Discussion #4

Today: Gradient Methods ☺

Re: HW

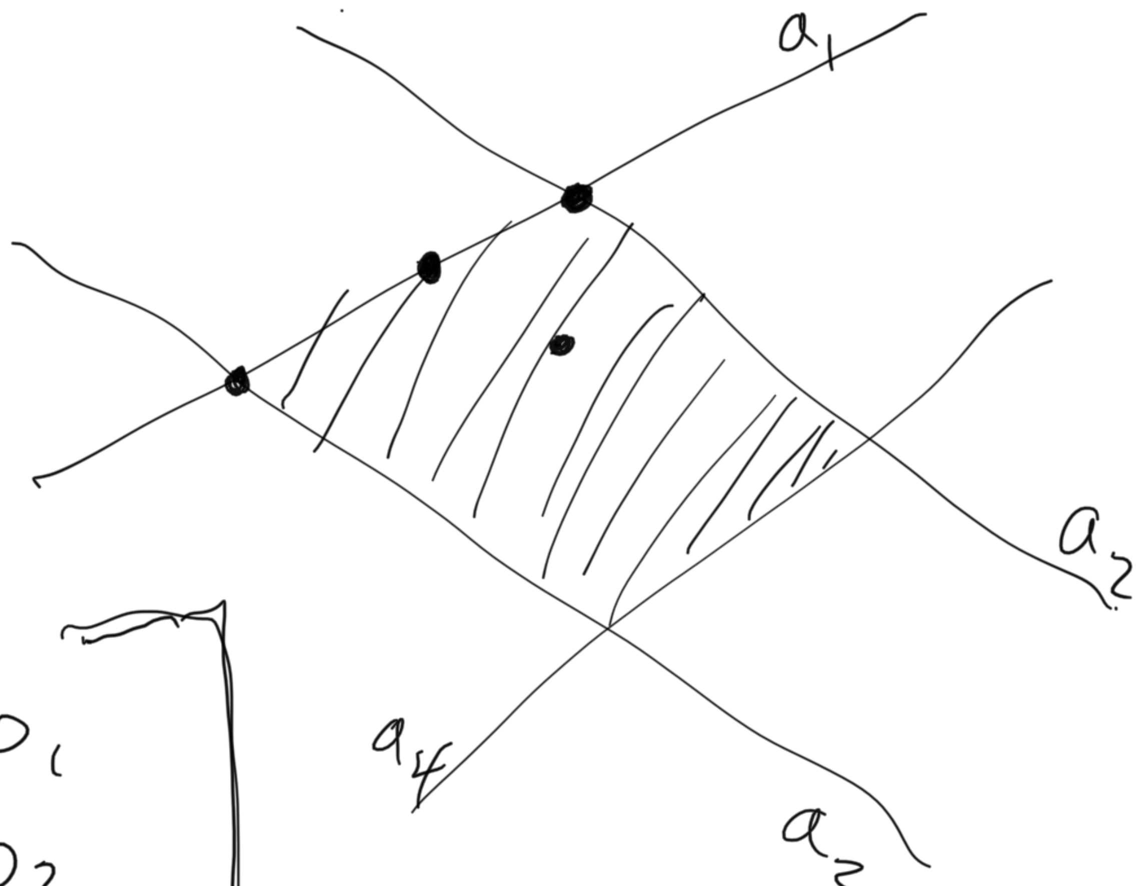
1. $S = \{x : \boxed{Ax \leq b}\}$

a) $x \in int(S)$, $Ax < b$

→ all directions feasible.

b) $x \in \partial S$,

$\partial S \neq \{x : Ax = b\}$



$Ax \leq b$

$a_1^T x \leq b_1$

$a_2^T x \leq b_2$

$\vdots$

$a_m^T x \leq b_m$

$a_4^T x = b_4$

$$\partial S = \{x : a_i^T x = b_i, \text{ for 1 or more } i \text{ in } 1, \dots, n\}$$

$x \in \partial S$, For all i such that $a_i^T x = b_i$

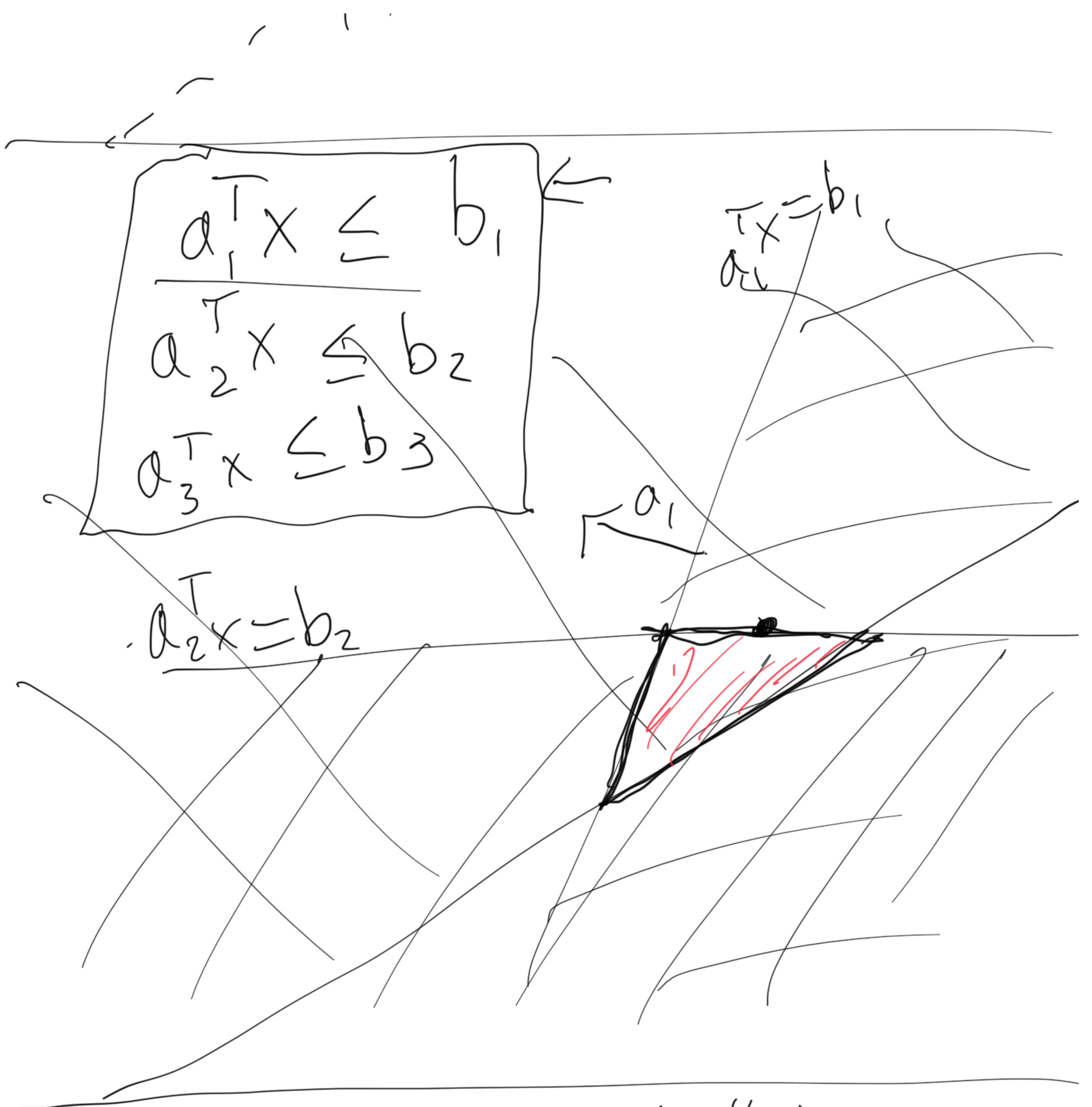
$$\underbrace{a_i^T (x}_{b_i} + \alpha d) \le b_i$$

$$\Rightarrow a_i^T d \le 0$$

$$Ad \le 0, \quad a_i^T d \le 0 \quad \cancel{\forall i = 1, \dots, n}$$

$i \in 1, \dots, n$
such that $a_i^T x = b$

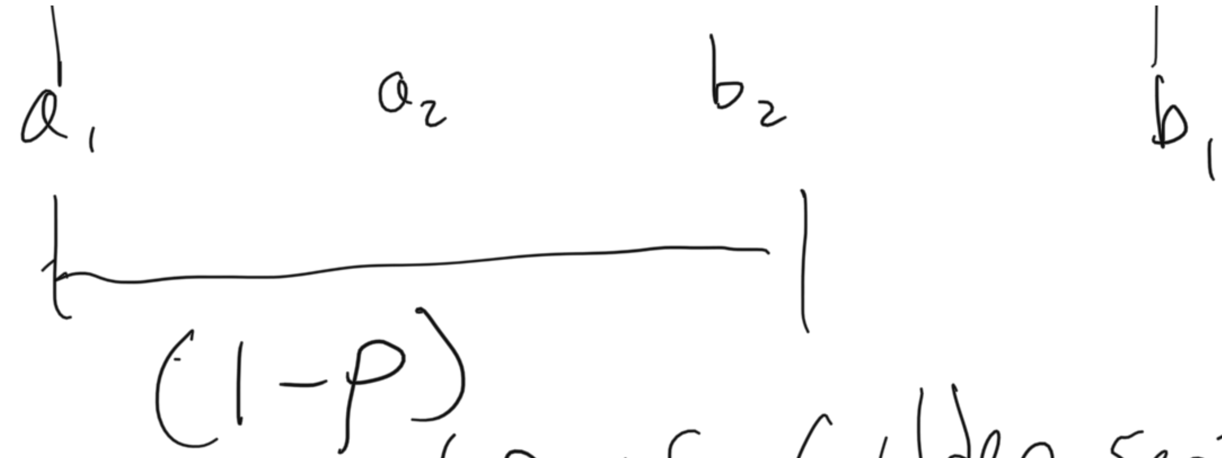
$$\boxed{\begin{aligned} a_1^T x &\le b_1 \\ a_2^T x &\le b_2 \\ a_3^T x &\le b_3 \end{aligned}}$$



#2 Fibonacci Method

Vs. Golden section

We want to discuss size of interval after N steps:



After every step of Golden section, interval decreases by a factor of $1-\rho$

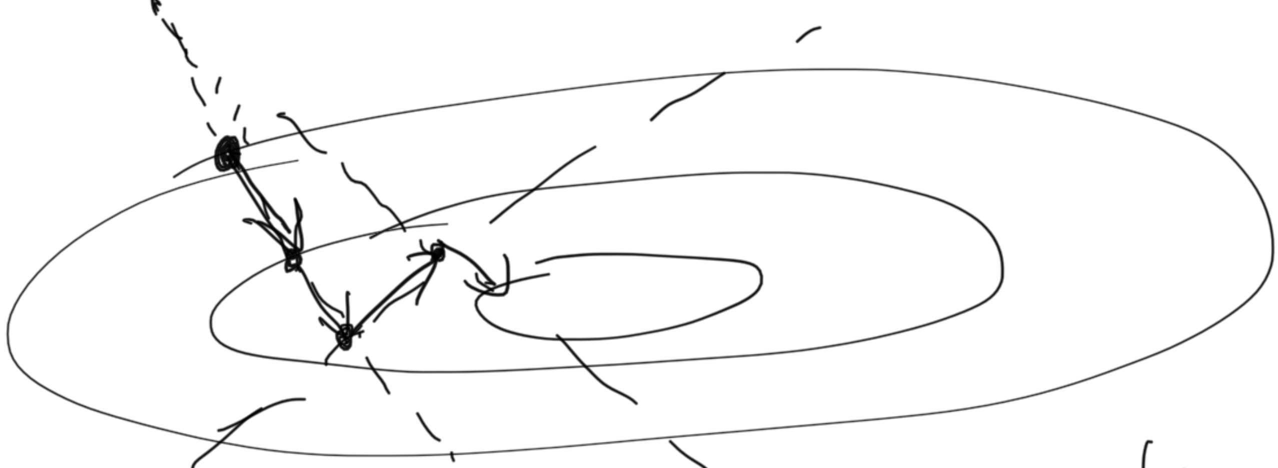
$\rho = \frac{3-\sqrt{5}}{2}$. After N steps, $(1-\rho)^N$

Thm: After N steps, Fibonacci Search reduces interval by $\frac{1+2\varepsilon}{F_{n+1}}$

N	Golden section	Fibonacci
1	$1-\rho$	$\frac{1+2\varepsilon}{F_2}$
2	$(1-\rho)^2$	$\frac{1+2\varepsilon}{F_3}$
3	$(1-\rho)^3$	$\frac{1+2\varepsilon}{F_4}$
4	$\boxed{(1-\rho)^4}$	$\frac{1+2\varepsilon}{F_5}$

	0.5	0.4
2	0.25	$\boxed{0.15}$
3	$\boxed{0.125}$	0.05

Gradient Methods



$$\boxed{f(x) = \frac{1}{2}x^T Q x - b^T x}$$

$$x_{k+1} = x_k - \boxed{\alpha_k} \nabla f(x_k)$$

How to choose

Stepsize:

1) $\alpha_k = \arg\min_{\alpha} f(x_k - \alpha \nabla f(x_k))$

2) α_k = const.

Tip for Analyzing Problems:
See How the error at each iteration develops as the algorithm proceeds.

$$\nabla f(x_k) = Qx - b = 0$$

$\Rightarrow$ solution is $x^* = Q^{-1}b$

Error at each iteration:

$$e_k = x_k - x^*$$

Note $Qe_k = Qx_k - Qx^*$

$$= Qx_k - b$$

$$= \nabla f(x_k)$$

For what α will fixed descent converge?

converge.

$$x_{k+1} = x_k - \alpha \underline{\nabla f(x_k)}$$

$$\underbrace{x_{k+1} - x^*}_{e_{k+1}} = \underbrace{x_k - x^*}_{e_k} - \alpha \underbrace{\nabla f(x_k)}_{Q e_k}$$

$$\underline{e_{k+1}} = e_k - \alpha Q e_k$$

$$= \boxed{(I - \alpha Q)} \underline{e_k}$$

$\vdots$

$$= (I - \alpha Q)^{k+1} e_0$$

Need $\boxed{(I - \alpha Q)^{k+1}} \to 0$

as $k \to \infty$

Need eigenvalues of $I - \alpha Q$

to have norm < 1

eigenvalues are $1 - \alpha \lambda_i$

eigenvalues

where λ_i is an eigenvalue of Q.

Need $|1-\alpha\lambda_i| < 1 \quad \forall i$

Note $Q > 0$, so $\lambda_i > 0 \quad \forall i$,

If $\alpha \leq 0$, $1-\alpha\lambda_i \geq 1$

$-1 < 1-\alpha\lambda_i < 1$

$\alpha\lambda_i < 2$

$\Leftrightarrow \alpha < \frac{2}{\lambda_i} \quad \forall i$

$$\boxed{0 < \alpha < \frac{2}{\lambda_{max}}}$$

When is it possible for gradient descent to converge in

1 step:

$$\underset{-x^*}{x_{k+1}} = \underset{-x^*}{x_k} - \alpha_k \nabla f(x_k)$$

$$\underset{\overset{\|}{0}}{e_{k+1}} = e_k - \alpha_k \underline{Q} e_k$$

$$0 = e_k - \alpha_k Q e_k$$

$$\underline{Q} e_k = \frac{1}{\alpha_k} e_k$$

Need e_k to be an eigenvector of Q