Course Outline by Day for M515

Tues, Sept 1

- Matrix notation: A, a_{ij}
- The matrices I and 0.
- Transpose A^T: switch rows and columns of A
- Column and row vectors (assume column): $A = [a_1, a_2, \ldots, a_n]$ means the $a_i's$ are the columns of A (they are column vectors)
- Scalars: real numbers
- Matrix arithmetic:
  matrix addition $A + B$: add corresponding entries (must be same size matrices)
  scalar multiplication αA: multiply each entry of A by α
  matrix multiplication AB: dot each row of A with each column of B
- The product Ax where x is a column vector: Use dot product definition or
  $Ax = x_1 a_1 + \ldots + x_n a_n$ where $x_i's$ are the entries of x and $a_i's$ are the columns of A
- Systems of equations and the three cases (one solution, no solutions, ∞ many solutions)
- Gaussian elimination, row ops:
  multiply row by number, switch rows, multiply row by number and add to another row
- Goal of elimination is to get to ref or rref
- Rank(M)=#non zero rows after rref
- Notation: R=real numbers, R^n=vectors of real numbers length n, $R^{m \times n}$= m by n matrices of real numbers, $R_r^{n \times m}$= m by n matrices of rank r
- Octave/Matlab example of rref

```
> A=[1 2;3 4]
A =
1 2
3 4
> b=[2 ; 2]
b =
2
2
> rref([A b])
ans =
1 0 -2
0 1 2
```

Thurs, Sept 3

- $B = [b_1, \ldots, b_p]$ then $AB = A[b_1, \ldots, b_p] = [Ab_1, \ldots, Ab_p]$
- Laws of matrix arithmetic:
  1. associative and commutative true for addition of matrices
  2. associative true but commutative not in general true for multiplication of matrices ($AB \neq BA$ in general)

3. distributive true (but keep matrices in same order)

- Special matrices: square, diagonal, upper triangular, lower triangular, symmetric
- Block matrices (can do in Matlab)
- Inner (dot) product $x^T y = x_1 y_1 + \ldots + x_n y_n$
- x, y orthogonal if $x^T y = 0$, orthonormal if $x^T x = 1$ and $y^T y = 1$ also
- Matrix A orthogonal if $A^T A = AA^T = I$ (used for rigid rotations)
- A orthogonal if and only if the columns of A form orthonormal set
- Det(M) defined for square matrices. For diag, upper triang or lower triang, Det(M)=product of diagonal elements
- Determinants and row ops (page 5 properties 3, 5, 7):
  1. multiply row by scalar multiplies $\det(A)$ by same scalar
  2. switch rows changes sign of $\det(A)$
  3. multiply row and add to another does not change $\det(A)$
- Use row operations to reduce matrix to triangular form, then use above properties to find det
- Cofactor method for determinants:

  Choose row or column, and multiply each entry by determinant of cofactor matrix (gotten by crossing out current row and column). Sign is determined by alternating $+, -, +, -, \ldots$ starting with upper left entry
- Matrix inverse A^{-1} satisfies $AA^{-1} = A^{-1}A = I$
- To find A^{-1} row reduce $[A\ I]$ to get $[I\ A^{-1}]$
- Important result: for $n \times n$ matrix A we have $Det(A) \neq 0$ iff A^{-1} exists iff $rank(A) = n$ iff $rref(A)$ has no zero rows

Sept 8

- Trace(A)=sum of diagonal elements
- $(AB)^T = B^T A^T$; $(AB)^{-1} = B^{-1} A^{-1}$
- More properties of Determinants (complete list on p5 of Laub):

  $\det(A^T) = \det(A)$

  $\det(AB) = \det(A) \cdot \det(B)$

  $\det(A) = 0$ if any row or column of A is all zeros

  $\det(A^{-1}) = \frac{1}{\det(A)}$
- rref and the three cases:
  1. overdetermined systems (no solutions): more equations than unknowns (after rref)
  2. underdetermined systems (∞ many solutions): more unknowns than equations (after rref)
  3. exactly one solution: same number of equations as unknows after (rref)
- $A = LU$ factorization: L lower triangular, 1's on diagonal; U upper triangular; U is row-reduced A; L contains multiplication factors

Sept 10

- See Matlab outline

Sept 15

- Fields: 2 operations (addition and multiplication), both associative and commutative, together distributive, additive identity 0, additive inverses (negative), multiplicative identity 1, multiplicitive inverses except for 0 (reciprocal)
- We will use R and possibly C (complex numbers) for our fields
- Vector spaces V: Field F of scalars (R or C) plus set of vectors V (usually R^n for some n)
- Vectors have one operation (addition), commutative and associative, there is an identity (the zero vector), and there are additive inverses
- Scalar multiplication exists and is commutative, associative, distributive (2 types), and $1 \cdot v = v$
- Examples of vector spaces (or not):
  1. $V = R^2$ (the coordinate plane), $F = R$
  2. $V = R^3$ (3D space), $F = R$
  3. $V =$first quadrant of R^2 (x and y both positive), $F = R$ is NOT a vector space
  4. $V = R^{2\times 2}$ (two by two matrices), $F = R$
- Subspace V of vector space W ($V \subseteq W$): closed under scalar multiplication and vector addition
  1. if $v \in V$ and $\alpha \in F$ then $\alpha v \in V$
  2. if $v_1 \in V$ and $v_2 \in V$ then $v_1 + v_2 \in V$
- Subspaces of $V = R^2$ are lines through origin
- Subspace of $V = R^3$ are lines through origin and planes through origin
- Symmetric matrices are subspace of matrices ($n \times n$)
- Anti-symmetric matrices ($A^T = -A$) are also a subspace

Sept 17

- Vectors $\{v_1, v_2, \ldots, v_n\}$ are linearly dependent if there exist $\alpha_i' s$, not all zero, for which $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$.
- Vectors $\{v_1, v_2, \ldots, v_n\}$ are linearly independent if $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0$ implies that all $\alpha_i' s$ are zero.
- If the only solution to $Va = \vec{0}$ is $a = \vec{0}$ then the columns of V are linearly independent. No solutions is not possible here. If there is one solution (the zero solution) the columns are linearly independent. If there are infinitely many, the columns are linearly dependent.
- To determine the linear dependence/independence of a set of vectors, create a matrix V with those vectors as the columns and use $rref$ on V. Rows are equations, columns are variables. After $rref$, eliminate any zero rows. If more variables than equations (more columns than rows) then the vectors are linearly dependent. If the same (in this case you get the identity matix) then the vectors are linearly independent.
- The span of a set of vectors $X = \{v_1, \ldots, v_n\}$ is denoted $sp(X)$, and is defined to be set of all possible linear combinations $\alpha_1 v_1 + \ldots + \alpha_n v_n$.
- A set of vectors $X = \{v_1, \ldots, v_n\}$ form a basis for a vector space V if
  1. The vectors in X are linearly independent
  2. $sp(X) = V$
- The number of vectors in a basis is independent of the basis
- The dimension of a vector space is $\dim(V) = \#$ of basis vectors

- Standard bases for various spaces ($0's$ and $1's$)

Sept 22

- Change of basis: If $V = \{v_1, v_2\}$ is a basis for R^2, then $\begin{bmatrix} 2 \\ 3 \end{bmatrix}_V = 2v_1 + 3v_2$
- Sum of subspaces $R + S = \{r + s : r \in R, s \in S\}$ is also a subspace
- Intersection of subspaces $R \cap S = \{v : v \in R$ and $v \in S\}$ is also a subspace
- Union of subspaces $R \cup S$ is generally not a subspace
- $T = R \oplus S$ is the direct sum of R and S if
  1. $R \cap S = \{0\}$
  2. $R + S = T$
- If $T = R \oplus S$ then
  1. Every $t \in T$ can be written uniquely as $t = r + s$ where $r \in R$ and $s \in S$
  2. $\dim(T) = \dim(R) + \dim(S)$
- An $m \times n$ matrix A can be thought of as a function (linear transformation) from R^n to R^m via matrix-vector multiplication. We write $A : R^n \to R^m$. Note that $A^T : R^m \to R^n$.
- $A(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 A v_1 + \alpha_2 A v_2$ is what makes the transformation (function) linear
- If $A : V \to W$, the range of A is $R(A) = \{w \in W : w = Av$ for some $v \in V\}$ and the null space of A is $N(A) = \{v : Av = \vec{0}\}$.
- $R(A)$ and $N(A)$ are subspaces

Sept 24

- $R(A)$ is equal to the span of the columns of A, and so is sometimes called the column space.
- Kronecker delta $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$. Set of vectors is orthonormal if $v_i^T v_j = \delta_{ij}$.
- Orthogonal complement of subspace S is $S^\perp = \{v : v^T s = 0$ for all $s \in S)$ which is also a subspace
- If S is a subspace of R^n then $S \oplus S^\perp = R^n$
- $N(A)^\perp = R(A^T)$ and $R(A)^\perp = N(A^T)$
- Four fundamental subspaces $N(A), N(A)^\perp, R(A), R(A)^\perp$. See page 23 of Laub.
- $A : V \to W$ is onto if $R(A) = W$
- $A : V \to W$ is one-to-one if $N(A) = \vec{0}$ or equivalently if $Av_1 = Av_2$ implies that $v_1 = v_2$.
- $rank(A) = \dim R(A)$
- $\dim R(A) = \dim N(A)^\perp$
- If $A : R^n \to R^m$ then $\dim N(A) + \dim R(A) = n$
- If $A : R^n \to R^m$, A is onto iff $rank(A) = m$ and A is 1-1 iff $rank(A) = n$

Sept 29

- Finding a basis for $R(A)$: The columns of A that correspond to the pivot columns of $rref(A)$ form a basis for $R(A)$
- Finding a basis for $N(A)$: Find $rref(A)$ and solve the resulting equations $Ax = 0$ for the

pivot variables in terms of the other variables. Then write the result as a linear combination of the basis vectors using the free variables as the scalars.

- For $R(A)^{\perp}$ find vectors perpendicular to $R(A)$ by writing equations and using $rref$, or use $R(A)^{\perp} = N(A^T)$
- For $N(A)^{\perp}$ find vectors perpendicular to $N(A)$ by writing equations and using $rref$, or use $N(A)^{\perp} = R(A^T)$
- Finding bases for other vector spaces: Create a general vector that is in the subspace, write equations to relate the components of the vector, and then write the general vector as a linear combination of constant vectors (which become the basis vectors).

Oct 1

- Gram-Schmidt: The orthogonal projection of v in the w direction is given by $proj(v,w) = \frac{w^T v}{w^T w} w$.

  Start with basis $v_1, v_2, \ldots, v_n$ and create new orthogonal basis $w_1, w_2, \ldots, w_n$:

  1. $w_1 = v_1$
  2. $w_2 = v_2 - proj(v_2, w_1)$
  3. $w_3 = v_3 - proj(v_3, w_1) - proj(v_3, w_2)$
  4. and so on

Oct 13

- Eigenvalues and eigenvectors:
  1. Solutions λ to $\det(A - \lambda I)$ are eigenvalues
  2. Given an eigenvalue λ, a solution v to $(A - \lambda I)v = 0$ is an eigenvector corresponding to λ
- Any multiple of an eigenvector is the same eigenvector

Oct 15

- Singular Value Decomposition $A = U\Sigma V^T = U_1 S V_1^T$. $\Sigma = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$, $U = [U_1 \ U_2]$, $V = [V_1 \ V_2]$. S is a diagonal matrix of the non-zero singular values $\sigma_1 \geq \sigma_2 \geq \ \ldots \ \geq \sigma_r$, where $r = rank(A)$. U_1 and V_1 each have r columns. The $\sigma' s$ are the square roots of the eigenvalues of $A^T A$. The columns of V_1 are the eigenvectors of $A^T A$ which correspond to non-zero eigenvalues. The columns of V_2 consist of any orthonormal basis for $N(A)$. The columns of U_1 are defined by $u_i = A v_i \frac{1}{\sigma_i}$ where the $v_i' s$ are the columns of V_1. The columns of U_2 are chosen to be orthonormal with respect to each other and to the columns of U_1.
- Some properties of the SVD:
  1. U and V are orthogonal matrices ($U^T U = UU^T = I$ and $V^T V = VV^T = I$).
  2. Σ and A are the same size.
  3. $A^T = V \Sigma^T U^T$ is an SVD for A^T.
  4. $rank(A)$ = number of non-zero $\sigma_i' s$

5. $A = \sum_{i=1}^{r} \sigma_i u_i v_i^T$ (sum of rank 1 matrices)

- Orthogonal bases of four fundamental subspaces:
  1. $R(U_1) = R(A) = N(A^T)^\perp$
  2. $R(U_2) = R(A)^\perp = N(A^T)$
  3. $R(V_1) = N(A)^\perp = R(A^T)$
  4. $R(V_2) = N(A) = R(A^T)^\perp$

  Note: Since columns of U and V are orthogonal, no row reduction is needed to find these bases; just take appropriate columns of U or V.

Oct 20

- Projection matrices

  If $V = X \oplus Y$ is a direct sum, then any $v \in V$ can be written uniquely as $v = x + y$ where $x \in X$ and $y \in Y$. Then the projection matrix $P_{X,Y}$ is defined by $P_{X,Y}\, v = x$. When $Y = X^\perp$ we write $P_{X,Y} = P_X$ and call the projection an orthogonal projection.
- P is an orthogonal projection if and only if P is symmetric and $P^2 = P$.
- If P_X is an orthogonal projection onto X then $I - P_X$ is an orthogonal projection onto $X^\perp$.
- The orthogonal projection onto $R(A)$, when the columns of A are linearly independent, is given by $P_{R(A)} = A(A^TA)^{-1}A^T$.
- The orthogonal projections onto the four fundamental subspaces of A are $P_{R(A)} = U_1U_1^T$, $P_{R(A)^\perp} = U_2U_2^T$, $P_{N(A)} = V_2V_2^T$, $P_{N(A)^\perp} = V_1V_1^T$.

Oct 22

- Inner products

  An inner product must satisfy three properties:
  1. $< v,v > \ \geq 0$ and $< v,v >= 0$ if and only if $v = 0$
  2. $< v,w >=< w,v >$ (commutativity)
  3. $< v, aw_1 + bw_2 >= a < v,w_1 > +b < v,w_2 >$ (linearity)
- The usual inner product for R^n is the dot product $< v,w >= v^Tw$
- A weighted inner product for R^n is $< v,w >= v^TQw$ where Q is symmetric and positive definite

  Note: Positive definite means $v^TQv > 0$ for all v except $v = 0$. This is equivalent to all eigenvalues being positive.
- For linear transformations A we have $< v,Aw >=< A^Tv,w >$
- An inner product for vector spaces consisting of matrices is given by $< A,B >= Trace(A^TB)$.
- An inner product induces a norm by $\|v\| = \sqrt{< v,v >}$.
- If P is an orthogonal projection then $\|Pv\| \leq \|v\|$

Oct 27

- Vector Norms

  A norm must satisfy three properties:
  1. $\|v\| \geq 0$ for all v and $\|v\| = 0$ if and only if $v = 0$.

2. $\|av\| = |a|\ \|v\|$ for scalars a.
3. $\|v + w\| \le \|v\| + \|w\|$.

- Three common vector norms:
  1. $\|v\|_1 = \sum |v_i|$
  2. $\|v\|_2 = \sqrt{\sum |v_i|^2}$

     The 2 norm is induced by the standard dot product.
  3. $\|v\|_\infty = \max |v_i|$
- Holder inequality: $|< v, w >| \le \|v\|_2 \|w\|_2$
- Pythagorean Identity: if v and w are orthogonal (inner product zero) then $\|v \pm w\|_2^2 = \|v\|_2^2 + \|w\|_2^2$
- Matrix Norms

  A matrix norm satisfies the same three properties as vector norms.
- Four common matrix norms:
  1. $\|A\|_F = \sqrt{\sum_i \sum_j a_{ij}^2}$ (Frobenius norm)
  2. $\|A\|_1 = \max_j \left(\sum_i |a_{ij}| \right)$ (max column sum of absolute values)
  3. $\|A\|_\infty = \max_i \left(\sum_j |a_{ij}| \right)$ (max row sum of absolute values)
  4. $\|A\|_2 = \max_i \sigma_i$ (largest singular value)
- Vector-Matrix norm inequalities
  1. $\|Av\| \le \|A\|\ \|v\|$ for $1,\ 2,\ \infty$ norms
  2. $\|AB\| \le \|A\|\ \|B\|$ for $1,\ 2,\ \infty$, and F norms
- Norms can be used to measure distance. The distance between vectors v and w is $\|v - w\|$ and the distance between matrices A and B is $\|A - B\|$.
- The condition number of an invertible matrix A is $cond(A) = \|A\| \cdot \|A^{-1}\|$
- $cond(A) \ge 1$
- If we want to solve $Ax = b$, any error Δb in b or ΔA in A induces an error Δx in the solution. It can then be shown that $\frac{\|\Delta x\|}{\|x\|} \le cond(A) \frac{\|\Delta b\|}{\|b\|}$ and $\frac{\|\Delta x\|}{\|x + \Delta x\|} \le cond(A) \frac{\|\Delta A\|}{\|A\|}$.

  Thus relative (percentage) errors in b or A can induce a relative error in the solution x which is $cond(A)$ times bigger than the relative error in b or A.

Oct 29

- Least Squares: Find the x for which $\|Ax - b\|_2$ is a minimum. This is the same as finding the Ax in $R(A)$ that is closest to b. Then $b - Ax$ (the vector from Ax to b) will be orthogonal to $R(A)$. Thus $0 = (Ay)^T(b - Ax) = y^T A^T (b - Ax) = y^T (A^T b - A^T Ax)$. Since this is true for any y, it must be that $A^T b - A^T Ax = 0$ or

  $$A^T b = A^T Ax.$$

  These are called the **normal equations** for the least squares problem.
- When $A^T A$ is invertible, we have the solution $x = (A^T A)^{-1} A^T b$.
- An SVD for A can be used to get the solution $x = V_1 S^{-1} U_1^T b + V_2 z_2$ for any compatible

vector z_2 (that is, V_2z_2 must be defined). Since the columns of V_2 form a basis for the null space of A, V_2z_2 represents any vector in the null space of A. When V_2 is absent, the solution is unique.

- To fit $y = f(t) = c_1p_1(t) + c_2p_2(t) + c_3p_3(t)$ to a set of data (t_1, y_1), (t_2, y_2), (t_3, y_3), (t_4, y_4):

  1. Form $A = \begin{bmatrix} p_1(t_1) & p_2(t_1) & p_3(t_1) \\ p_1(t_2) & p_2(t_2) & p_3(t_2) \\ p_1(t_3) & p_2(t_3) & p_3(t_3) \\ p_1(t_4) & p_2(t_4) & p_3(t_4) \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$, $x = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$

  2. Solve the least squares problem of finding the x for which $\|Ax - y\|_2$ is a minimum using $x = (A^TA)^{-1}A^Ty$ or $x = V_1S^{-1}U_1^Ty + V_2z_2$ (z_2 anything).

Nov 5

- For triangular matrix, eigenvalues appear on diagonal