of vectors associating with every pair $\alpha, \beta \in \mathcal{A}$ a vector $\alpha\beta \in \mathcal{A}$ so that the following conditions are satisfied:

(a) $\alpha(\beta\gamma) = (\alpha\beta)\gamma$ for all $\alpha, \beta, \gamma \in \mathcal{A}$.

(b) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ and $(\alpha + \beta)\gamma = \alpha\gamma + \beta\gamma$ for all $\alpha, \beta, \gamma \in \mathcal{A}$.

(c) $c(\alpha\beta) = (c\alpha)\beta = \alpha(c\beta)$ for all $c \in F$ and all $\alpha, \beta \in \mathcal{A}$.

If there exists an element $1 \in \mathcal{A}$ such that $1\alpha = \alpha 1 = \alpha$ for all $\alpha \in \mathcal{A}$, then $\mathcal{A}$ is said to be a **linear algebra with identity** and 1 is said to be the (multiplicative) **identity** of $\mathcal{A}$. If $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathcal{A}$, then $\mathcal{A}$ is said to be a **commutative** F–algebra.

Examples of F–algebras with identity are $L(V, V)$ when V is a vector space over F ($L(V, V)$ is noncommutative if V has dimension two or larger) and $F^{n\times n}$ (also noncommutative when $n \geq 2$).

The linear algebra of all **formal power series over** F, denoted F^∞, is defined as follows. Let $\mathbb{N}$ denote the set $\{0, 1, 2, \dots\}$ of all non–negative integers. Define

$$F^\infty = \{f \mid f : \mathbb{N} \to F \text{ is a function}\}.$$

If $f \in F^\infty$ and $n \geq 0$ is an integer, then $f(n)$ will be denoted by f_n. Thus, each $f \in F^\infty$ can be thought of as an infinite sequence and is written as $f = (f_0, f_1, \dots, f_n, f_{n+1}, \dots)$. By Example 3, page 30, F^∞ is a vector space over F when, for $f, g \in F^\infty$ and $c \in F$, one defines $f + g, cf : \mathbb{N} \to F$ by $(f + g)(n) = f(n) + g(n)$ and $(cf)(n) = cf(n)$ for all $n \in \mathbb{N}$. In sequence notation this means that, if $f = (f_0, f_1, f_2, \dots)$ and $g = (g_0, g_1, g_2, \dots)$, then $f + g = (f_0 + g_0, f_1 + g_1, f_2 + g_2, \dots)$ and $cf = (cf_0, cf_1, cf_2, \dots)$. Multiplication is defined as follows: given $f, g \in F^\infty$, fg is the function with $(fg)(n) = (fg)_n$ given by

$$(fg)_n = f_0 g_n + f_1 g_{n-1} + \cdots + f_n g_0 = \sum_{i=0}^{n} f_i g_{n-i}$$

for all integers $n \geq 0$. In order to show that multiplication is associative, let $f, g, h \in F^\infty$ and $n \in \mathbb{N}$. Then

$$[f(gh)]_n = \sum_{i=0}^{n} f_i (gh)_{n-i} = \sum_{i=0}^{n} f_i \sum_{j=0}^{n-i} g_j h_{n-i-j}\,.$$

In comparison,

$$[(fg)h]_n = \sum_{i=0}^{n} (fg)_i h_{n-i} = \sum_{i=0}^{n} \sum_{j=0}^{i} f_j g_{i-j} h_{n-i}$$

which equals

$$\sum_{i=0}^{n}(f_0g_ih_{n-i} + \cdots + f_kg_{i-k}h_{n-i} + \cdots + f_ig_0h_{n-i})\,.$$

Note that for each $i = 0, 1, \ldots n$, $f_{i+1}, \ldots, f_n$ are not present in the i^{th} term of this sum. Factoring out the f_k, $k = 0, \ldots, n$, one obtains

$$[(fg)h]_n = f_0(\sum_{i=0}^{n} g_ih_{n-i}) + \cdots + f_k(\sum_{i=k}^{n} g_{i-k}h_{n-i}) + \cdots + f_n(\sum_{i=n}^{n} g_{i-n}h_{n-i})\,,$$

so

$$[(fg)h]_n = \sum_{k=0}^{n} f_k \sum_{i=k}^{n} g_{i-k}h_{n-i}\,.$$

Define $j = i - k$. Then, as i ranges from k to n, j ranges from 0 to $n - k$, and $n - i = n - k - j$. Thus,

$$[(fg)h]_n = \sum_{k=0}^{n} f_k \sum_{j=0}^{n-k} g_jh_{n-k-j} = \sum_{k=0}^{n} f_k(gh)_{n-k} = [f(gh)]_n\,.$$

Thus, $\mathcal{A} = \mathcal{F}^\infty$ satisfies condition (a) of the definition of a linear algebra. Homework problems will ask you to verify conditions (b) and (c). Note that $(fg)_n = f_0g_n + \cdots + f_ng_0 = g_0f_n + \cdots + g_nf_0 = (gf)_n$ for all $f, g \in F^\infty$ and all $n \geq 0$, which implies that F^∞ is commutative.

Define $1 = (1, 0, 0, \ldots) \in F^\infty$, i.e. $1_n = \delta_{0n}$ for all $n \geq 0$. Then $(1f)n = 1 \cdot f_n = f_n$ for all n proving $1f = f$ and F^∞ is an algebra with identity 1.

For polynomials, a special symbol is usually set aside, like x or X. The book chooses x to denote the function $(0, 1, 0, 0, \ldots) \in F^\infty$, i.e. $x_n = \delta_{1n}$ for all $n \geq 0$. Define $x^0 = 1$, the identity of F^∞, define $x^1 = x$, and $x^2 = x \cdot x$, etc. We proved by induction on n that for all $n \geq 0$, $(x^n)_k = \delta_{nk}$. Define $F[x]$ to be the subspace of F^∞ spanned by the set $\{1 = x^0, x, x^2, \ldots, x^n, x^{n+1}, \ldots\}$ of all powers of x. The elements of $F[x]$ are said to be **polynomials** over F.

Session 37, Wednesday, November 19. Lecture on §4.2. Let $\mathcal{A}$ be a linear F–algebra with identity $1_{\mathcal{A}}$. For each $\alpha \in \mathcal{A}$, define $\alpha^0 = 1_{\mathcal{A}}$, $\alpha^1 = \alpha$, and $\alpha^k = \alpha \cdot \ldots \cdot \alpha$, the product of k copies of α, when k is a positive integer. Let $\alpha \in \mathcal{A}$, let m, n be non–negative integers, and let $c_i, d_j \in F$. The left and right distributive laws (i.e. condition (b) in the definition of a linear algebra, page 117) imply that

$$(\sum_{i=0}^{m} c_i\alpha^i)(\sum_{j=0}^{n} d_j\alpha^j) = \sum_{i=0}^{m}\sum_{j=0}^{n}(c_i\alpha^i)(d_j\alpha^j)\,.$$

By condition (c) of the definition on page 117, this equals

$$\sum_{i=0}^{m}\sum_{j=0}^{n} c_id_j\alpha^i\alpha^j = \sum_{i=0}^{m}\sum_{j=0}^{n} c_id_j\alpha^{i+j} = \sum_{t=0}^{m+n}(\sum_{i+j=t} c_id_j)\alpha^t\,.$$

If we define $c_i = 0$ for $i = m+1, \dots, m+n$ and $d_j = 0$ for $j = n+1, \dots, m+n$, then the latter sum can be re–written and we have the following equality

$$(\sum_{i=0}^{m} c_i\alpha^i)(\sum_{j=0}^{n} d_j\alpha^j) = \sum_{t=0}^{m+n}(\sum_{i=0}^{t} c_i d_{t-i})\alpha^t \,.$$

Let $f \in F[x]$. By definition of $F[x]$, there exist an integer $n \geq 0$ and scalars $c_0, \dots, c_n \in F$ such that $f = c_0x^0 + c_1x^1 + \cdots + c_nx^n$. Since each x^i is the sequence with 1_F in coordinate i and zeros elsewhere,

$$f = \sum_{i=0}^{n} c_i x^i = (c_0, c_1, \dots, c_n, 0, 0, \dots)\,,$$

i.e. $f_i = f(i) = c_i$ for $i = 0, \dots, n$ and $f_i = f(i) = 0$ for $i > n$. It follows that the set $S = \{1 = x^0, x, x^2, \dots\}$ of all powers of x is linearly independent. Also, every $f \in F[x]$ has the property that the sequence $(f_0, f_1, \dots)$ contains a nonzero entry in at most finitely many coordinates.

Given $f \in F[x]$, if $f \neq 0$ then the **degree of** f, denoted $\deg f$, is n if $f_n \neq 0$ and $f_i = 0$ for all $i > n$. A **scalar polynomial** is a polynomial of the form cx^0. The zero polynomial has no degree defined. The nonzero scalar polynomials are exactly the polynomials of degree zero. A polynomial f of degree n is **monic** if $f_n = 1$. If $f, g \in F[x]$ are nonzero, then (i) $fg \neq 0$, (ii) $\deg(fg) = \deg f + \deg g$, (iii) f and g monic implies fg monic, (iv) fg is scalar iff both f and g are scalar, and (v) if $f + g \neq 0$, then $\deg(f + g) \leq \max(\deg f, \deg g)$ (Theorem 1, page 120). The set $F[x]$ of all polynomials over the field F is a commutative linear F–algebra with identity, $1_{F[x]} = 1_F x^0$ the monic scalar polynomial (Corollary 1).

Session 38, Friday, November 21. Continue Lecture on §4.2. Let $\mathcal{A}$ be a linear algebra over F with identity and let $f \in F$. If $f = \sum_{i=0}^{n} f_i x^i$ and $\alpha \in \mathcal{A}$, define

$$f(\alpha) = \sum_{i=0}^{n} f_i\alpha^i \,.$$

Then, for all $f, g \in F[x]$, all $c \in F$, and all $\alpha \in \mathcal{A}$, (i) $(cf + g)(\alpha) = cf(\alpha) + g(\alpha)$; and (ii) $(fg)(\alpha) = f(\alpha)g(\alpha)$ (Theorem 2). This theorem can be interpreted as follows. Fix $\alpha \in \mathcal{A}$, and define a function $E_\alpha : F[x] \to \mathcal{A}$ by $E_\alpha(f) = f(\alpha)$ for all $f \in F[x]$. (The function E_α is said to be the **evaluation map at** α.) Then E_α is a linear transformation. Morover, E_α preserves multiplication in the sense that $E_\alpha(fg) = E_\alpha(f)E_\alpha(g)$ for all $f, g \in F[x]$. Such a function between F–algebras is said to be an **algebra homomorphism**.

Session 39, Monday, November 24. Lecture on §4.4. Given polynomials f and d over the field F with $d \neq 0$, there exist unique polynomials $q, r \in F[x]$ such that (i) $f = qd+r$, and (ii) either $r = 0$ or $\deg r < \deg d$ (Theorem 4). In class, we omitted the proof that q and r are unique. Here it is: if $f = q_1d + r_1 = q_2d + r_2$ with $q_i, r_i \in F[x]$ and $r_i = 0$ or of degree less than $\deg d$ for $i = 1, 2$, then $(q_1 - q_2)d = r_2 - r_1$. If $r_2 - r_1 \neq 0$, then $q_1 - q_2 \neq 0$, and $\deg(q_1 - q_2)d \geq \deg d$ [HK, page 120, Theorem 1(ii)]. But $r_2 - r_1 \neq 0$ implies $\deg(r_2 - r_1) < d$ [HK, page 120, Theorem 1(v)], a contradiction. Thus, $r_1 = r_2$ and [HK, page 121, Corollary 2] implies $q_1 = q_2$.

A polynomial $d \in F[x]$ is said to **divide** the polynomial f over F if there exists a polynomial $q \in F[x]$ such that $f = qd$. Note that a field F is a commutative linear algebra over F with identity $1 = 1_F$. Thus, for every $c \in F$ and every $f \in F[x]$, $f(c)$ is defined; if $f(c) = 0$, then c is said to be a **root** of f. Given a polynomial f over F and a scalar $c \in F$, c is a root of f iff the polynomial $x - c \in F[x]$ divides f (Corollary 1). A nonzero polynomial $f \in F[x]$ of degree n has at most n roots in F (Corollary 2).

Session 40, Monday, December 1. On request, Exerces 4, 6, and 7 on page 123 were worked.

Lecture on §4.3, page 125f. By [HK, page 30, Example 3], the set F^F of all functions $\phi : F \to F$ from F to F is a vector space over F under the operations of pointwise addition and scalar multiplication. Define, for all $\phi, \psi \in F^F$, $\phi\psi : F \to F$ by $(\phi\psi)(\alpha) = \phi(\alpha)\psi(\alpha)$ for all $\alpha \in F$ (note that $\phi\psi$ is the pointwise multiplication and **not** the composition of functions!) and verify that F^F is a commutative linear algebra over F with identity as defined on page 117; the identity is the constant function $1 : F \to F$ defined by $1(\alpha) = 1_F$ for all $\alpha \in F$. A function $\phi : F \to F$ is said to be a **polynomial function on** F if there exists a polynomial $f \in F[x]$ such that $\phi(\alpha) = f(\alpha)$ for all $\alpha \in F$. If $\phi(\alpha) = f(\alpha)$ for all $\alpha \in F$, then the text writes $\phi = f^\sim$. Given $f, g \in F[x]$ and a scalar $c \in F$, then by Theorem 2, $((cf + g)^\sim)(\alpha) = (cf + g)(\alpha) = cf(\alpha) + g(\alpha) = (cf^\sim + g^\sim)(\alpha)$ and $((fg)^\sim)(\alpha) = (fg)(\alpha) = f(\alpha)g(\alpha) = f^\sim(\alpha)g^\sim(\alpha) = (f^\sim g^\sim)(\alpha)$ for all $\alpha \in F$. Thus, $(cf+g)^\sim = cf^\sim + g^\sim$ and $(fg)^\sim = f^\sim g^\sim$. It follows that the set $\mathcal{P}$ of all polynomial functions on F is a subspace of F^F which is closed under the multiplication defined on F^F and, hence, is a linear algebra over F on its own right. It also follows that the association $f \mapsto f^\sim$, $f \in F[x]$, is a linear transformations from $F[x]$ to the space $\mathcal{P}$ of all polynomial functions on F and that this association preserves products. If the field F contains infinitely many elements, then the association $f \mapsto f^\sim$, $f \in F[x]$, is an isomorphism from the algebra $F[x]$ of all polynomials over F to the algebra of polynomial functions on F according to the definition on page 125 (Theorem 3, page 126). Indeed, the only item missing in what has been proved so far is to show that the linear transformation $f \mapsto f^\sim$, $f \in F[x]$, is one–to–one which is equivalent to its nullspace being $\{0_{F[x]}\}$. Let $p \in F[x]$ such that $p^\sim = 0$. Then $p^\sim(\alpha) = p(\alpha) = 0$ for every $\alpha \in F$. Thus, when F is infinite, p has infinitely many roots. The only polynomial over F with infinitely many roots is the zero polynomial [HK, page 129, Corollary 2]. Thus $p = 0$ proving $f \mapsto f^\sim$, $f \in F[x]$, is nonsingular and, hence, one–to–one.

If F is a field containing only n elements for some (finite) integer n, then $F[x]$ and the algebra of polynomial functions on F can never be isomorphic since $F[x]$ is always infinite and there are only finitely many (namely n^n)

functions from F to F. Example: for the two–element field F of Exercise 5 on page 5 and $p = x + x^2 \in F[x]$, one has $p(\alpha) = 0$ for all $\alpha \in F$. Thus $p^\sim = 0^\sim = 0$ is the zero function on F proving the association $f \mapsto f^\sim$ is not one–to–one and hence not an isormorphism.

Session 41, Wednesday, December 3. Review for Test 3. On request, Exercises 7 (§3.4, page 95), 5 (§4.3, page 127), and Extra Problem 5 were worked on the board. In studying for Test 3 and also for the Final, you are encouraged to re–work Tests 1, 2, and 3 and make sure your solutions are correct (see the solutions on the WEB). Also, the theoretical aspects in the text and in class should be reviewed. For example: (i) If V is spanned by a finite set of m vectors in V, then any set of $m + 1$ or more vectors in V is linearly dependent. (ii) If S is a linearly independent set of vectors in V and the span of S is not V, then for any vector $\alpha \in V$ which does not belong to the span of S the set $S \cup \{\alpha\}$ is linearly independent. Using these facts from Chapter 2, you should be able to prove that (1) Every vector space V which has a finite spanning set has a finite basis, i.e. V is finite dimensional. In fact, if V is spanned by m vectors, every basis of V contains at most m vectors. (2) Any two bases of a finite dimensional vector space contain the same (finite) number of elements (this justifies the definition of dimension) (3) If V has finite dimension n, then a) every linearly independent set of vectors in V contains at most n vectors; b) every spanning set contains at least n vectors; any linearly independent set of n vectors in V is a basis; and d) every set of n vectors in V that span V is a basis (if the n vectors spanning V were not linearly independent, one could be omitted and V would be spanned by $n - 1$ vectors resulting in $\dim V < n$ – a contradiction).

Session 42, Friday, December 5. Test 3. Topics: Chapter 3, Section 3.1 through 3.4 only; Chapter 4, Sections 4.1, 4.2, 4.4 through Corollary 2 only, and 4.3, pages 125, 126 only: definition of an algebra isomorphism and Theorem 3.

The End