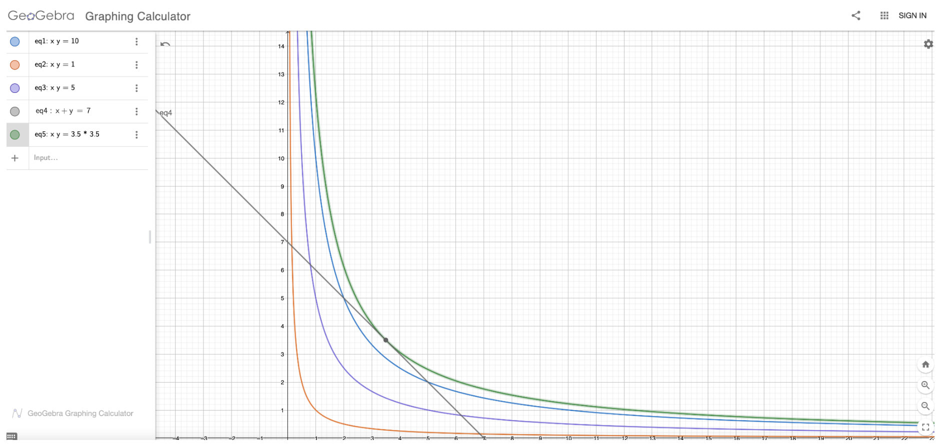
CAS CS 365
Lec 24

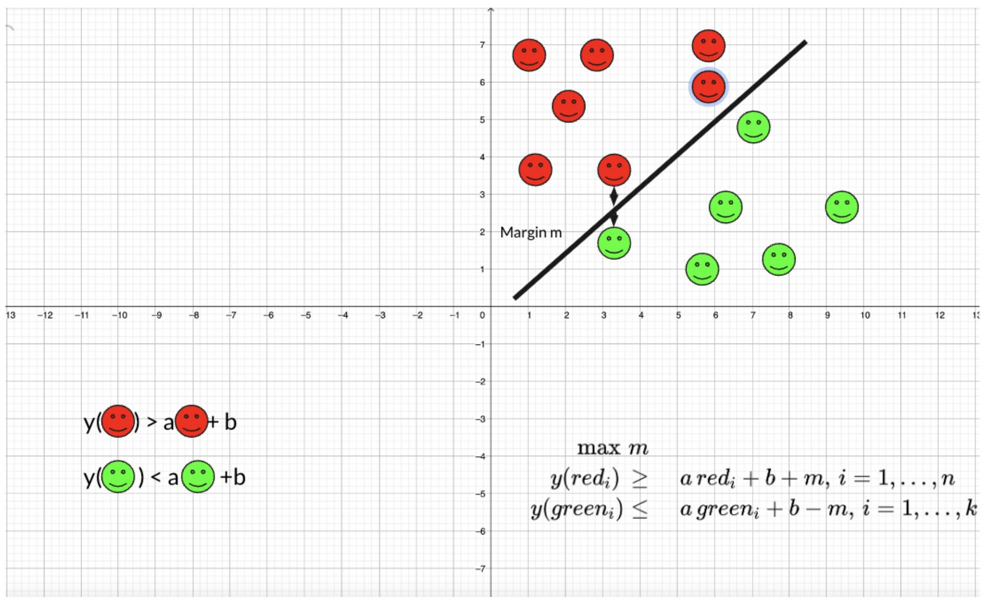
Optimization

1. Optimal solution a = b = l/4 (h= b)
    a. 

    b. The line is set of all rectangles
    c. We want to find the optimal solution that maximizes the area
    d. All the curves are xy = c, where c is area
    e. The point where the two curves intersect and are tangent to each other is the maximum area
2. Transportation problem
    a. Minimize the cost of goods transported from
        i. A set of m sources to…
        ii. … a set of n destinations
            1. Subject to the supply and demand of the sources and destination respectively
    b. Given:
        i. $a_1,\ldots,a_m$: units to transfer from sources
        ii. $b_1,\ldots,b_m$: units to receive by destinations
        iii. c_{ij}: cost of transferring a unit from source i to destination j
    c. Find the auntities x_{ij} to be transferred from source i to destination j for i = 1,...,m and j = 1,...,n

d.
$$
\begin{aligned}
\min \ & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\
& \sum_{j=1}^{n} x_{ij} = a_i\,, \quad i = 1,\dots,m \\
& \sum_{i=1}^{m} x_{ij} = b_j, \quad j = 1,\dots,n \\
& x_{ij} \geq 0
\end{aligned}
$$

3. A (not so) Toy ML problem

   a. 

   b. Try to make simplest vector that separates (model selection)

4. Minimization

   a. $$\min_{x \in F} \ f(x)$$

   b. Let $f:R^n \rightarrow R$

      i. When $F = R^n$, the optimization is unconstrained

      ii. When $F = \{x \in \mathbb{R}^n : \ h(x) = 0,\ g(x) \leq 0\}$ where h: $R^n \rightarrow R^m$, g: $R^n \rightarrow R^k$ are real functions, the problem is called constrained

Lec 24

Inflection point
Saddle point

$f(x) = x^3$
↳ $x=0$ is saddle point

$\nabla f(x) = 0$

A → positive semi-definite if $x^T A x \geq 0$ for all $x \in \mathbb{R}^n$
and $\lambda_1, \dots, \lambda_n \geq 0$

ex) $\nabla f = [6(x_1 - x_2) + 3(x_1 + x_2)^2, -6(x_1 - x_2) + 3(x_1 + x_2)^2] = 0$

$H_f = \frac{df^2}{dx^2} = \begin{bmatrix} 6 + 6(x_1 + x_2) & -6 + 6(x_1 + x_2) \\ -6 + 6(x_1 + x_2) & 6 + 6(x_1 + x_2) \end{bmatrix}$

$H_f(0,0) = \begin{bmatrix} 6 & -6 \\ -6 & 6 \end{bmatrix}$, which is positive semidefinite

ex $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ → rank of 1
↳ one of λ is 0

$\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
↳ $\lambda = 2$

⇒ λ is 0 and 2

$(x_1 \ x_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$

$x_1 = -x_2$

ex) $\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
↳ $\lambda = 0$

$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$
↳ $\lambda = 2$

back to

$$H_f(0,0) = \begin{bmatrix} 6 & -6 \\ -6 & 6 \end{bmatrix}$$

↳ this doesn't mean it's local min

- It can't be local min if we can go further down

Consider the eigenvector $= \begin{bmatrix} 1 \\ 1 \end{bmatrix} = z$ of $H_f(x^*)$ with eigenvalue 0.

$f(x^* + az) = 8a^3$ (so for $a<0$, we get a value $< f(x^*)$)

This is why we call these conditions necessary

---

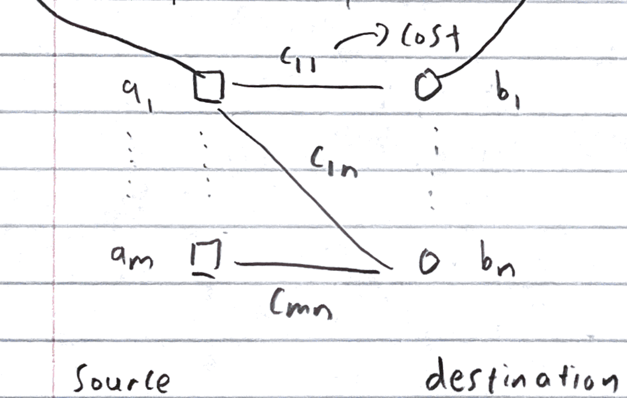
the amount of units it sends away

$x_{11} + x_{12} + \dots + x_{1n} = a_1$

Transportation problem

the total # of units it receives

$x_{11} + x_{21} + \dots + x_{m1} = b_1$



$\Rightarrow \min \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$

(now we have constraints)

$\sum_{j=1}^{n} x_{ij} = a_i, \ i = 1, \dots, m$

$\sum_{i=1}^{m} x_{ij} = b_j, \ j = 1, \dots, n$

$x_{ij} \geq 0$

↳ to be feasible, $\sum a_i = \sum b_j$

if $a_i < b_j$ (more demand), no solution

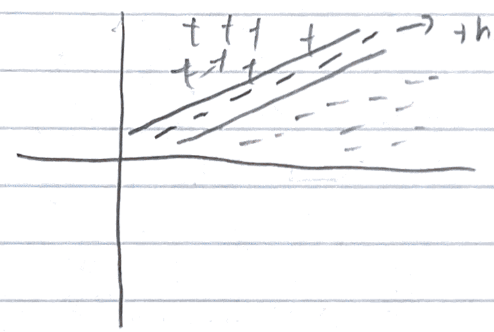
if $a_i > b_j$ (more source), it should be $\sum x_{ij} \leq a_i$

- Toy ML problem

$\exists a,b$ such that $y_1 \geq ax+b$ if (x,y) is green

$y < ax+b$ if (x,y) is red

maximizes the margin (equally far away from closest red/green points)



→ the dotted line is equally far away (balance)

Practice

what is the best fit function of the following form that passes through the given points

$$f(x) = A\cos x + B\sin x + C\cos(2x) + D$$

$$\sum_{i=1}^{n} (y_i - A\cos x - B\sin x - C\cos(2x) - D)$$

$$\min_x \|y - Ax\|_2^2$$

↳ $x = \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}$

assuming first point is $x = -5.27$

↳ $A = \begin{bmatrix} \cos(-5.27) & \sin(-5.27) & \cos(2 \cdot -5.27) & 1 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$

$$\|Ax - y\|_2^2$$

↳ first point is 2.28 when $x = -5.27$

↳ $\left\| \begin{bmatrix} 2.28 \\ \vdots \end{bmatrix} \right\|_2^2$

$\nabla f(x^*) = 0$, but Hessian has both positive and negative λ
$\Rightarrow x^*$ is saddle point

$f(x,y) = x^2 - y^2$
$\nabla f(x,y) = 0 \Rightarrow (2x, -2y) = (0,0)$
$H_f = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$
$\hookrightarrow \lambda = -2, 2 \Rightarrow$ saddle point

matrix with $\lambda = 0, -2, 2$
$\hookrightarrow \mathbb{R}^3 : \{u_1, u_2, u_3\}$

$A = 2u_1u_1^T + (-2)u_2u_2^T + (0)u_3u_3^T$

x^TAx