ECHELON FORM a matrix is in echelon form if 1. all nonzero rows are above any rows of all 0's, 2. each leading entry of a row is in a col to the right of the leading entry of the row above it, 3. all entries in a col below a leading entry are zeroes REDUCED ECHELON FORM a matrix is in reduced echelon form if it is in echelon form, and the leading entry of each nonzero row is 1, and each leading 1 is the only nonzero entry in its column PIVOT POSITION a pivot position is a location that corresponds to a leading 1 in the reduced echelon form. a pivot column is a column that contains a pivot position ALGEBRAIC PROPERTIES OF RN (i) u+v = v+u (ii) (u+v)+w=u+(v+w) (iii) u+0 = 0+u = u (iv) u + (-u) = -u + u = 0 (v) c(u+v) = cu + cv (vi) (c+d)u = cu + du (vii) c(du) = (cd)u (viii) 1u = u THEOREM if Ap=b, then the solutions to Ax=b are of the form p + v, where v is a solution Ax=0 THEOREM the columns of A are linearly independent iff Ax=0 has only the trivial solution LINEAR TRANSFORMATION a transformation T is linear if T(u+v) = T(u) + T(v) and T(cu) = cT(u) TRANSPOSE RULES $(A^T)^T = A$, $(A+B)^T = A^T + B^T$, $(rA)^T = rA^T$, $(AB)^T = B^T A^T$ TWO BY TWO INVERSE inverse of [[a b][c d]] is 1/(ad-bc) [[d -b][-c a]] INVERSE RULES $(AB)^{-1} = B^{-1}A^{-1}$, $(A^T)^{-1} = (A^{-1})^T$ ELEMENTARY MATRIX the result of a row operation on A can be written as EA where E is an invertible matrix created by performing the same row operation on I. FINDING INVERSE row reduce [A I] to get [I A^{-1}] BASIS must be linearly independent INVERTIBLE MATRIX THEOREM the following statements are equivalent: (a) A is an invertible matrix (b) A is row equivalent to I (c) A has n pivot positions (d) Ax=0 has only the trivial solution (e) columns of A form a linearly independent set (f) x $\mapsto$ Ax is one-to-one (g) Ax=b has at least one solution for each b (h) columns of A span R^n (i) x $\mapsto$ Ax maps R^n to itself (j) there is an nxn matrix C such that CA = I (k) there is an nxn matrix D such that AD=I (l) A^T is an invertible matrix (m) columns of A form a basis of R^n (n) Col A = R^n (o) dim Col A = rank A = n (p) Nul A = {0} (q) 0 is not an eigenvalue of A (r) det(A) $\neq$ 0 (s) (Col A)$^\perp$ = {0} (t) (Nul A)$^\perp$ = R^n (u) Row $A = R^n$ (v) A has n nonzero singular values COLUMN ROW EXPANSION calculate $AB = col_1(A)row_1(B) + ... + col_n(A)row_n(B)$ LU FACTORIZATION write A=LU where A is mxn, L is mxm lower triangular with 1's on diagonal, U is mxn echelon form. solve Ax=b by solving Ly = b then Ux=y. LU FACTORIZATION ALGORITHM reduce A to echelon form U, using only row replacements that add a multiple of one row to another row below it. Place entries in L such that the same sequence of row operations reduces L to I. SUBSPACE a subspace is a set containing 0 which is closed under addition and scalar multiplication COLUMN SPACE the column space of A is the set Col A of all linear combinations of the columns of A NULL SPACE the null space of A is the set Nul A of all solutions to Ax=0 THEOREM the pivot columns of A form a basis of Col A RANK of a matrix A, rank A, is the dimension of Col A RANK THEOREM if A has n columns, rank A + dim Nul A = n DETERMINANT AND ROW OPERATIONS if a multiple of one row of A is added to another row to produce B, det B = det A. if two rows of A are interchanged to produce B, then det B = -det A. If one row of A is multiplied by k to produce B, then det B = k det A DETERMINANT TRANSPOSE PROPERTY det A^T = det A MULTIPLICATIVE PROPERTY det AB = det A det B DETERMINANT LINEARITY the determinant is a linear function of each of the columns of A (multilinear) TRIANGULAR MATRIX DETERMINANT is the product of the entries on the main diagonal CRAMERS RULE if A is invertible then sol'n of Ax=b is $x_i = detA_i(b)/detA$ where $A_i(b)$ is the matrix obtained from A by replacing column i with b COFACTOR the ij cofactor of A, C_{ij}, is $(-1)^{i+j}detA_{ij}$ where A_{ij} is the matrix A with row i and column j deleted ADJUGATE adj A is a matrix B where $b_{ij} = C_{ji}$. $A^{-1} = adjA/detA$. VOLUME UNDER TRANSFORMATION if $T : R^2 \rightarrow R^2$ is linear given by matrix A and $S \subset R^2$ then {area of T(S)} = $|detA|$ {area of S}. If instead $T : R^3 \rightarrow R^3$ and $S \subset R^3$, then {volume of T(S)} = $|detA|$ {volume of S} VECTOR SPACE is a nonempty set V of objects, closed under addition and scalar multiplication, satisfying u+v=v+u, (u+v)+w=u+(v+w),with a 0 so u + 0 = u, with a -u so that u + (-u) = 0, and where c(u+v) = cu + cv, (c+d)u = cu + du, c(du) = (cd)u, 1u = u SPANNING SET THEOREM if S = $\{v_1, ..., v_p\} \subset V$ and H = Span(S), then if v_k is a linear combination of the other vectors in S, Span($S - \{v_k\}$) = H. If H $\neq$ {0}, some subset of S is a basis for H UNIQUE REPRESENTATION THEOREM if B is a basis for V, then for each x, there are unique $c_1, ..., c_n$ s.t. x=$c_1 b_1 + ... + c_n b_n$. COORDINATES the c's from the uniq. rep. thm. are called the coordinates of x in the basis B. the column vector $[x]_B = (c_1, ..., c_n)$ CHANGE OF COORDINATES MATRIX for basis B of R^n, the matrix $P_B = [b_1 ... b_n]$ is the change of coordinates matrix. $x = P_B[x]_B$, and $P_B^{-1}x = [x]_B$ where P_B is invertible because its columns span R^n THEOREM if two matrices A and B are row equivalent then their row spaces are the same. If B is in echelon form its nonzero rows span it. GENERAL CHANGE OF COORDINATES MATRIX if C and B are bases of V, then there is a unique nxn matrix $P_{C \leftarrow B}$ such that $[x]_C = P_{C \leftarrow B}[x]_B$, and $P_{C \leftarrow B} = [[b_1]_C ... [b_n]_C]$. Also, $P_{C \leftarrow B}^{-1} = P_{B \leftarrow C}$ FINDING EIGENVALUES λ is an eigenvalue of A iff det(A - λI) = 0 (the characteristic polynomial) THEOREM the eigenvalues of a triangular matrix are the entries on its main diagonal THEOREM if A and B are similar by row operations, then they have the same characteristic polynomial and eigenvalues DIAGONALIZATION is writing $A = PDP^{-1}$ where D is a diagonal matrix whose (nonzero) entries along the diagonal are the eigenvalues of A, and the columns of P are the corresponding eigenvectors in order. DIAGONALIZATION THEOREM an nxn matrix A is diagonalizable iff A has n linearly independent eigenvectors (it isn't necessarily invertible, e.g. if any eigenvalue is 0) THEOREM an nxn matrix with n distinct eigenvalues is diagonalizable. this is sufficient but not necessary for it to be diagonalizable EIGENVALUE MULTIPLICITY the multiplicity of λ is the multiplicity of the root λ of the polynomial $det(A - I\lambda) = 0$ THEOREM if A is nxn with eigenvalues $\{\lambda_k\}$ then 1. dimension of eigenspace for λ_k is $\leq$ the multiplicity of λ_k, 2. A is diagonalizable iff the sum of dimensions of eigenspaces is n, 3. if A is diagonalizable and B_k is a basis for the eigenspace for λ_k then $\cup\{B_k\}$ forms an eigenvector basis for R^n MATRIX OF TRANSFORMATION RELATIVE TO BASES B AND C is $M = [[T(b_1)]_C ... [T(b_n)]_C]$ so that $[T(x)]_C = M[x]_B$. If B=C then M is the B-matrix for T. DIAGONAL MATRIX REPRESENTATION if $A = PDP^{-1}$ and B is the basis for R^n formed from the columns of P, then D is the B-matrix for the transformation $x \mapsto Ax$ THEOREM let A be a real 2x2 matrix with complex eigenvalue $\lambda = a - bi(b \neq 0)$ and associated eigenvector v. Then $A = PCP^{-1}$ where $P =$ [Re v Im v] and $C = [[a\ -b][b\ a]]$ PYTHAGOREAN THEOREM $u \cdot v = 0$ iff $||u+v||^2 = ||u||^2 + ||v||^2$ ORTHOGONAL COMPLEMENT of vector space W is the vector space $W^\perp$ of vectors orthogonal to (every vector in) W THEOREM $(RowA)^\perp = NulA$ by the row-column rule for Ax, and thus obviously $(ColA)^\perp = NulA^T$ ANGLE $u \cdot v = ||u||||v||cos\theta$ ORTHOGONAL SET of vectors $\{u_i\}$ is one where $i \neq j \rightarrow i \cdot j = 0$ ORTHOGONAL BASIS is a basis that is an orthogonal set THEOREM if $\{u_i\}$ is an orthogonal basis and $y = c_1 u_1 + ... + c_p u_p$ then $c_j = y \cdot u_j / u_j \cdot u_j$ ORTHOGONAL PROJECTION if $y = \hat{y} + z$ where $\hat{y} = \alpha u$ and $z \cdot u = 0$ then $\hat{y} = proj_u y = (u \cdot y / u \cdot u)u$ ORTHONORMAL SET is an orthogonal set of unit vectors. If they are a basis then they are an orthonormal basis THEOREM an mxn matrix U has orthonormal columns iff $U^T U = I$ THEOREM if U is mxn with orthonormal columns, then $||Ux|| = ||x||$ and $(Ux) \cdot (Uy) = x \cdot y$ and $(Ux) \cdot (Uy) = 0$ iff $x \cdot y = 0$ ORTHOGONAL MATRIX is a square matrix U with orthonormal columns, which makes it invertible, and means it has orthonormal rows, too. ORTHOGONAL DECOMPOSITION THEOREM for a given W, y can be written uniquely as $y = \hat{y} + z$ where $\hat{y}$ is in W and z is in $W^\perp$. $\hat{y} = proj_W(y) = (y \cdot u_1/u_1 \cdot u_1)u_1 + ... + (y \cdot u_p/u_p \cdot u_p)u_p$ where $\{u_i\}$ is an orthogonal basis of W and $z = y - \hat{y}$. BEST APPROXIMATION THEOREM if $\hat{y} = proj_W(y)$ and $v \in W$, $v \neq y$, then $||y - \hat{y}|| < ||y - v||$ THEOREM if U = $[u_1 ... u_p]$, where the columns are an orthonormal basis for W, then $proj_W(y) = (y \cdot u_1)u_1 + ... + (y \cdot u_p)u_p = UU^T y$ GRAM SCHMIDT PROCESS given a basis $\{x_i\}$ for W, i=1 to p, define $v_1 = x_1$, $v_{i+1} = x_{i+1} - proj_{W_i} x_{i+1}$, where $W_i = Span(v_1, ..., v_i)$. Then $\{v_i\}$ for i=1 to p is an orthogonal basis for W. In addition $Span(v_1, ..., v_k) = Span(x_1, ..., x_k)$ for each k=1 to p. Note: $proj_{W_i} x_{i+1} = \sum_{j=1}^{i}(x_{i+1} \cdot v_j / v_j \cdot v_j)v_j$ QR FACTORIZATION if A is an mxn matrix with linearly independent columns, then A=QR where Q is mxn whose columns form an orthonormal basis of Col A and R is an nxn upper triangular invertible matrix with positive entries on its diagonal. Q can be found by the Gram Schmidt process. R can be found by solving $x_k = r_{1k}v_1 + ... + r_{kk}v_k$ since x_k is in $Span(v_1, ..., v_k)$ LEAST SQUARES SOLUTION of Ax=b, where A is mxn, is $\hat{x}$ such that $||b - A\hat{x}|| \leq ||b - Ax||$ for all x. To solve it let $\hat{b} = proj_{ColA} b$ and $\hat{x}$ such that $A\hat{x} = \hat{b}$. THEOREM the set of least-squares solutions of Ax=b coincides with the nonempty set of solutions of the normal equations $A^T Ax = A^T b$ THEOREM the matrix $A^T A$ is invertible iff the columns of A are linearly independent. In this case, $\hat{x} = (A^T A)^{-1} A^T b$ is the only least-squares solution of Ax=b THEOREM if A has linearly independent columns and A=QR is a QR factorization, then the least squares solution to Ax=b is $\hat{x} = R^{-1} Q^T b$ INNER PRODUCT on a vector space V by definition satisfies 1. $\langle u, v \rangle = \langle v, u \rangle$ 2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ 3. $\langle cu, v \rangle = c\langle u, v \rangle$ 4. $\langle u, u \rangle \geq 0$ with $\langle u, u \rangle = 0$ iff $u = 0$. INNER PRODUCT SPACE is a vector space with an inner product. The Gram Schmidt process can be used to find an orthonormal basis for any inner product space. CAUCHY-SCHWARTZ INEQUALITY for all u,v in V, $|\langle u, v \rangle| \leq ||u||||v||$ TRIANGLE INEQUALITY for all u,v, in V, $||u + v|| \leq ||u|| + ||v||$ A POLYNOMIAL INNER PRODUCT $\langle p, q \rangle = p(t_0)q(t_0) + ... + p(t_n)q(t_n)$ A CONTINUOUS FUNCTION INNER PRODUCT $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ WEIGHTED LEAST SQUARES if the diagonal of W is $w_1, ..., w_n$ and we seek to minimize the weighted error $w_1^2(y_1 - \hat{y}_1^2) + ... + w_n^2(y_n - \hat{y}_n^2)$ then we seek a least-squares solution of WAx=Wy TREND ANALYSIS to separate a linear or quadratic trend from a function f sampled at n data points, use the inner product $\langle p, q \rangle = p(t_0)q(t_0) + ... + p(t_n)q(t_n)$, find an orthogonal basis $p_0, ..., p_3$ for the third degree polynomial space P_3 using Gram-Schmidt on $1, t, t^2, t^3$, and use that basis to orthogonally project f onto P_3. The coefficients of the projection $\hat{g} = c_0 p_0 + ... + c_3 p_3$ in the orthogonal basis are called trend coefficients, where c_1 represents the linear trend, c_2 represents the quadratic trend, and so on. THEOREM if A is symmetric then eigenvectors from different eigenspaces are orthogonal ORTHOGONALLY DIAGONALIZABLE a matrix A is orthogonally diagonalizable if there is an orthogonal matrix P where $A = PDP^T = PDP^{-1}$ THEOREM nxn matrix A is orthogonally diagonalizable iff A is symmetric SPECTRAL THEOREM FOR SYMMETRIC MATRICES if A is symmetric, 1. A has n real eigenvalues counting multiplicities, 2. dimension of each eigenspace equals multiplicity of eigenvalue, 3. eigenspaces are mutually orthogonal, 4. A is orthogonally diagonalizable SPECTRAL DECOMPOSITION of A is $A = \lambda_1 u_1 u_1^T + ... + \lambda_n u_n u_n^T$ where $u_1 ... u_n$ are mutually orthonormal eigenvectors. $u_i u_i^T$ is a projection matrix in the sense that $u_i u_i^T x = proj_{ui} x$ QUADRATIC FORM is a function $R^n \rightarrow R$ given by $Q(x) = x^T Ax$ for some matrix A, which is called the matrix of the quadratic form. CHANGE OF VARIABLE IN QUADRATIC FORM of vector x is an equation of the form $x = Py$ where P is an invertible matrix. Thus $x^T Ax = y^T(P^T AP)y$. If P orthogonally diagonalizes A then $P^T AP = P^{-1}AP$ is diagonal. PRINCIPAL AXES THEOREM if A is a symmetric matrix then there is an orthogonal change of variable, $x = Py$, that transforms $x^T Ax$ into the quadratic form $y^T Dy$ with no cross-product terms CLASSIFYING QUADRATIC FORMS a quadratic form Q is positive definite if $Q(x) > 0$ for $x \neq 0$, negative definite if $Q(x) < 0$ for $x \neq 0$, and indefinite if Q assumes both positive and negative values QUADRATIC FORMS AND EIGENVALUES the form $x^T Ax$ is positive definite iff the eigenvalues are all positive, negative definite iff the eigenvalues are all negative, and indefinite iff A has both positive and negative eigenvalues CONSTRAINED OPTIMIZATION let $m = min\{x^T Ax : ||x|| = 1\}, M = max\{x^T Ax : ||x|| = 1\}$. Then M is the greatest eigenvalue of A and m is the least eigenvalue of A. The value of $x^T Ax$ is M or m when x is a corresponding unit eigenvector u_1 or u_p. THEOREM the maximum value of $x^T Ax$ subject to $x^T x = 1$,$x^T u_1 = 0$ is the second greatest eigenvalue λ_2, attained when $x = u_2$ THEOREM the maximum value of $x^T Ax$ subject to $x^T x = 1$,$x^T u_1 = 0$,...,$x^T u_{k-1} = 0$ is the kth greatest eigenvalue λ_k, attained when $x = u_k$ SINGULAR VALUES of an mxn matrix A are the square roots of the eigenvalues of $A^T A$ denoted $\sigma_1 > \sigma_2 > ... > \sigma_n$. The singular values are the lengths of the vectors $Av_1, ..., Av_n$ where $v_1, ..., v_n$ is an orthonormal basis for R^n consisting of eigenvectors of $A^T A$ THEOREM if $\{v_1, ..., v_n\}$ is an orthonormal basis of R^n consisting of eigenvectors of $A^T A$, where A has r nonzero singular values, then $\{Av_1, ..., Av_r\}$ is an orthogonal basis for Col A, and rank $A = r$ SINGULAR VALUE DECOMPOSITION let A be mxn with rank r. Then there exists an mxn matrix $\Sigma = [[D0][00]]$ where D is an rxr diagonal matrix whose entries are the first r singular values of A, and an mxn orthogonal matrix U, and an nxn orthogonal matrix V, such that $A = U\Sigma V^T$. The columns of V are an orthogonal basis $v_1, ..., v_n$ for R^n consisting of eigenvectors of $A^T A$, and the columns of U are $u_i = (Av_i)/||Av_i|| = (Av_i)/\sigma_i$. THEOREM the vectors $u_{r+1}, ..., u_m$ form an orthonormal basis for Nul A^T and the eigenvectors $v_{r+1}, ..., v_m$ form an orthonormal basis for Nul A. CONDITION NUMBER of an invertible matrix is the ratio σ_1/σ_n of its largest and smallest singular values COVARIANCE MATRIX if $X_1 + ... + X_n = 0$ (i.e. the mean has been subtracted so the X's are in mean-deviation form) and $X = [X_1 ... X_n]$ is the pxN observation matrix, then $S = 1/(N-1)XX^T$ is the covariance matrix PRINCIPAL COMPONENT ANALYSIS is an orthogonal change of variable $X = PY$ such that the covariance matrix for Y, namely $P^T SP$, is diagonal. This is accomplished by orthogonally diagonalizing $S = PDP^T$ to get $P^T SP = D$ where $P = [u_1 ... u_p]$ is the change of variables. The u_i are called the principal components.