Jack Le - (713)-449-5966
call if found

Lecture 1 (Vectors)

1. Vector sum: $\vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, $\vec{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$, $\vec{v}+\vec{w} = \begin{bmatrix} v_1+w_1 \\ \vdots \\ v_n+w_n \end{bmatrix}$
2. Scalar Multiplication: $c\vec{v} = \begin{bmatrix} cv_1 \\ \vdots \\ cv_n \end{bmatrix}$
3. Linear Combination: $\vec{W} = a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_n\vec{v}_n$
4. Convex Combination: Lin comb when $a_1 + a_2 + \dots + a_n = 1$
   ↳ for two vectors, convex is $(1-t)\vec{v} + t\vec{w}$, $0 \le t \le 1$
5. Vector Properties:
   - $\to \vec{v}+\vec{w} = \vec{w}+\vec{v}$ $\quad \to (\vec{v}+\vec{w}) + \vec{v} = \vec{v} + (\vec{w}+\vec{v})$
   - $\to (ab)\vec{v} = a(b\vec{v})$ $\quad \to a(\vec{v}+\vec{w}) = a\vec{v} + a\vec{w}$
   - $\to (a+b)\vec{v} = a\vec{v} + b\vec{v}$
6. Magnitude: $\|\vec{v}\| = \sqrt{v_1^2 + \dots + v_n^2}$
   $\|-\vec{v}\| = \|\vec{v}\|$ and $\|c\vec{v}\| = |c|\,\|\vec{v}\|$
7. Displacement: $\vec{v} - \vec{w}$ is displacement vector, $\|\vec{v}-\vec{w}\|$ is distance btwn $\vec{v}$ & $\vec{w}$

Lecture 2 (Dot Products)

1. $\cos\theta = \dfrac{\vec{v}\cdot\vec{w}}{\|v\|\,\|w\|}$ (angle between $\vec{v}$ and $\vec{w}$)
2. Dot prod: $\vec{x}\cdot\vec{y} = x_1y_1 + \dots + x_ny_n = \sum_{i=1}^{n} x_iy_i$
3. Orthogonal: $\vec{x}\cdot\vec{y} = 0$
4. Dot prod properties:
   - $\to \vec{v}\cdot\vec{w} = \vec{w}\cdot\vec{v}$ $\quad \to \vec{v}\cdot(c\vec{w}) = c(\vec{v}\cdot\vec{w})$
   - $\to \vec{v}\cdot\vec{v} = \|\vec{v}\|^2$ $\quad \to a\vec{v}\cdot b\vec{w} = ab(\vec{v}\cdot\vec{w})$
   - $\to \vec{v}\cdot(\vec{w}_1 + \vec{w}_2) = \vec{v}\cdot\vec{w}_1 + \vec{v}\cdot\vec{w}_2$
   - $\to \vec{v}\cdot(c_1\vec{w}_1 + c_2\vec{w}_2) = c_1(\vec{v}\cdot\vec{w}_1) + c_2(\vec{v}\cdot\vec{w}_2)$
5. Correlation Coefficient: $r = \dfrac{\hat{X}\cdot\hat{Y}}{\|X\|\,\|Y\|}$, $-1 \le r \le 1$

Lecture 3 (Planes)

To find side of plane of point, plug in point and compare to d.

Equation form: $ax + by + cz = d$ $\quad$ $ax+by+cz > d$, $ax+by+cz < d$ are sides

Normal vector: find point P and $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ ← orthogonal to plane

Parametric: $P + te + t'e'$

Equation ⟷ Normal

1. coefficients of equation is $\vec{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$
2. use ~~displacement from $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ to point~~ or plug in point to equation to find d.

Equation → Parametric

1. find 3 points from equation, (x, y, z)
2. pick one to be point P
3. find displacement $\vec{e} = \vec{Q} - \vec{P}$, $\vec{e}' = \vec{R} - \vec{P}$
4. Put into parametric form

Parametric → Normal

1. solve for $\vec{n}$, $\vec{n}\cdot\vec{e} = 0$, $\vec{n}\cdot\vec{e}' = 0$
2. use point to find plane

Lecture 4 (Subspaces)

1. Span: $\text{span}(\vec{v}_1, \dots, \vec{v}_k) = \{c_1\vec{v}_1 + \dots + c_k\vec{v}_k \;;\; c_1, \dots, c_k \in \mathbb{R}\}$
   ↳ $\vec{0}$ is always in span, $\text{span}(\vec{0}) = \{\vec{0}\}$
   ↳ spans are not unique. Multiple vectors can make same span
   ↳ linear subspace is same thing as span; must contain origin
   ↳ all linear combs of vectors in subspace are also in subspace
2. Dimension: for subspace V, $\dim(V)$ = smallest # of vectors to span V.
   For $V = \text{span}(v_1, \dots, v_k)$, $\dim(V) = k$ if each v_i is not LC of other vectors.

Lecture 5 (Basis / Orthogonality)

1. Basis: basis of subspace V is spanning set of $\dim(V)$ vectors (no redundant vectors)
   ↳ $\{e_1, e_2, e_3\}$ is basis of $\mathbb{R}^3$; there are many more
2. Dimension Criterion: ① span of one vector has $\dim(V) = 1$
   ② $\dim(V) < k$ if vectors are scalar mults or linear combos ← pick a vector & show it's a linear comb of others
   ↳ remove all scalar mults & linear combos, remaining vectors is dimension
3. Orthogonal Basis: if $\vec{v}_1, \dots, \vec{v}_k$ is orthogonal collection (all ⊥ to each other), then it is a basis for $\text{span}(\vec{v}_1, \dots, \vec{v}_k)$. $\dim(\text{span}(\vec{v}_1, \dots, \vec{v}_k)) = k$ → orthogonal basis
4. Orthonormal basis: orthogonal basis w/ unit vectors → $\frac{v}{\|v\|}$
5. Standard basis: for $\mathbb{R}^3$, it's → span $\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$
6. Fourier formula: for orthogonal collection $v_1, \dots, v_k \in \mathbb{R}^n$ and $v \in \text{span}(v_1, \dots, v_k)$
   $\vec{v} = \sum_{i=1}^{k}\left(\dfrac{v\cdot v_i}{v_i\cdot v_i}\right)v_i \to c_i = \dfrac{v\cdot v_i}{v_i\cdot v_i} \to$ if orthonormal, $v = \sum (v\cdot v_i)v_i$

Lecture 6 (Projections)

1. $\text{Proj}_{\vec{w}}\vec{x} = \left(\dfrac{\vec{x}\cdot\vec{w}}{\vec{w}\cdot\vec{w}}\right)\vec{w}$ ✓ point in $L = \text{span}(\vec{w})$ closest to $\vec{x}$.
2. Projecting onto lines using identity:
   $\text{Proj}_{\vec{w}}(c_1x_1 + \dots + c_kx_k) = c_1\text{Proj}_{\vec{w}}(x_1) + \dots + c_k\text{Proj}_{\vec{w}}(x_k)$
   $= \text{Proj}_{\vec{w}}(\vec{x})$
   ↳ compute $\text{Proj}_{\vec{w}}(\vec{e}_1), \dots, \text{Proj}_{\vec{w}}(\vec{e}_k)$
   ↳ then $\vec{v} = v_1\vec{e}_1 + v_2\vec{e}_2 + \dots + v_k\vec{e}_k$
3. Orthogonal Projection Theorem: (subspace V)
   ↳ for orthogonal basis $\vec{v}_1, \dots, \vec{v}_k$ of V
   $\text{Proj}_V(\vec{x}) = \text{Proj}_{\vec{v}_1}(\vec{x}) + \dots + \text{Proj}_{\vec{v}_k}(\vec{x})$
4. Orthogonal Projection Theorem (Ver. 2): $\vec{x} = \vec{v} + \vec{v}'$, where $\vec{v} = \text{Proj}_V(\vec{x})$ and $v' = x - \text{Proj}_V(\vec{x})$ [V is lin subspace, $\vec{x} \in V$]

Lecture 7 (Orthogonal Basis)

1. $\vec{y}$ and $\vec{x}' = \vec{x} - \text{Proj}_{\vec{y}}(\vec{x})$ is orthogonal basis of $\text{span}(\vec{x}, \vec{y})$
2. Linear Regression Steps:
   1) find X and Y
   2) find $\hat{X} = X - \text{Proj}_1(X) = X - \bar{x}1$
   3) use $\{\hat{X}, 1\}$ as orthog basis for the space $V = \text{span}(X, 1)$
   4) Project Y onto V
      $\text{Proj}_V Y = \text{Proj}_{\hat{X}} Y + \text{Proj}_1 Y$
      $= a\hat{X} + \bar{y}1$
   5) sub $(X - \bar{x}1)$ for $\hat{X}$
      $= a(X - \bar{x}) + \bar{y}$
      $= aX - a\bar{x} + \bar{y}$
   6) $y = mx + b$, $m = a$, $b = \bar{y} - a\bar{x}$

Lecture 8 (Level Sets)

1. Scalar-valued function: $f: \mathbb{R}^n \to \mathbb{R}$
2. Vector-valued function: $f: \mathbb{R}^n \to \mathbb{R}^m$ } same
   ↳ component funcs: $f_1, \dots, f_m: \mathbb{R}^n \to \mathbb{R}$
3. Composition: $(f\circ g)(x) = f(g(x))$ $\quad \mathbb{R}^n \xrightarrow{g} \mathbb{R}^p \xrightarrow{f} \mathbb{R}^m$
   ↳ output size of g match input of f
4. Graph: $f: \mathbb{R}^n \to \mathbb{R}$
   $\text{Graph}(f) = \{(x_1, \dots, x_n, z) \in \mathbb{R}^{n+1} : z = f(x_1, \dots, x_n)\}$
5. Level Sets: for $f: \mathbb{R}^n \to \mathbb{R}$ and $c \in \mathbb{R}$, level set of f @ level c is pts $\in \mathbb{R}^n$ s.t. $f(x_1, \dots, x_n) = c$
6. Contour Plot: depicts level sets in z plane for many values of c in $\mathbb{R}^2$.
   let $f(x, y) = x^2 + y^2$ $\quad$ you can add constants to get 2 different functions w/ same level set
   $c = 0: \; 0 = x^2 + y^2$
   $c = 1: \; 1 = x^2 + y^2$

Dimension Criterion

One Vector: $\text{span}(\vec{v})$ has $\dim = 1$

Two Vectors: $\dim(\text{span}(\vec{v}, \vec{w})) = 2$ if $\vec{v}$ and $\vec{w}$ not scalar mult. else $\dim = 1$

Three Vectors: $\dim(V) = 3$ except

① all three vectors scalar mult → $\dim = 1$

② two vectors scalar mult → $\dim = 2$

③ no scalar mult, but one v_i is linear comb of other two → $\dim = 2$

↳ every v_i will be linear comb of other two if one is

Ex: span $\swarrow \mathbb{R}^3$ subspace in $\mathbb{R}^4$

$$U = \left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{R}^4 : \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2 \\ 3 \\ 1 \end{bmatrix} = 0 \right\}$$

$$-x + 2y + 3z + w = 0$$
$$x = 2y + 3z + w$$

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2y+3z+w \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 2y \\ y \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 3z \\ 0 \\ z \\ 0 \end{bmatrix} + \begin{bmatrix} w \\ 0 \\ 0 \\ w \end{bmatrix}$$

$$= y\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + z\begin{bmatrix} 3 \\ 0 \\ 1 \\ 0 \end{bmatrix} + w\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$U = \text{span}\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Ex: line as span of single vector

$y = 2x$

$\hookrightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$L = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$

To show that three vectors are not on a line, get displacement vector $\vec{PQ}$ and $\vec{PR}$ and make sure they are not scalar multiples.

Two var, all 2^{nd} partials

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

Ex: find 3 5-vectors where span$(v_1, v_2, v_3) = U$

$U = \{ x \in \mathbb{R}^5 : x \cdot v_1 = 0,\ x \cdot v_2 = 0 \}$

$x \cdot v_1 = 0$, $v_1 = (1, -2, 0, -4, 3)$

$x_1 - 2x_2 - 4x_4 + 3x_5 = 0$

$x_1 = 2x_2 + 4x_4 - 3x_5$

$x \cdot v_2 = 0$, $v_2 = (0, 5, -1, 2, 2)$

$5x_2 - x_3 + 2x_4 + 2x_5 = 0$

$x_3 = 5x_2 + 2x_4 + 2x_5$

$$x = \begin{bmatrix} 2x_2 + 4x_4 - 3x_5 \\ x_2 \\ 5x_2 + 2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2\begin{bmatrix} 2 \\ 1 \\ 5 \\ 0 \\ 0 \end{bmatrix} + x_4\begin{bmatrix} 4 \\ 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} + x_5\begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

so, $v_1 = \uparrow$, $v_2 = \uparrow$, $v_3 = \uparrow$

Lecture 9 (Partial Derivatives)

1. Partial Notation: $\frac{\partial f}{\partial x_1}(a,b)$, $\left.\frac{\partial f}{\partial x_1}\right|_{(a,b)}$, $f_{x_1}(a,b)$ $\leftarrow$ all three are same
   $\leftarrow$ move in x_1 dir, all other constant
2. Definition: $\frac{\partial f}{\partial x_1}(a,b) = \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h}$
3. Symbolic: Don't plug constants until final calculation
   $\hookrightarrow$ think of other x_j's as constant
4. Numerical: replace each x_j with constants before differentiating
5. Partial Derivatives on Contour plot
   $\hookrightarrow$ $f_x(a,b)$ is slope experienced walking on $z = f(x,y)$ from W to E at pt (a,b)
   $\hookrightarrow$ $f_y(a,b)$ is slope walking on $z = f(x,y)$ from S to N
6. Second partial: $\frac{\partial^2 f}{\partial x_i^2 \partial x_j^2} = \frac{\partial f}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right)$

   $\frac{\partial^2 f}{\partial x \partial y} = f_{xy} = f_{yx} = (f_y)_x = (f_x)_y$
7. Clairaut-Schwarz: $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} \to f_{xy} = f_{yx}$
8. Functions that satisfy $f_{xx} + f_{yy} = 0$ are called harmonic
9. Chain Rule: $\frac{\partial}{\partial x}(F^2) = \frac{\partial}{\partial F}\frac{\partial F}{\partial x} = 2F\frac{\partial F}{\partial x} = 2FF_x$

Lecture 10: Extrema & Critical Points

- Local max @ (a,b) if $f(a,b) \geq f(x,y)$ for all (x,y) nearby
- Local min @ (a,b) if $f(a,b) \leq f(x,y)$
- Crit Point: $\nabla f = 0$, possible extrema
- Saddle Point: local min in one dir, max in other dir if point is local/global extrema in both, not saddle
- Extrema on region:
  1) find crit points in interior $(\nabla f = 0)$
  2) find crit points on boundary (single-var)
  3) find crit points on corners (intersection of boundaries)
  4) plug all into f

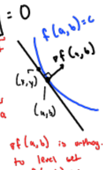
Lecture 11: Gradients

$\nabla f(x,y) = \begin{bmatrix} f_x \\ f_y \end{bmatrix}$ grad points in dir of maximal increase

Linear Approx: $f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \begin{bmatrix} x_1 - a_1 \\ \vdots \\ x_n - a_n \end{bmatrix}$

Tangent Line: $\underbrace{\nabla f(a,b)}_{\perp \text{ to T.L.}} \cdot \begin{bmatrix} x-a \\ y-b \end{bmatrix} = 0$

Tangent Plane: $\underbrace{\nabla f(a,b,c)}_{\text{scalar multiple of normal of plane}} \cdot \underbrace{\begin{bmatrix} x-a \\ y-b \\ z-c \end{bmatrix}}_{\text{displacement vectors}} = 0$



Unit Vector: $\underbrace{\nabla f(\vec{a}) / \|\nabla f(\vec{a})\|}_{\text{dir } f \text{ incr most rapidly at a}}$

Gradient Descent: $\vec{a}_1 = \vec{a}_0 + \overset{\text{learning rate}}{t} \nabla f(\vec{a}_0)$
$\vec{a}_2 = \vec{a}_1 + t \nabla f(\vec{a}_1)$

$\nabla f(a,b)$ is orthog. to level set $f(a,b) = c$

Lecture 12: Lagrange Multipliers

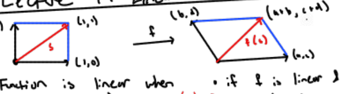
Maximize f subject to constraint $g = c$.

Step 1: check $\nabla g(\vec{x}) = 0$, find possible pts

Step 2: check $\nabla f = \lambda \nabla g$ — If dividing by VAR, always check VAR=0 as a case
- Method 1: find $\lambda = \frac{\text{num}}{\text{denom}}$
  - case 1) check denom = 0 for all
  - case 2) denom non-zero, set equations equal
- Method 2: cross multiply
  - case 1) check $\lambda = 0$ with all
  - case 2) $\lambda \neq 0$, eliminate lambda, solve new system with cross-multiplied equation
    - case 2.1) if 3-variable, try getting two to equal 0 $(y = x = 0)$, also $y = x \neq 0$

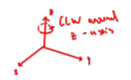
Step 3: plug all points into f

Lecture 14: Linear Transformations

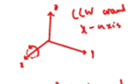


- Function is linear when
  $f(c\vec{x}) = c f(\vec{x}) \rightarrow cf(\vec{0}) = \vec{0}$
  $f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y})$
- if f is linear & g is linear, $f \circ g$ is linear
- $T_A(\vec{x}) = A\vec{x}$, $T_B(\vec{x}) = B\vec{x}$, $T_A \circ T_B(\vec{x}) = A(B\vec{x})$
- $T_{A+B}(\vec{x}) = (A+B)\vec{x} = A\vec{x} + B\vec{x} = T_A(\vec{x}) + T_B(\vec{x})$
- $T_{cA}(\vec{x}) = (cA)\vec{x} = c(A\vec{x}) = cT_A(\vec{x})$
- Matrix multiplication: row of A dot col of B
- CCW rot in $\mathbb{R}^2$: $A_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ rotations abt origin in $\mathbb{R}^2$ commute

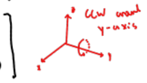
$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$



$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$



$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$



Lecture 16: Markov Matrices

- Markov Matrix: square matrix, non-negative, columns sum to 1
- Markov matrices with all positive entries will always stabilize (can still stabilize w/ 0 entries)
- If stabilizes, initial distribution doesn't matter
- Steps: suppose P_n, find P_{n+1} using it

  $\vec{P}_{n+1} = \begin{bmatrix} A_{n+1} \\ B_{n+1} \end{bmatrix} = \begin{bmatrix} 1/2 A_n + 1/3 B_n \\ 1/2 A_n + 2/3 B_n \end{bmatrix} = \begin{bmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{bmatrix}\begin{bmatrix} A \\ B \end{bmatrix} = M\vec{P}_n$

  $\vec{P}_k = M^k \vec{P}_0$
- If M has all positive entries, all columns of the stabilized matrix M^k will be same
  $\hookrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ stabilizes, but columns not all same

Lecture 17: Multi-var Chain Rule

- $(D(f \circ g))(\vec{a}) = Df(g(\vec{a})) \cdot Dg(\vec{a})$
- single var: $(f(g(x)))' = f'(g(x)) \cdot g'(x)$
- Linear Approx:
  $(f \circ h)(\vec{x}) \approx (f \circ h)(\vec{a}) + (D(f \circ h)(\vec{a}))(\vec{x} - \vec{a})$
- Alternate form:
  $\frac{\partial f}{\partial x_j} = \sum_{k=1}^{p} \frac{\partial f}{\partial y_k} \cdot \frac{\partial y_k}{\partial x_j}$ — y_k is components of f, x_j is outer variable

  Ex: $h(r, \theta) = f(r\cos\theta, r\sin\theta)$

  $\frac{\partial h}{\partial r} = \frac{\partial}{\partial r} f(r\cos\theta, r\sin\theta) = \frac{\partial}{\partial r} f(x,y)$
  $= \sum_{k=1}^{2} \frac{\partial f}{\partial x_k}\frac{\partial x_k}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r}$
  $= \frac{\partial f}{\partial x}\frac{\partial}{\partial r}(r\cos\theta) + \frac{\partial f}{\partial y}\frac{\partial}{\partial r}(r\sin\theta)$

  $\frac{\partial h}{\partial r} = \cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y}$
- $D(f \circ g \circ h)(\vec{x}) = Df((g \circ h)(\vec{x})) \cdot Dg(h(\vec{x})) \cdot Dh(\vec{x})$

Lagrange Multiplier Example

Find max/min of $f(x,y,z) = 6xy + z^3$ subject to the constraint $g(x,y,z) = x^2 + y^2 + z^2 = 36$.

① $\nabla g = 0 \to (2x, 2y, 2z) = 0 \to x = y = z = 0$ $\uparrow$ violates constraint

② $\nabla f = \lambda \nabla g$

$\nabla f = \begin{pmatrix} 6y \\ 6x \\ 3z^2 \end{pmatrix} = \lambda \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} \to$ $6y = \lambda 2x$, $6x = \lambda 2y$, $3z^2 = \lambda 2z$

1) cross multiply eqns

$6y \cdot \lambda 2y = 6x \cdot \lambda 2x$

case 1: $\lambda = 0 \to \nabla f = 0$
$x, y, z = 0 \to$ violates constraint

case 2: $\lambda \neq 0$
$y^2 = x^2$ keep this equation

$6x\lambda 2z = 3z^2 \lambda 2y$

case 1: $\lambda = 0$

case 2:
$2xz = z^2 y$ keep this

2) solve system

$y^2 = x^2 \to y = \pm x$
$2xz = z^2 y$
$x^2 + y^2 + z^2 = 36$

case 1:
$y = x = 0$
$0^2 + 0^2 + z^2 = 36$
$z = \pm 36$

case 2:
$y = x \neq 0$
$2z = z^2$
- $z = 0$: $x^2 + x^2 = 36$, $x^2 = 18$, $x = \pm 3\sqrt{2}$
- $z = 2$: $x^2 + x^2 = 32$, $x^2 = 16$, $x = \pm 4$

case 3:
$y = -x \neq 0$
$-2z = z^2$
- $z = 0$: $x^2 + x^2 = 36$, $x^2 = 18$, $x = \pm 3\sqrt{2}$
- $z = -2$: $x^2 + x^2 = 32$, $x^2 = 16$, $x = \pm 4$

3) Plug into f

$f(0, 0, -36) = -216$ min
$f(0, 0, 36) = 216$ max
$f(\pm 3\sqrt{2}, \pm 3\sqrt{2}, 0) = 108$
$f(\pm 4, \pm 4, -2) = 104$
$f(\pm 3\sqrt{2}, \mp 3\sqrt{2}, 0) = -108$
$f(\pm 4, \mp 4, 0) = -104$

Lecture 13: Linear Functions & Matrices

- Affine: $f(\vec{x}) = a_1 x_n + \ldots + b$ $(f(0) = b$, has constants$)$
- Linear: $f(\vec{x}) = a_1 x_n + \ldots$ $(f(0) = 0$, all components linear$)$
- Linear function: $f(\vec{x}) = \underset{m \times n}{A}\,\underset{n \times 1}{\vec{x}}$
- Affine function: $f(\vec{x}) = \underset{m \times n}{A}\,\underset{n \times 1}{\vec{x}} + \underset{m \times 1}{\vec{b}}$
- Transformation $T(\vec{x}) = A\vec{x}$, find A by getting columns through standard basis
  $A = \begin{bmatrix} | & & | \\ T(e_1) & \cdots & T(e_n) \\ | & & | \end{bmatrix}$ apply T to e_1 to get 1^{st} col
- Derivative/Jacobian Matrix
  $Df = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots \\ \partial f_2/\partial x_1 & \ddots & \\ \vdots & & \ddots \end{bmatrix} = \begin{bmatrix} - \nabla f_1 - \\ - \nabla f_2 - \\ \vdots \\ - \nabla f_n - \end{bmatrix}$ row is gradient represented horizontally
- Linear Approximation:
  $f(\vec{x}) \approx f(\vec{a}) + Df(\vec{a})(\vec{x} - \vec{a})$ for $\vec{x}$ near $\vec{a}$
  $f(\vec{a} + \vec{h}) \approx f(\vec{a}) + Df(\vec{a})\vec{h}$ for $\vec{h}$ near $\vec{0}$

Lecture 15: Matrix Algebra

- Diagonal: $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$
- Lower/Upper Triangular: all entries above/below diagonal are zero
- Properties:
  $A(B+C) = AB + AC$
  $A(BC) = (AB)C$
  $A(cB) = (cA)B = c(AB)$
  $I_m A = A = A I_n$
- $AB = \vec{0}$ does not mean A and B is zero matrix
- $AB = AC$, B not always C
- $(AB)\vec{x} = A(B\vec{x})$
- $(A+B)C = AB + BC$
- Diagonal Matrix:
  $\hookrightarrow$ only non-zero at diag.
  $\begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ $\begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix}$
- Multiplying by diag matrix
  AD: scales cols by diag.
  DA: scales rows by diag.

Derivatives (a is a scalar)

$\frac{d}{dx}(au) = a\frac{du}{dx}$

$\frac{d}{dx}(a^x) = a^x \ln a$

$\frac{d}{dx}(\ln(u)) = \frac{1}{u}\frac{du}{dx}$

$\frac{d}{dx}(\log_a(u)) = \frac{1}{u \ln a}\frac{du}{dx}$