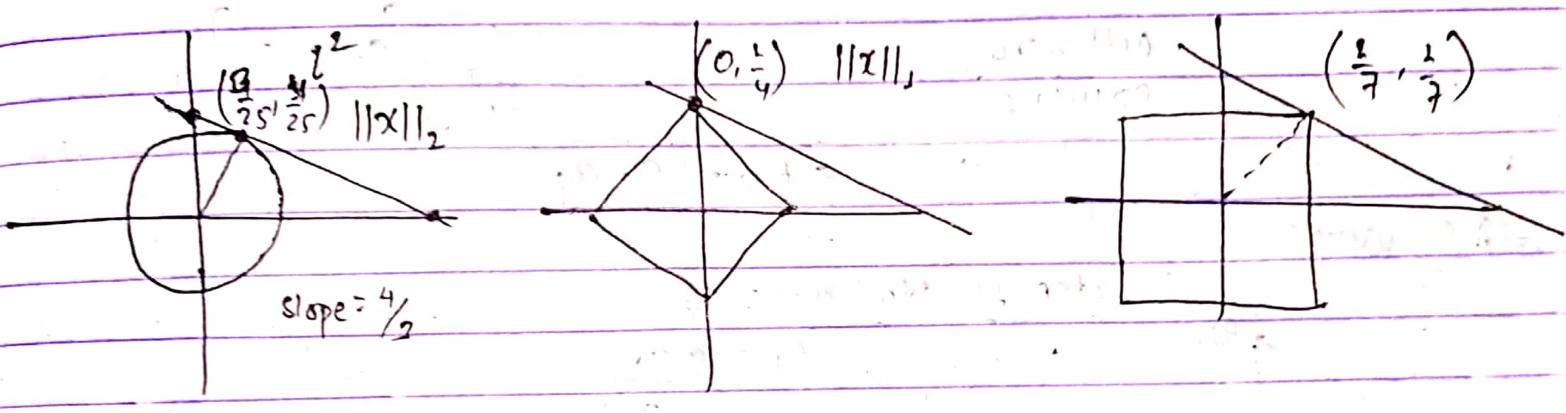
Lec. 11 : Minimization and Gram-Schmidt

Minimize $||x||_1$, $||x||_2$ and $||x||_\infty$ with respect to

$$3x_1 + 4x_2 = 1.$$

The unit balls for different norms are



You expand/shrink unit balls to become tangent to the constraint line. We've encountered this in previous units

Project: Visualization in 2d and 3d. for different constraints

Scanned with CamScanner

Gram-Schmidt:

① Usual way.

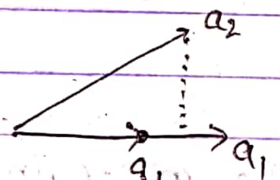
$$A = QR \Rightarrow R = Q^{-1}A = Q^{T}A$$

orthogonal columns

$$\begin{bmatrix} q_1^T \\ \vdots \\ q_n^T \end{bmatrix} \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}$$

$$\therefore R_{ij} = q_i^T a_j$$

18.06 way:



step 1: Normalize a_1

$$q_1 = \frac{a_1}{\|a_1\|}$$

step 2: Subtract the component of a_2 along q_1 from a_2 to obtain the component of a_2 perpendicular to q_1

i.e. $A_2 = a_2 - (a_2^T q_1)\, q_1$

$$x\vec{q}_1 + \vec{e} = \vec{a}_2$$

$$x\, q_1^2 = \vec{a}_2 \cdot \vec{q}_1$$

$$\therefore x = \frac{a_2^T q_1}{\|q_1\|}$$

$$\therefore x = a_2^T q_1 \quad \left[q_1 \text{ is unit vector} \right]$$

Step 3: Normalize A_2

$$q_2 = \frac{A_2}{\|A_2\|}$$

For the vectors that follow,

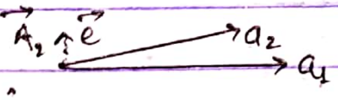
$$A_k = a_k - \left[(a_k^T q_1)\, q_1 + \cdots (a_k^T q_{k-1})\, q_{k-1} \right]$$

$$q_k = \frac{A_k}{\|A_k\|}$$

Rich:

Gram-Schmidt is applied in the usual order, at least in theory. So, we allow column exchanges in this method.

- Suppose a_1 is of descent size
- a_2 could be close to a_1



- $\vec{e}$ is too small, $q_2 = \dfrac{A_2}{\|A_2\|}$ could blow up because $\|A_2\|$ is too small

So,

$$\left.\begin{aligned} A_2 &= a_2 - (a_2^T q_1) \cdot q_1 \\ &= a_3 - (a_3^T q_1)\, q_1 \\ &\ \vdots \\ &= \underbrace{a_n - (a_n^T q_1)\, q_1}_{\text{we would have to compute these anyway.}} \end{aligned}\right\}\ \text{Choose the } A_2 \text{ that is largest}$$

Project: Implement in code.

KRYLOV

$Ax = b$

big and sparse matrix

Compute $b, Ab, A^2 b, \ldots A^{j-1} b$

= The combinations give krylov's space K_j

x_j = Best vector [closest vector in Krylov's space]

What if these are nearly dependent?

→ orthogonalize [Arnoldi's method].

Scanned with CamScanner

- Why is orthogonal (normal) basis great?

$$x = QC = c_1 q_1 + c_2 q_2 + \cdots c_n q_n = \begin{bmatrix} Q \end{bmatrix} \begin{bmatrix} c \end{bmatrix}$$

$$C = Q^{-1} x = Q^{T} x \quad \text{[for orthogonal basis]}$$

Arnoldi's algorithm,

Scanned with CamScanner