DEFVecSp : $1.x + y = y + x2.(x + y) + z = x + (y + z)3.Add.Id.0 \in V|0 + x = x4.\exists Add.Inv.(-x)|x+(-x) = 0(5.)F.Dis.Lawa(x+y) = ax+ay(6.)S.Dis.Law(a+b)(x) = ax + bx(7.)Mul.Id.1x = x(8.)(ab)x = a(bx)$ **THM1.1.13** If W is a subset of a vector space V s.t. $\mathbf{x}, \mathbf{y} \in$ @ and for any $a, b \in \mathbb{F}$ the vector $a\mathbf{x} + b\mathbf{y} \in W$, then W is a subspace of V. **DEFLinHull** of $S\langle S\rangle$, smallest subspace of V that contains S,i.e. intersection of all subspaces of V that contain S. **THM1.2.6Span**(S) $= \langle S\rangle$. **DEF** $\bigoplus$ Where W_1, W_2 are subspaces of V, then $W_1 + W_2 = W_1 \bigoplus W_2$ if $W_1 \cap W_2 = 0$. **DEFComplementarysubspaces**W_1 and W_2 if $V = W_1 \bigoplus W_2$ **THMReplacement**: V is a vector space spanned by $S = s_1, \cdots, s_m$. If $T = t_1, \cdots, t_n$ is a L.I. subset of V, then $\le m$ and $\exists S' \subset S$ having $m - n$ elements such that $T \cup S'$ spans V. **THMExtension**: W is a subspace of VIf $T = t_1, \cdots, t_n$ and $S = s_1, \cdots, s_m$ span W and V, respectivley, then $\exists S' \subset S$ having $m - n$ elements such that $T \cup S'$ is a basis for V. **DEFQuotientSpaces**: W subspace of V. The set $x + W|x \in V(orequivalently[[x]]|x \in V)$ of all cosets of W in V is denoted V/W and is called the quotient of V modulo W. **DEF** $\boxplus$ $\boxdot$:Let W be a subspace of V. Define operations $\boxplus : V/W \times V/W \to V/W$ and $\boxdot : \mathbb{F} \times V/W \to V/W$ given by (i) $(x + W) \boxplus (y + W) = (x + y) + W$ and $a \boxdot (x + W) = (ax) + W$. These are the operations of vector addition and scalar multiplication on V/W. **CHAP2** **DEFLineartransformation** Let V and V be vector spaces over $\mathbb{F}$. A map $L : V \to W$ is a linear transformation from V into W if $L(ax_1 + bx_2) = aL(x_1) + bL(x_2)$ for $x_1, x_2 \in V$ and $a, b \in \mathbb{F}$ **COR2.1.17** A linear transforamtion is invertible if and only if it is bijective. **Prop.2.1.24** : If $V \cong W$ are isomorphic vector spaces, with isoorphism $L : V \to W$, then: (i) A linear equation holds in V iff it also holds in W: that is $\sum_{i=1} a_i\mathbf{x_i} = \mathbf{0}$ holds in V iff $\sum_{i=1} a_iL_i\mathbf{x_i} = \mathbf{0}$ holds in W. (ii) A set B $= \{\mathbf{v_i}, \ldots, \mathbf{v_n}\}$ is a basis of V iff LB $= \{L\mathbf{v_i}, \ldots, L\mathbf{v_n}\}$ is a basis for W. Moreover, the dimension of V is equal to the dimension of V. (iii) The subspaces of V are in vijective correspondence with the subspaces of W. (iv) If K: W $\to$ U is any linear transformation, then the composition KL:V$\to$ U is also a linear transformation and we have $\mathscr{N}(KL) = L^{-1}\mathscr{N}(K) = \{v|L(\mathbf{v}) \in \mathscr{N}(K)\}$ and $\mathscr{R}(KL) = \mathscr{R}(K)$ **THMF.Iso**. If V and X are vector spaces and $L : V \to X$ is a linear transformation, then $V/\mathscr{N}(L) \cong \mathscr{R}(L)$. in particular, if L is surjective, then $V/N(L) \cong X$. **THM2.2.7** If V is a finite-dimensional vector space and W is a subspace of V, then $\dim(V) = \dim(W) + \dim(V/W)$ **THMRank − Nullity** Let V and W be finite-dimensional vector spaces. If $L : V \to$ is a linear transformation then $\dim(V) = \dim\mathscr{R}(L) + \dim\mathscr{N}(L) = \operatorname{rank}(L) + \operatorname{nullity}(L)$. **CORSec.Iso.Thm**. Assume V_1 and V_2 are subspaces of V. Then $V_1/(V_1 \cap V_2) \cong (V_1 + V_2)/V_2$. **CORDim.Formula** If V_1 and V_2 are finite-dimensional subspaces of a vector space V, then $\dim(V_1)+\dim(V_2) = \dim(V_1\cap V_2)+\dim(V_1+V_2)$ **DEFSimilarMatrices** Two square matrices $A, B \in M_n(\mathbb{F})$ are said to be similar if there exists a nonsingular $P \in M_n(\mathbb{F})$ such that $B = P^-1AP$. **DEFBernsteinPolynomials** Given $n \in \mathbb{N}_{\ge 0}J$, the Bernstein polynomials $B^n_j(x)^n_{j=0}$ of degree n are defined as $B^n_j(x) = \binom{n}{j}x^j(1 - x)^{n-j}$, where $\binom{n}{j} = \frac{n!}{j!(n-j)!}$ **LEM2.5.3**For $j = 0, 1, \ldots, n$ $B^n_j(x) = \sum_{i=j}^n(-1)^{i-j}\binom{n}{i}\binom{i}{j}x^i$ **THM2.5.4** For any $n \in \mathbb{N}$, the set T_n of degree n Bernstein polynomials $T_n = B^n_j(x)^n_{j=0}$ forms a basis for $\mathbb{F}[x]^n$ **DEFTrace** The trace is the sum of the elements along the main diagonal **PROP2.6.2**All of the elementary matrices are invertible. **DEFRowEquivalence**The B is said to be row equivalent to the matrix A if there exists a finite collection of elementary matrices $E_1, E_2, \ldots, E_n$ such that $B = E_1E_2 \ldots E_n$ **DEFREF** A is REF if (i) leading coefficient of each row is strictly to the right of the previous row's leading coefficient (ii) All nonzero rows are above any zero rows and **RREF** if (iii) the leading coefficient of every row is 1 (iv) The leading coefficient of every row is the only nonzero entry in its column. **DEFPermutation** Different arrangements of a set. Even if it has an even number of inversions, odd if an odd number of inversions. Sign is 1 if even, -1 if odd. **DEFInversion** A pair $(\sigma(i), \sigma(j))$ such that $i < j$ and $\sigma(i) > \sigma(j)$. **THM2.8.7** If $A, B \in M_n(\mathbb{F})$, then $\det(AB) = \det(A)\det(B)$ **COR2.8.8**$\det(A^{-1} = (\det(A))^{-1})$ **Cramer'sRule** If $A \in M_n(\mathbb{F})$ is nonsingular, then the unique solution to $Ax = b$ is $x = A^{-1}b = \frac{\operatorname{adj}(A)}{\det(A)}b$. Moreover, if $A_i(b) \in M_n(\mathbb{F})$ is the matrix A with the i-th column replaced by b, then the i-th coordinate of x is $x_i = \frac{\det(A_i(b))}{\det(A)}$ **Exam2** DEF**innerproduct** on V for $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$, $a, b \in \underline{\mathbb{F} : (i)\langle x, x\rangle \ge 0, eq.iff\mathbf{x} = 0(ii)}$ $\langle x, a\mathbf{y} + b\mathbf{z}\rangle = a\langle x, y\rangle + b\langle x, z\rangle(iii)\langle x, y\rangle = \overline{\langle y, x\rangle}$PROP:(i)$\langle \mathbf{x} + \mathbf{y}, \mathbf{z}\rangle = \langle \mathbf{x}, \mathbf{z}\rangle + \langle \mathbf{y}, \mathbf{z}\rangle$(ii)$\langle a\mathbf{x}, \mathbf{y}\rangle = \bar{a}\langle \mathbf{x}, \mathbf{y}\rangle$FunFact: $0 \le \langle x-\lambda y, x-\lambda y\rangle = \langle x, x\rangle-\langle \lambda y, x\rangle-\langle x, \lambda y\rangle+\langle \lambda y, \lambda y\rangle = \langle x, x\rangle-\bar{\lambda}\langle y, x\rangle - \lambda\langle x, y\rangle + \lambda\bar{\lambda}\langle y, y\rangle$Frobenius inner product: $\langle A, B\rangle = \operatorname{tr}(A^HB)$Orthogonal if $\langle \mathbf{y}, \mathbf{x}\rangle = 0$.Cauchy-Schwarz: $|\langle \mathbf{x}, \mathbf{y}\rangle| \le \|\mathbf{x}\|\|\mathbf{y}\|$PROOF:suppose $\mathbf{u}, \mathbf{v} \ne 0$, choose$\lambda = \frac{|\langle \mathbf{v}, \mathbf{u}\rangle|}{\langle \mathbf{v}, \mathbf{u}\rangle}$, and $|\lambda| = 1$,Thus, $0 \le \|\frac{\lambda\mathbf{u}}{\|\mathbf{u}\|} - \frac{\mathbf{v}}{\|\mathbf{v}\|}\|^2 = |\lambda|^2 - 2\Re(\langle \frac{\lambda\mathbf{v}}{\mathbf{v}}, \frac{\mathbf{u}}{\|\mathbf{u}\|}\rangle) + 1 = 2 - 2\frac{\lambda\langle \mathbf{v}, \mathbf{u}\rangle}{\|\mathbf{vu}\|} \implies |\langle \mathbf{u}, \mathbf{v}\rangle = \lambda\|\langle \mathbf{v}, \mathbf{u}\rangle\| \le \|\mathbf{u}\|\|\mathbf{v}\|$ **PR. 3.1.23**: If $\mathbf{u}$ is unit, $\operatorname{proj}_{\mathbf{u}} : V \to V$ is a linear operator and: (i) $\operatorname{proj}_{\mathbf{u}} \circ \operatorname{proj}_{\mathbf{u}} = \operatorname{proj}_{\mathbf{u}}$ (ii)$\mathbf{r} = \mathbf{v} - \operatorname{proj}_{\mathbf{u}}(\mathbf{v})$ is orthogonal to $\operatorname{span}\{\mathbf{u}\}$ (iii) $\operatorname{proj}_{\mathbf{u}}(\mathbf{v})$ unique vector in $\operatorname{span}\{\mathbf{u}\}$ nearest $\mathbf{v}$. Angle:

$\cos(\theta) = \frac{\langle \mathbf{x},\mathbf{y}\rangle}{\|\mathbf{x}\|\|\mathbf{y}\|}$ Pyth: if x,y are orthogonal $\|\mathbf{x}+\mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$ **TH. 3.2.3**: If $(V, \langle .,.\rangle)$ and $\{\mathbf{x}_i\}_{i=1}^m$ is a finite orthonormal set: (i) If $x = \sum_{i=1}^m a_i \mathbf{x}_i$, then $\langle \mathbf{x}_i, \mathbf{x}\rangle = a_i$ for all $i = 1, 2, \dots, m$ (ii) If $x = \sum_{i=1}^m a_i\mathbf{x}_i$ and $y = \sum_{i=1}^m b_i\mathbf{y}_i$, then $\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{i=1}^m \overline{a}_i b_i$ (iii) If $x = \sum_{i=1}^m a_i\mathbf{x}_i$, then $\|x\|^2 = \sum_{i=1}^m |a_i|^2$. linear map is orthonormal if for every x,y $\langle \mathbf{x}, \mathbf{y}\rangle_V = \langle \mathbf{Lx}, \mathbf{Ly}\rangle_W$ Proj onto unit: $\text{proj}_{\mathbf{u}}(\mathbf{x}) = \langle \mathbf{u}, \mathbf{x}\rangle\mathbf{u}$ THM: Q, Q_1 orthonormal: (i) $\|Q\mathbf{x}\| = \|\mathbf{x}\|(ii)QQ_1$ is orthonormal$(iii)Q^{-1} = Q^H$is orthonormal$(iv)Q^H A = QQ^H = I(v)$columns are orthonormal$(vi)|det(Q)| = 1$ **DEF**: Bessel's Inequality: If $(V, \langle .,.\rangle)$ and $\{\mathbf{x}_i\}_{i=1}^m$ is a finite orthonormal set: $\|\mathbf{v}\| \geq \sum_{i=1}^m |\langle \mathbf{x}_i, \mathbf{v}\rangle|^2 = \|\text{proj}_X \mathbf{v}\|^2$. **DEF**: Pythagorean Theorem: If $(V, \langle .,.\rangle)$ and $\{\mathbf{x}_i\}_{i=1}^m \subset V$ is a finite orthonormal subset with span X, for every $\mathbf{v} \in V$: $\|\mathbf{v}\|^2 = \|\text{proj}_X(\mathbf{v})\|^2 + \|\mathbf{v} - \text{proj}_X(\mathbf{v})\|^2 = \sum_{i=1}^m |\langle \mathbf{x}_i, \mathbf{v}\rangle|^2 + \|\mathbf{v} - \sum_{i=1}^m \langle \mathbf{x}_i, \mathbf{v}\rangle\mathbf{x}_i\|^2$ Gram-Schmidt: Given $(x_1 \dots x_n)$ to get orthonormal basis $(q_1 \dots q_n)$ 1) $q_1 = \frac{x}{\|x\|}$ 2) $p_1 = \text{proj}_{q_1}(x_2) = \langle \mathbf{q_1}, \mathbf{x_2}\rangle q_1$, and $q_2 = \frac{x_2 - p_1}{\|x_2 - p_1\|}$ Repeat, $p_{n-1} = \langle \mathbf{q_1}, \mathbf{x_n}\rangle q_1 \cdots + \langle \mathbf{q_{n-1}}, \mathbf{x_n}\rangle q_{n-1}$, and $q_n = \frac{x_n - p_{n-1}}{\|x_n - p_{n-1}}$ QR: $A = QR$ where Q is orthonormal columns of A, calculated through Gram-Schmidt. Note, $A = QR \implies Q^H A = R$. So calculate R. HyperPlane: Defined as perpendicular to a particular v, thus $proj_Y(x) = x - \langle \mathbf{v}, \mathbf{x}\rangle\mathbf{v}$ and the transformation is given by $I - 2\frac{vv^H}{v^H v}$ NORMS: (i) Positivity $\|x\| \geq 0$ with equality if and only if $x = 0$ (ii) Scale preservation $\|ax\| = |a|\|x\|$ (iii) Triangle inequality (Follows from Cauchy Schwarz) $\|x+y\| \leq \|x\| + \|y\|$ Every inner-product has norm $\|x\| = sqrt\langle \mathbf{x}, \mathbf{x}\rangle$ Norms: $\|x\|_p = (\sum |x|^p)^{1/p}$ $\|A\|_F = \sqrt{\text{tr}(A^H A)}$ Induced Norm on Linear Transformations: $\|T\|_{V,W} = sup_{\|x\|_V=1}\|Tx\|_W$ THM:If $T \in \mathscr{B}(X,Y)$, $S \in \mathscr{B}(Y,Z)$ then $ST \in \mathscr{B}(X,Z)$ and $\|ST\|_{X,Z} \leq \|S\|_{Y,Z}\|T\|_{X,Y}$ Remark: for $n \geq 1$ we have $\|T^n\| \leq \|T\|^n$. If $\|T\| \leq 1$ then $\|T\|^n$ approaches 0. EX: $\|A\|_p = sup_{x\neq 0}\frac{\|Ax\|_p}{\|x\|_p}, 1 \leq p \leq \infty$ Young's: $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ if $\frac{1}{p} + \frac{1}{q} = 1$ Arithmatic Geotmetric mean: $a^\theta b^{1-\theta} \leq \theta a + (1-\theta)b$ for $0 \leq \theta \leq 1$. Holder's: if $\frac{1}{p} + \frac{1}{q} = 1$ where $1 \leq p \leq \infty$ then, $\sum |xy| \leq (\sum |x|^p)^{1/p}(\sum |y|^q)^{1/q} = \|x\|_p\|y\|_q$ (Note, $p = q = 2$ implies Cauchy Swarz) Minkowski: $\|x+y\|_p \leq \|x\|_p + \|y\|_p$ Finite Dimensional Riesz Thm: Let $L: V \to \mathbb{F}$, $\exists! y \in V s.t. L(x) = \langle \mathbf{y}, \mathbf{x}\rangle \forall x \in V$, and $\|L\| = \|y\| = \sqrt{\langle \mathbf{y}, \mathbf{y}\rangle}$ Adjoint: Adjoint of L is a linear transformation s.t. $\langle \mathbf{w}, \mathbf{Lv}\rangle_W = \langle \mathbf{L^*w}, \mathbf{v}\rangle_v$ $\forall v \in V, w \in W$. THM: Let $L: V \to W$ be finite, adjoint L^* exists and is unique. Proof: Let $L_w: V \to \mathbb{F}$ be defined by $L_w(v) = \langle \mathbf{w}, \mathbf{L(v)}\rangle_W$. By Riesz, $\exists! u \in V s.t. L_w(v) = \langle \mathbf{u}, \mathbf{v}\rangle_V \forall V$. Let $L^*: W \to V$ be $L^*(w) = u$. Thus, $\langle \mathbf{w}, \mathbf{L(v)}\rangle_W = \langle \mathbf{L^*(w)}, \mathbf{v}\rangle_V \forall v \in V, w \in W$. Show linearity and uniqueness. Prop 3.7.12 $(S+T)^* = S^* + T^*$ and $(\alpha T)^* = \bar{\alpha}T^*$ OrthComplement of $S \subset V$ is the set $S^\perp = \{y \in V | \langle \mathbf{x}, \mathbf{y}\rangle = 0, \forall x \in S\}$, Note $S^\perp$ is a subset of V. If W is finite dim. subspace of V, then $V = W \oplus W^\perp$. Fund. Subspaces: $\mathscr{R}(L)^\perp = \mathscr{N}(L^*)$ and $\mathscr{N}(L)^\perp = \mathscr{R}(L^*)$ COR: Let V,W be finite, $L: V \to W$ $V = \mathscr{N}(L) \oplus \mathscr{R}(L^*)$ and $W = \mathscr{R}(L) \oplus \mathscr{N}(L^*)$. Least squares: $\hat{\mathbf{x}} = (A^H A)^{-1}A^H\mathbf{b}$ is unique minimizer. Semi-Spectral mapping: If λ_i are the eigenvalues of a semisimple matrix $A \in M_n(\mathbb{F})$, and if f(x) is any polynomial, then $\{f(\lambda_i\}$ are the eigenvalues of $f(A)$. ExcerCh3: For reals, $\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$ and $\|x\|^2 + \|y\|^2 = \frac{1}{2}(\|x+y\|^2 + \|x-y\|^2)$ CH4: Eigenvalues and eigenvectors depend only on the linear transformation and not the choice of basis of it. Let C_{TS} be the transition matrix, $A \in S, B \in T$ s.t. $[x]_T = C[x]_S$ and $B = CAC^{-1}$. Thus, $B[x]_T = CAC^{-1}C[x]_S = C\lambda[x]_S = \lambda C[X]_S = \lambda[X]_T$ THM: Following are equivalent, (i) λ is an eigenvalue (ii) There is a nonzero x such that $(\lambda I - A)x = 0$ (iii) $\Sigma_\lambda(A) \neq \{\mathbf{0}\}$ (iv) $\lambda I - A$ is is singular, thus det = 0. Prop: If A, B are similar, that is $A = PBP^{-1}$. They are the same operator, and (i) have the same charcatistic poly, eigenvalues, the eigenbases are isomorphic. Invariant: A subspace is invariant if for $L: V \to V$, $L(W) \subset W$. A is simple if eigenvalues are distinct, semi-simple if eigenbasis spans A. Diagonalizable iff semisimple. $A = PDP^{-1}$. P is eigenvectors, D is eigenvalues. Fibonnaci Numbers, Make the matrix, calculate eigens, note 300th number is $v_{301} = Av_{300} = A^{300}v_0 = P^{-1}D^{300}Pv_0$ where v_0 is starting vector of sequence, of course, round to nearest integer. Power Method: pick vector, muliply by A, normalize, iterate. Implies dominant eigenvector and value if semi-simple. Rayleigh quotient does the same thing, but faster. $\frac{\langle \mathbf{x}, \mathbf{Ax}\rangle}{\|x\|^2}$ Implies λ and eigenvector, which converge to *some* eigenvalue and vector. Orthonormally similar if $B = U^H AU$, U is an orthonormal iff $\langle \mathbf{Ux}, \mathbf{Uy}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle = x^H y$. Hermitian matricies are linear operators that preserve length and angle on different bases. Lem: If A is Hermitian, ortho similar to B, then B is Hermitian. Schur's Lemma: Every nxn matrix, A, is orthonormally similar to an upper triangular matrix. Proved by induction. Spct Thm I: Every Hermitian matrix A is orthonormally similar to a real diagonal matrix. Proof: By Schur's, A is ortho to an upper triangular, T. Since A is hermitian, so is T, and $T^H = T$ thus all eigenvalues(diagonals) are real. Normal Matrix: Spct Thm holds for all Normals. Normal when $A^H = AA^H$. Spct Thm II: A matrix A is normal iff it is orthonormally similar to a diagnoal matrix, equivalently, if it has an orthonormal eigenbasis. Proof: by Shur's T is uppertriangular Other direction, just multiply out. Positive Definite: Has positive eigen values,there exists a unique lower triangular matrix L, with real and strictly positive diagonal elements such that $M = LL^*$, that's cholesky decomp. It is invertible, and it's inverse is positive definite. Sum and product of semi definites are semi definite. Q^TMQ is positive semi definite. Determinant is bounded by product of diagonal elements. $A = S^HS$, and if A is definite, S is nonsingular. SVD: if A is of rank r, $\exists$orthonormalU, V and real valued $\Sigma = diag(\sigma_1, \dots, \sigma_r, \dots, 0, 0,)$ s.t. $A = U\Sigma V^H$ where σ_i is positive real valued. Σ is unique. To calculate, $\Sigma = \sqrt{\lambda_i}$ of A^HA. V is construction of eigenvectors of A^HA, with unfilled spots filled by gramschmidt (if eigenbasis doesn't exist). U is determined by $u_i = \frac{1}{\sigma}Av_i$ where u_i, v_i are columns of U, V respectively. Compact form is where we drop all zeros of sigma, and fit U, V accordingly. Moore-Penrose: $A^\dagger = V_1\Sigma_1^{-1}U_1^H$ of compact, or $V\Sigma^\dagger U^*$ where $\Sigma^\dagger$ is recipricol of nonzeros on diagonal, transpose. Schmidt-Eckart-Young: for A of rank r, and each s<r, $\sigma_{s+1} = inf_{rank(B)=s}\|A - B\|_2$ with minimizer $B^\diamond = \sum_{i=1}^{s}\sigma_iu_iv_i^H$ where each σ is the singular value of A, with corresponding u,v columns of U,V in the compact form of SVD. Scmidt-Eckart-Young-Mirsky: $(\sum_{j=s+1}^{r})^{1/2} = inf_{rank(B)=s)}\|A - B\|_F$ Cor: $A = U\Sigma V^H$ is the SVD where A has rank r >s then for any $mxn\Delta$ such that $A + \Delta$ has rank s we have: $\|\Delta\|_2 \geq \sigma_{s+1}$ and $\|\Delta\|_F \geq (\sum_{k=s+1}^{r}\sigma_k)^{1/2}$ and equality holds if $\Delta = -\sum_{i=s+1}^{r}\sigma_iu_iv_i^H$. The infimum of $\|\Delta\|$ rank$(I - A\Delta) < m$is i/σ_1 with minimizer $\Delta^* = \frac{1}{\sigma_1}v_1u_1^H$ Small gains: if $A \in M_n$, then $I - A\Delta$ is nonsingular, provided that $\|A\|_2\|\Delta\|_2 < 1$ For $A \in M_{mxn}(\mathbb{F})$: (i) $\|A\|_2 = \sigma_1$(largest signular value) (ii) if A is invertible, then $\|A^{-1}\|_2 = \frac{1}{\sigma_n}$ (iii) $\|A^H\|_2^2 = \|A^T\|_2^2 = \|A^HA\|_2 = \|A\|_2^2$ (iv) if U, V are orthonormal then $\|UAV\|$ (v) $\|UAV\|_2 = \|A\|_F$ (vi) $\|A\|_F = (\sigma_1^2 + \dots + \sigma_r^2)^{1/2}$. (vii) $|det(A)| = \Pi\sigma_i$. The SVD gives an orthonormal basis of four fundamental subspaces. first r columns of V are basis of $\mathscr{R}(A^H)$, the last n-r columns of V span $\mathscr{N}(A)$, first r columns of U span $\mathscr{R}(A)$ last m-r span $\mathscr{N}(A^H)$. Remark: Let V,W be normed linear spaces, if $L \in \mathscr{B}$ then the induced norm on L satisfies $\|Lx\|_W \leq \|L\|_{V,W}\|x\|_V$. Fun Fact: $0 \leq \langle \mathbf{x} - \lambda\mathbf{y}, \mathbf{x} - \lambda\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{x}\rangle - \langle \lambda\mathbf{y}, \mathbf{x}\rangle - \langle \mathbf{x}, \lambda\mathbf{y}\rangle + \langle \lambda\mathbf{y}, \lambda\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{x}\rangle - \hat{\lambda}\langle \mathbf{y}, \mathbf{x}\rangle - \lambda\langle \mathbf{x}, \mathbf{y}\rangle + \hat{\lambda}\lambda\langle \mathbf{y}, \mathbf{y}\rangle$. Polar Decomp: If $A \in M_{mxn}(\mathbb{F})$ and $m \geq n$. Then there exists a Q with orthonormal columns and positive semidefinite P such that $A = PQ$. Because $A = U\Sigma V^H = UV^HV\Sigma V^H$. Let $Q = UV^H$ and $P = V\Sigma V^H$ Norms: 1-Norm max sum of columns. ∞-Norm max sum of rows. 2-Norm is largest singular value. Semi-Simple: Diagonalizable, Eigenbasis exists, distinct eigenvalues iff simple. Semi Positive Definite: $\langle \mathbf{x}, \mathbf{Ax}\rangle \geq 0$, non-negative eigenvalues, chebesky decomp, sub-matrix positive, sort of kind of hermitian. Normal: Orthonormally similar to Diagonal matrix, $A^HA = AA^H$, Hermetians, skews, orthonormals... **CHAPTER 5**: **DEF: Metric** (i)Positive Definiteness $d(\mathbf{x}, \mathbf{y}) \geq 0$ ($=$ iff $\mathbf{x} = \mathbf{y}$) (ii)Symmetry: $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (iii)Triangle Inequality: $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + (\mathbf{z}, \mathbf{y})$. **DEF: Open Ball** with center at $\mathbf{x}_0$ radius $r > 0$ to be the set $B(\mathbf{x}_0, r) = \{\mathbf{x} \in X | d(\mathbf{x}, \mathbf{x}_0 < r\}$. **DEF: Neighborhood** of point $\mathbf{x}_0$ is a subset $E \subset X$ if $\exists$ an open ball $B(\mathbf{x}_0, r)$ such that $B(\mathbf{x}_0, r) \subset E$. Here $\mathbf{x}_0$ is an **interior point**. And E° is the **set of interior points** of E. **DEF: Open Set** if $E \subset X$ and every point $\mathbf{x} \in E$ is an interior point. **THRM 5.1.10:** If $\mathbf{x} \in B(\mathbf{x}_0, r)$ then $B(\mathbf{x}, r - \varepsilon) \subset B(\mathbf{x}_0, r)$ where $\varepsilon = d(\mathbf{x}_0, \mathbf{x})$. **THRM 5.1.12:** The union of any collection of open sets is open, and the intersection of any finite collection of open sets is open. **THRM 5.1.14:** (i)E° is open (ii) If G is an open subset of E, then $G \subset E^\circ$. (iii) E is open iff $E = E^\circ$ (iv) E° is the union of all open sets contained in E. **DEF: Continuous(1)** Let (X, d) and (Y, ρ) be metric spaces. A function $f : X \to Y$ is continuous at a point $\mathbf{x}_0 \in X$ if for all $\varepsilon > 0$, there exists $\delta > 0$ such that $\rho(f(\mathbf{x}), f(\mathbf{x}_0)) < \varepsilon$ whenever $d(\mathbf{x}, \mathbf{x}_0) < \delta$. The function f is continuous on a subset $E \subset X$ if it is continuous at all $\mathbf{x}_0 \in E$. The set of continuous functions from X to Y is denoted $C(X; Y)$. **THRM 5.1.18 Continuous(2):** A function $f : X \to Y$ is continuous on X iff the preimage $f^{-1}(U)$ of every open set $U \subset Y$ is open in X. **COR/PROP 5.1.19/20:** Composition, sum, product, and scalar multiples of continuous functions are continuous. **DEF: Limit Point** is a point $\mathbf{p} \in E$ of set $E \subset X$ if every neighborhood of $\mathbf{p}$ intersects $E \setminus \{\mathbf{p}\}$. **DEF: Isolated Point** is a point $\mathbf{p} \in E$ where $E \subset X$ and $\mathbf{p}$ is not a limit point. **DEF: Dense** We say that E is dense in X if every point in X is either in E or is a limit point of E. **DEF: Closed** We say that F is closed if it contains all of its limit points. **THRM 5.2.8:** If $\mathbf{p}$ is a limit point of $E \subset X$, then every neighborhood of $\mathbf{p}$ contains infinitely many points. **THRM 5.2.9:** A set $U \subset X$ is open iff its compliment U^c is closed. **COR 5.2.12:** Intersection of any collection of closed sets is closed, and the union of a finite collection of closed sets is closed. **COR 5.2.14:** A function $f : X \to Y$ is continuous iff for each closed set $F \subset Y$ the preimage $f^{-1}(F)$ is closed on X. **DEF 5.2.16:** The **closure** of E, denoted $\bar{E}$, is the set E together with its limit points. We define the **boundary** of E, denoted Bd E, as the closure minus the interior, that is, Bd $E = \bar{E} \setminus E^\circ$. **DEF Limit:** A function $f : X \to Y$ has a limit $\mathbf{y}_0 \in Y$ at $\mathbf{x}_0 \in X$ if for all $\varepsilon > 0$, there exists a $\delta > 0$ such that $\rho(f(\mathbf{x}), \mathbf{x}_0)) < \varepsilon$ wherever $0 < d(\mathbf{x}, \mathbf{x}_0) < \delta$. We denote the limit as $\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = \mathbf{y}_0$. **DEF Limit (sequence):** The limit of the sequence $(\mathbf{x}_i)_{i=0}^{\infty}$ if for all $\varepsilon > 0$, there exists $N > 0$

such that $d(\mathbf{x}, \mathbf{x}_n) < \varepsilon$ whenever $n \geq N$. We write $\mathbf{x}_k \to \mathbf{x}$ or $lim_{k\to\infty}\mathbf{x}_k = \mathbf{x}$, and say that the sequence converges to $\mathbf{x}$. **DEF Cluster Point:** We say that $\mathbf{x} \in X$ is a cluster point of the sequence $(\mathbf{x}_i)_{i=0}^{\infty}$ if for all $\varepsilon > 0$ and $N > 0$, there exists $n \geq N$ such that $d(\mathbf{x}, \mathbf{x}_n) < \varepsilon$. **PROP 5.2.30:** A convergent sequence has exactly one cluster point. In particular, if a sequence has a limit, it is unique. **THRM 5.2.32:** The function $f : X \to Y$ is continuous at a point $\mathbf{x}_* \in X$ iff for each sequence $(\mathbf{x}_i)_{i=0}^{\infty} \subset X$ that converges to $\mathbf{x}_*$, the sequence $(f(\mathbf{x}_k))_{k=0}^{\infty}$ converges to $f(\mathbf{x}_*) \in Y$. **DEF: Cauchy sequence** A sequence $(\mathbf{x}_i)_{i=0}^{\infty}$ in X is a Cauchy sequence if for all $\varepsilon > 0$ there exists an $N > 0$ such that $d(\mathbf{x}_m, \mathbf{x}_n) < \varepsilon$ whenever $m, n \geq N$. **PROP 5.3.4:** Any sequence that converges is a Cauchy sequence. **PROP 5.3.6:** Cauchy sequences are **bounded** (for all $\mathbf{x}$ $\exists$ $M \in \mathbb{N}$ such that for all $\mathbf{p}$ we have $d(\mathbf{x}, \mathbf{p}) < M$). **PROP 5.3.8:** Any Cauchy sequence which has a convergent subsequence is convergent. **DEF: Complete** A metric space (X, d) is complete if every Cauchy sequence converges. **THRM 5.3.13:** The fields $\mathbb{R}$ and $\mathbb{C}$ are complete with respect to the usual metric $d(z, w) = |z - w|$. **THRM 5.3.14:** If $\{(X_{,i}\,, d_i)\}_{i=0}^{n}$ is a finite collection of complete metric spaces, then the Cartesian product $X = X_1 \times X_2 \times \cdots \times X_n$ is complete when endowed with the p-metric for $1 \leq p \leq \infty$. **THRM 5.3.15:** Let $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{C}$. For every $n \in \mathbb{N}$ and $p \in [1, \infty]$, the linear space $\mathbb{F}^n$ with the norm $\|\cdot\|_p$ is complete. **DEF: Uniform Continuity** A function $f : X \to Y$ is uniformly continuous on $E \subset X$ if for all $\varepsilon > 0$ there exists $\delta > 0$ such that $\rho(f(\mathbf{x}), f(\mathbf{x})) < \varepsilon$ whenever $\mathbf{x}, \mathbf{x} \in E$ and $d(\mathbf{x}, \mathbf{x}) < \delta$. **PROP 5.3.21:** If $f : X \to Y$ is a bounded linear transformation of normed linear spaces, then f is uniformly continuous. **THRM 5.3.23:** Let (X, d) and (Y, ρ) be metric spaces and $f : X \to Y$ be uniformly continuous. If $(\mathbf{x}_i)_{i=0}^{\infty}$ is a Cauchy sequence, then so is $(f(\mathbf{x}_k))_{k=0}^{\infty}$. **COR 5.3.24:** Every finite-dimensional normed linear space is complete with respect to the metric induced by the norm. **LEM 5.3.25:** If Y is a dense subspace of of a normed linear space Z such that every Cauchy sequence in Y converges in Z, then Z is complete. **DEF: Open Cover** A collection $\{G_\alpha\}_{\alpha\in J}$ of open sets is an open cover of the set E if $E \subset \bigcup_{\alpha\in J} G_\alpha$. **DEF: Compact** A set E is compact if every open cover has a finite subcover, that is, for every open cover $\{G_\alpha\}_{\alpha\in J}$ there exists a finite subcollection $\{G_\alpha\}_{\alpha\in J'}$, where $J' \subset J$ is a finite subset, such that $E \subset \bigcup_{\alpha\in J'} G_\alpha$. **PROP 5.4.3:** A closed subset of a compact set is compact. **LEM 5.4.4:** Every infinite subset of a compact set K contains a limit point of K. **DEF: Bounded** We say that a subset K of a metric space X is bounded if there exists an $\mathbf{x} \in X$ and an $M > 0$ such that $K \subset B(\mathbf{x}, M)$. **THRM 5.4.6:** A compact subset of a metric space is closed and bounded. **Heine-Borel Theorem:** Consider $\mathbb{R}^n$ with the usual (Euclidean) metric. If a subset is closed and bounded, then it is compact. **PROP 5.4.9:** The continuous image of a compact set is compact; that is, if $f : X \to Y$ is continuous and $k \subset X$ is compact, then $f(K) \subset Y$ is compact. **Extreme-Valued Theorem:** Let (X, d) be a metric space. If $f : X \to \mathbb{R}$ is continuous and $K \subset X$ is a nonempty compact set, the $f(K)$ contains its infimum and supremum. **THRM 5.4.12:** Let (X, d) and (Y, ρ) be metric spaces and $K \subset X$ be compact. If $f : K \to Y$ is continuous, then f is uniformly continuous on K. **THRM 5.4.14:**The following are equivalent:(i) X is compact (ii) Every collection $\mathscr{C}$ of closed sets in X with the finite intersection property has a non-empty intersection (iii) X has the Bolzano-Weierstrass property (every infinite sequence $(\mathbf{x}_k)_{k=0}^{\infty} \subset X$ has at least one cluster point) (iv) X is sequentially compact (every sequence $(\mathbf{x}_k)_{k=0}^{\infty} \subset X$ has a convergent subsequence) (v) X is totally bounded (for all $\varepsilon > 0$ the cover $\mathscr{C} = \{B(\mathbf{x}, \varepsilon)\}_{\mathbf{x}\in X}$ has a finite subcover) and every open cover has a positive Lebesgue number (which depends on the cover). **Generalized Heine-Borel Theorem:** A metric space X is compact iff it is complete and totally bounded. **DEF: Pointwise Convergence** For any sequence of functions $(f_n)_0^{\infty}$, we can evaluate all the functions at a single point in the domain, which gives the sequence $(f_n(x))_0^{\infty}$. If for every x, the sequence converges, then we can define a new function given by $f(x) = \lim_{n\to\infty} f_n(x)$. In this case the sequence converges point wise to f(x). **DEF: Uniform Convergence** Let $(f_k)_{k=0}^{\infty}$ be a sequence of bounded functions from some domain X to a normed space Y. If $(f_k)_{k=0}^{\infty}$ converges to f in the L^∞ norm, then we say that the sequence $(f_k)_{k=0}^{\infty}$ converges uniformly to f. **PROP 5.5.5:** Uniform convergence impies pointwise convergence. **DEF: Banach Space** A complete normed linear space is called a Banach Space. **THRM 5.5.8:**For any $a < b \in \mathbb{R}$ the space $(C([a, b]; \mathbb{F}), \|\cdot\|_{L^\infty})$ is a Banach space. **COR 5.5.9:** If a sequence of continuous functions converges uniformly to a function f, then f is also continuous. **DEF: Convergence(Banach)** Let $(X, \|\cdot\|_X)$ be a Banach space and consider the sequence $(\mathbf{x}_k)_{k=0}^{\infty} \subset X$. We say that the series $\sum_{k=0}^{\infty} \mathbf{x}_k$ converges in X if the sequence $(\mathbf{s}_k)_{k=0}^{\infty}$ of partial sums defined by $s_n = \sum_{k=1}^{n} \mathbf{x}_k$ converges in X; otherwise we say that the series diverges. **DEF: Absolute Convergence(Banach)** Let $(\mathbf{x}_k)_{k=0}^{\infty}$ be a sequence in the Banach space $(X, \|\cdot\|_X)$. The series $\sum_{k=0}^{\infty} \mathbf{x}_k$ is said to absolutely converge if the series $\sum_{k=0}^{\infty} \|\mathbf{x}_k\|$ converges in $\mathbb{R}$. **PROP 5.5.13:** Let $(\mathbf{x}_k)_{k=0}^{\infty}$ be a sequence in the Banach space $(X, \|\cdot\|_X)$. If the series $\sum_{k=0}^{\infty} \mathbf{x}_k$ converges absolutely the it converges in X. **EX 5.5.20:** Let $(X, \|\cdot\|_X)$ be a Banach space. LeT $A \in \mathscr{B}(X)$ be a bounded operator with $\|A\| < 1$. The **Neumann Series** of A is the sum $\sum_{k=0}^{\infty} A^k$. If $\|\cdot\|$ is the operator norm, then:

$\sum_{k=0}^{\infty}\|A^k\| \le \sum_{k=0}^{\infty}\|A\|^k = \frac{1}{1-\|A\|} < \infty$. **PROP 5.5.21:** Let $(X, \|\cdot\|_X)$ be a Banach space. If $A \in \mathscr{B}(X)$ satisfies $\|A\| < 1$ then $I - A$ is invertible. Moreover, we have that $(I-A)^{-1} = \sum_{k=0}^{\infty} A^k$, and thus $\|(I-A)^{-1}\| \le (1-\|A\|)^{-1}$. **Topological Properties** preserved under homeomorphism: open sets, continuity, compactness, convergence, Connectedness **Not Topological:** Completeness, Cauchy (unless using topologically equivalent norms) **DEF: Bounded Functions** Let $(X, \|\cdot\|_X)$ be a Banach space. Write $L^{\infty}([a,b];X)$ to denote the set of bounded functions $f : [a,b] \to X$ equipped with teh sup norm $\|f\|_{l^{\infty}} = \sup_{t\in[a,b]}\|f(t)\|$. **THRM 5.7.2:** $L^{\infty}([a,b];X)$ is a Banach space. **DEF: Step Function** Let $(X, \|\cdot\|_X)$ be a Banach space. A map $f : [a,b] \to X$ is a step function if there is a (finite) partition $a = t_0 < t_1 < \cdots < t_{N-1} < t_N = b$ such that we may write f in the form: $f(t) = \sum_{i=1}^{N-1} \mathbf{x}_i \mathbb{1}_{[t_{i-1},t_i]} + \mathbf{x}_N \mathbb{1}_{[t_{N-1},t_N]}$, where each $\mathbf{x}_i \in X$ and $\mathbb{1}_E$ is the indicator function $\mathbb{1}_E(t) = 1$ if $t \in E$, 0 if $t \notin E$. **Continuous Linear Extension Theorem:** Let $(Z, \|\cdot\|_Z)$ be a normed linear space, $(X, \|\cdot\|_X)$ a Banach space, and $S \subset Z$ a dense subspace of Z. If $T : S \to X$ is a bounded linear transformation, then T has a unique linear extension to $\bar{T} \in \mathscr{B}(Z;X)$ satisfying $\|\bar{T}\| = \|T\|$. **DEF: Integral Properties** If $(X, \|\cdot\|)$ is a banach space, if $f \in \overline{S([a,b];X)}, \subset L^{\infty}([a,b];X)$ and $\alpha, \beta, \gamma \in [a,b]$, with $\alpha < \gamma < \beta$, then (i)$\|\int_a^b f(t)dt\| \le \int_a^b \|f(t)\|dt \le (b-a)sup_{t\in[a,b]}\|f(t)\|$ (ii)Restricitng f to $[\alpha,\beta]$ defines a function that we also denote by $f \in \overline{s([a,b];X)}$. We have $\int_a^b f(t)\mathbb{1}_{[\alpha,\beta]}dt = \int_{\alpha}^{\beta} f(t)dt$ (iii)$\int_{\alpha}^{\beta} f(t)dt = \int_{\alpha}^{\gamma} f(t)dt + \int_{\gamma}^{\beta} f(t)dt$. (iv) $F(t) = \int_a^t f(s)ds$ is continuous on $[a,b]$.

Ch.6 The directional derivative of f at x with respect to v is the limit $\lim_{t\to 0}\frac{f(x+tv)-f(x)}{t}$ this limit is often denoted $D_v f(x)$. **partial derivatives** the ith partial derivative of f at the point x is given by the limit $D_i f(x) = \frac{f(x+he_i)-f(x)}{h}$. **Frechet derivative***: Let $(X, ||.||)_X$ and $(Y, ||.||)_Y$ be banach spaces and let $U \subset X$ be an open set. A map f is differentiable at $x \in U$ if there exists a bouned linear transformation $Df(x) : X \to Y$ s. t. $\lim_{h\to 0}\frac{||f(x+h)-f(x)-Df(x)h||_Y}{||h||_X} = 0$. X and Y are banach spaces, $U \subset X$ be an open set, and f be differentiable on U. if Df given by $x \to Df(x)$ is also continuous, we say that f is continously differentiable on U. if f is differentiable on U, then f is locally **lipshitz**, that is $\forall\ x_0 \in U\ \exists B(x_0,\delta) \subset U$ and $L > 0$ s.t. $||f(x) - f(x_0)||_Y \le L||x - x_0||_X$ whenever $||x - x_0||_X < \delta$.

Mean Value Theorem: Let $(X, \|\cdot\|_X)$ be a Banach space, $U \subset X$ be an open set, and $f : U \to \mathbb{R}$ be differentiable on U. If for $\mathbf{x}, \mathbf{x} \in U$, the entire line segment $\ell(\mathbf{x},\mathbf{x}) := \{(1-t)\mathbf{x} + t\mathbf{x} \in [0,1]\}$ is also in U, then there exists $\mathbf{c} \in \ell(\mathbf{x},\mathbf{x})$ such that: $f(\mathbf{y}) - f(\mathbf{x}) = Df(\mathbf{c})(\mathbf{yx})$. **Fundamental Theorem of Calculus:** Let $(X, \|\cdot\|_X)$ be a Banach space: (i) If $f \in C([a,b];X)$, then for all $t \in (a,b)$ we have that: $\frac{d}{dt}\int_a^t f(s)ds = f(t)$ (ii) If $F : [a,b] \to X$ is continuously differentiable on (a,b) and $DF(t)$ extends to a continuous function on $[a,b]$, then: $\int_a^b DF(s)ds = F(b) - F(a)$. **Integral Mean Value Theorem:** Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be a Banach spaces, $U \subset X$ be an open set, and $f : U \to Y$ be continuously differentiable on U. If the line segment $\ell(\mathbf{x},\mathbf{x}) = \{t\mathbf{x} + (1-t)\mathbf{x} | t \in [0,1]\}$ is contained in U, then: $f(\mathbf{x}) - f(\mathbf{x} = \int_0^1 Df(t\mathbf{y} + (1-t)\mathbf{x})(\mathbf{x} - \mathbf{x})dt$, alternatively if we let $\mathbf{y} = \mathbf{x} + \mathbf{h}$, then $f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \int_0^1 Df(\mathbf{x} + t\mathbf{h})\mathbf{h}dt$ or $\|f(\mathbf{x}) - f(\mathbf{x})\|_Y \le \sup_{\mathbf{c}\in\ell(\mathbf{x},\mathbf{x})}\|Df(\mathbf{c})\|_{X,Y}\|\mathbf{x} - \mathbf{x}\|_X$ **Change of Variable Formula:** Let $(X, \|\cdot\|_X)$ be a Banach space and $f \in C([a,b];X)$. If $g : [c,d] \to [a,b]$ is continuous and g' is continuous on (c,d) and can be continuously extended to $[c,d]$, then $\int_c^d f(g(s))g'(s)ds = \int_{g(c)}^{g(d)} f(\tau)d\tau$

Lemma*: Given $f : X \to Y$ and linear $L : X \to Y$, to prove that f is differentiable at x with derivative L is equivalent to proving that for every $\epsilon > 0$ there is $||\xi|| < \delta$ we have $||f(x+\xi) - f(x) - L\xi||_Y \le \epsilon||\xi||_X$. **Linearity**: assume X and Y are Banach spaces, that U is an open neighborhood in X, and $f : U \to Y$ and $g : U \to Y$. If f and g are differentiable on U and $a, b \in \mathbb{F}$, then $af + bg$ is also differentiable on U, and $D(af(x) + bg(x)) = aDf(x) + bDg(x)$. **Product Rule**: Let X be banach space, that U is an open neighborhood of X, and that $f : U \to \mathbb{F}$ and $g : U \to \mathbb{F}$ if f and g are differentiable on U, then the product map fg is also differentiable on U and $D(f(x)g(x)) = g(x)Df(x) + f(x)Dg(x)$ for each $x \in U$. Rules of differentiation: (i) $u(x), v(x)$ are differentaible from R^n to R^m and $f : \mathbb{R}^n \to \mathbb{R}$ then $Df(x) = u(x)^T Dv(x) + v(x)^T Du(x)$. (ii) if $g : \mathbb{R}^n \to \mathbb{R}$ s.t. $g(x) = x^T Ax$ then $Dg(x) = x^T(A + A^T)$.**Chain Rule**: Assume X, Y, and Z are Banach, U and V are open neighborhoods of X and Y respectively, and $f : U \to V$ and $g : V \to Z$ with $f(U) \subset V$ if f is differentiable on U and g is differentiable on V, then the composite map $h = f \circ g$ is also differentiable on U and $Dh(x) = Dg(f(x))Df(x)$, for each $x \in U$

Sec6.5 Let $U \subset R^n$ be open. If $f : U \to \mathbb{R}^m$ is twice continuously differentiable on U, then $D^2 f(x)(x_1, x_2) = D^2 f(x)(x_2, x_1)$. Or $\frac{\partial^2 f_k}{\partial x_i \partial x_j} = \frac{\partial^2 f_k}{\partial x_j \partial x_i}$ Taylor's Theorem: For X,Y be Banach. $U \subset X$ an open set, and $f : U \to Y$ be k times differentiable. If $x \in U$ and $h \in X$ are such that the line $l(x, x+h)$ is contained in U, then

$f(x+h) = f(x) + Df(x)h + \frac{D^2 f(x)h^2}{2!} + R_k$. Example: for $f(x,y) = e^{x+y}$ at $(0,0)$ in direction of (h_1, h_2). Note, $f(0,0) = 1$. $D_h f(x) = \nabla f(0,0) \cdot h = h_1 D_1(0,0) + h_2 D_2(0,0) = h_1 + h_2$.