Necessary condition KKT:
if x^*, λ^*, ν^* optimal & strong duality & diff'able then KKT holds
Sufficient condition: if $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ satisfy KKT + convex + diff'able, then $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ are optimal
ie objective convex, domain convex, f_i convex, h_i affine

Today:

- KKT sufficiency proof
- linear programs
- optima at extreme points of polyhedron

KKT conditions (sufficient)

- Convex problem (f_i's convex, h_i's affine, $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ that satisfy KKT), differentiable

primal feasibility
$$f_i(\tilde{x}) \le 0 \quad i=1,\dots,m$$
$$h_i(\tilde{x}) = 0 \quad i=1,\dots,p$$
dual feasibility
$$\tilde{\lambda}_i \ge 0 \quad i=1,\dots,m$$

comp. slackness:
$$\tilde{\lambda}_i f_i(\tilde{x}) = 0 \quad i=1,\dots,m$$

stationarity
$$\nabla f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^p \tilde{\nu}_i h_i(\tilde{x}) = 0$$

$\Rightarrow$ $\tilde{x}$ primal optimal
$\tilde{\lambda}, \tilde{\nu}$ dual optimal

$$p^* = \min f_0(\vec{x})$$
s.t.
$$f_i(\vec{x}) \le 0 \quad i=1,\dots,m$$
$$h_i(\vec{x}) = 0 \quad i=1,\dots,p$$
$$d^* = \max_{\vec{\lambda} \ge 0} g(\vec{\lambda}, \vec{\nu})$$

if KKT, then optimal. previously we proved if cond. then KKT

$\hookrightarrow$ want to prove that this is true (if we have these conditions holding, we get the optimal point)

Proof

Ayah

FOC: $\nabla f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^p \tilde{\nu}_i h_i(\tilde{x}) = 0$

Consider:
$$L(x, \tilde{\lambda}, \tilde{\nu}) = f_0(x) + \sum_{i=1}^m \tilde{\lambda}_i f_i(x) + \sum_{i=1}^p \tilde{\nu}_i h_i(x)$$

$\tilde{\lambda} \ge 0$, f_i's convex, h_i's affine $\Rightarrow$ $L(x, \tilde{\lambda}, \tilde{\nu})$ is a convex fcn of x

$\therefore$ if $\tilde{x}$ satisfies FOC (gradient), then it must be a minimizer

KKT $\Rightarrow \nabla_x L(x, \tilde{\lambda}, \tilde{\nu})\Big|_{x=\tilde{x}} = 0 \Rightarrow \tilde{x}$ is a minimizer of $L(x, \tilde{\lambda}, \tilde{\nu})$

↳ Q/ what can we say about the dual?

$$\underbrace{g(\tilde{\lambda}, \tilde{\nu})}_{\text{lower bound for } \forall \lambda, \nu \text{ (proved previously)}} = \min_x L(x, \tilde{\lambda}, \tilde{\nu})$$

$$= \min_x \left[f_0(x) + \sum \tilde{\lambda} f_i(x) + \sum \tilde{\nu}_i h_i(x) \right]$$

$$= f_0(\tilde{x}) + \underbrace{\sum \tilde{\lambda}_i f_i(\tilde{x})}_{0} + \underbrace{\sum \nu_i h_i(\tilde{x})}_{0}$$

↳ b/c this $= f_0(\tilde{x})$

$$g(\tilde{\lambda}, \tilde{\nu}) = f_0(\tilde{x})$$

↳ says that the duality gap btwn primal & dual optimal is 0 $\Rightarrow$ strong duality must hold

$\tilde{x}$, $(\tilde{\lambda}, \tilde{\nu})$ have 0 duality gap

$p^* \geq \underbrace{g(\tilde{\lambda}, \tilde{\nu})}_{}$ lower bound on p^* — not necessarily a unique min

$\Rightarrow p^* = f_0(\tilde{x}) \Rightarrow d^* = g(\tilde{\lambda}, \tilde{\nu}) = f_0(\tilde{x})$

↳ uses certificate property of duality

↳ if you hit a dual point, you must reach the minimum bc the dual $g(\tilde{\lambda}, \tilde{\nu})$ is the lower bound on p^*

- slater's condition doesn't necessarily hold if KKT holds

Linear programs

↳ nice to get optimization problems in LP form bc we can solve them (in the avg case) in polynomial time using the simplex algorithm

$$\min \ \vec{c}^T \vec{x}$$
$$\text{s.t. } A\vec{x} \leq \vec{b}$$

- equality constraint → inequality constraint

$$\vec{a}_i^T \vec{x} = b \rightarrow \vec{a}_i^T \vec{x} \geq \vec{b} \ \wedge \ \vec{a}_i^T \vec{x} \leq \vec{b}$$

$$\vec{a}_i^T \vec{x} \geq \vec{b} \rightarrow -\vec{a}_i^T \vec{x} \leq -\vec{b}$$

Standard Form of an LP

→ typically used for simplex algo

$$\min \ \vec{c}^T \vec{x}$$
$$A\vec{x} = \vec{b}$$
$$\vec{x} \geq 0$$

- all LPs can be translated into standard form:

① Eliminate inequality constraints (by adding slack variables)

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \rightarrow \sum_{j=1}^{n} a_{ij} x_j + s_i = b_i \ ; \ s_i \geq 0$$

Q/ what if we have unconstrained variables?

② How to get $x_i \geq 0$, $\forall x_i$?

x_i unconstrained

↳ can rewrite as diff btwn 2 positive things:

$$x_i = x_i^+ - x_i^- \quad \text{where } x_i^+ \geq 0 \quad x_i^- \geq 0$$

$$\min \ 2x_1 + 4x_2$$
$$x_1 + \underline{x_2} \geq 3$$
$$3x_1 + \underline{2x_2} = 14$$
$$x_1 \geq 0$$

no constraint on x_2

→ ① convert to equality:

$$x_1 + x_2 - \underbrace{x_3}_{\text{slack variable}} = 3 \qquad x_3 \geq 0$$

② $x_2 = x_2^+ - x_2^-$

rewrite

$$\min \ 2x_1 + 4x_2^+ - 4x_2^-$$
$$x_1 + x_2 - x_3 = 3$$
$$3x_1 + 2x_2^+ - 2x_2^- = 14$$
$$\left.\begin{array}{l} x_1 \geq 0 \\ x_2^+ \geq 0 \\ x_2^- \geq 0 \\ x_3 \geq 0 \end{array}\right\} \text{dimensionality increase} \Rightarrow \text{computational price} \uparrow$$

Graphical Representation of LPs & their Duals

minimize $-x_1 - x_2$

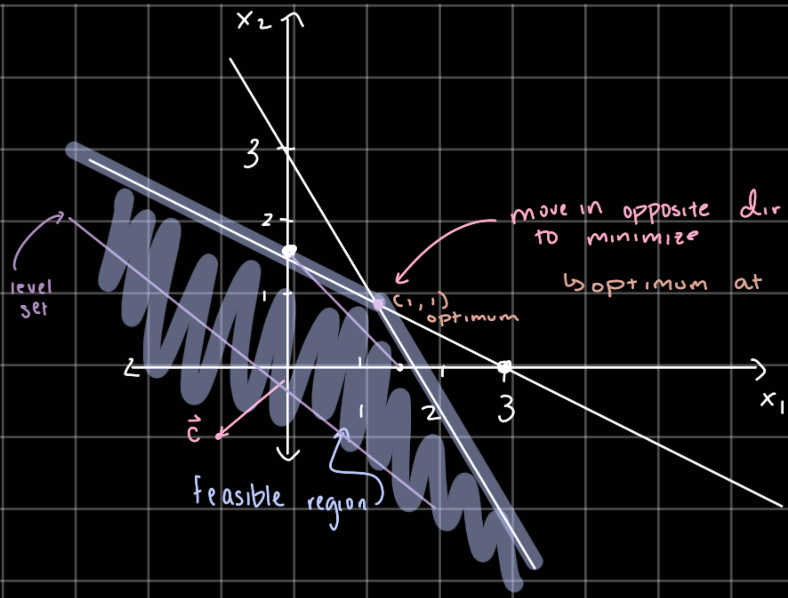
s.t. $x_1 + 2x_2 \leq 3$ (1)

$2x_1 + x_2 \leq 3$ (2)

$\xrightarrow{\text{standard form}}$

$$\min \begin{bmatrix} -1 \\ -1 \end{bmatrix}^T \vec{x}$$

s.t. $$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \vec{x} \leq \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$



$-x_1 - x_2 = -1$
$-x_1 - x_2 = 0$
$-x_1 - x_2 = 1$ } level sets ↳ orthogonal to $\vec{c}$

→ generally, optimum point of LPs at vertex of bounding polyhedron

Dual:

$$\max \; -b^T\lambda$$
$$A^T\lambda + c = 0$$
$$\lambda \geq 0$$

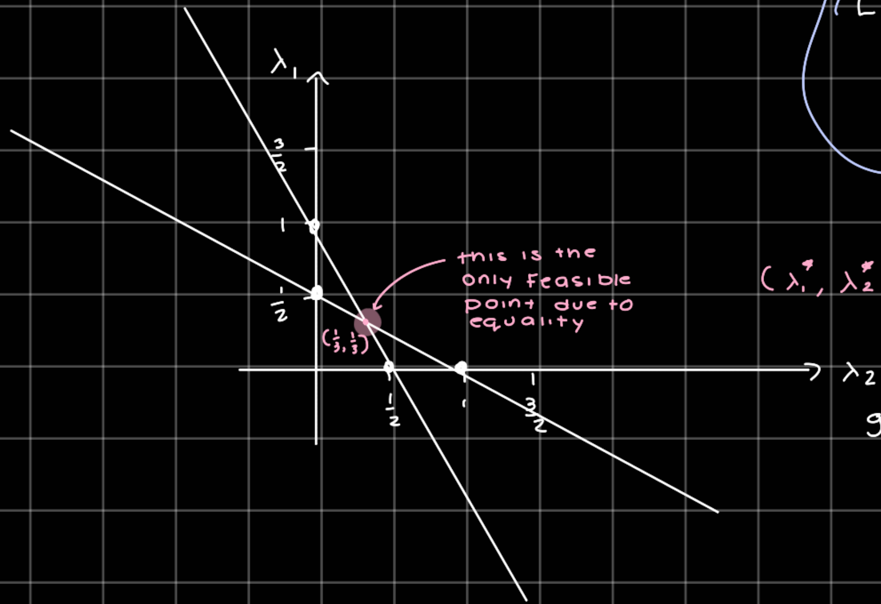
$\xrightarrow{\text{the dual's}}$

$$\max \; -\begin{bmatrix} 3 \\ 3 \end{bmatrix}^T \vec{\lambda}$$

$$\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\lambda + \begin{bmatrix} -1 \\ -1 \end{bmatrix} = 0$$

$$\vec{\lambda} \geq 0$$

$\lambda_1 + 2\lambda_2 = 1$
$2\lambda_1 + \lambda_2 = 1$



$(\lambda_1^*, \lambda_2^*) = (\frac{1}{3}, \frac{1}{3})$

$$g(\lambda^*) = -\begin{bmatrix} 3 \\ 3 \end{bmatrix}^T \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

$$= -2$$

→ if we were given the dual & we wanted to find (x_1, x_2) then we would use compl slackness → have equality so we know equality holds at this point.

*? .... active constraints....?

Proof (optimizing LPs at their vertices)

Definition:

Polyhedron: set $\{\vec{x} \in \mathbb{R}^n \mid A\vec{x} - \vec{b}\}$ ↳ set created by linear constraints

$A \in \mathbb{R}^{n \times n}$

$\vec{b} \in \mathbb{R}^m$

↳ "standard form" (using slack variables):

set $\{\vec{x} \in \mathbb{R}^\ell \mid C\vec{x} = \vec{d} \; ; \; \vec{x} \geq \vec{0}\}$

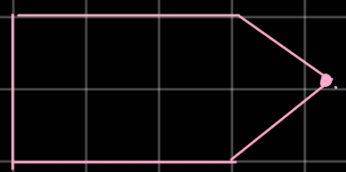
↳ dimensionality might change due to adding slack variables

Extreme point (of a polyhedron P)

$x \in P$ is an extreme point (vertex of P) if we cannot find 2 vectors $\vec{y}, \vec{z} \neq \vec{x}$ $\quad \vec{y}, \vec{z} \in P$; $\lambda \in [0, 1]$

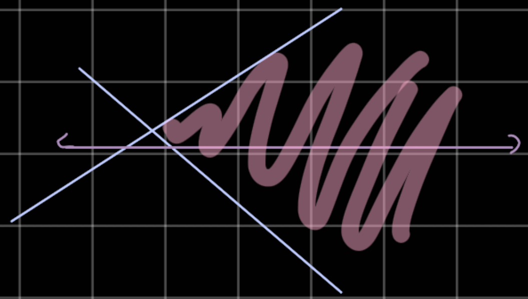
s.t. $\vec{x} = \lambda \vec{y} + (1-\lambda)\vec{z}$

↳ ie that $\vec{x}$ is a convex combo of them.

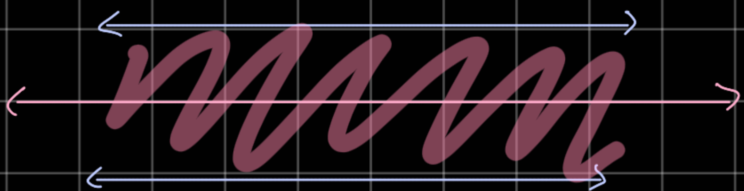


↖ need a point here to avg out but not in polyhedron ↳ this is a vertex

P has an extreme point IFF P doesn't contain a line.



← no way an infinite line can fit into this interior



→ this works, but no extrema

$\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ satisfy KKT

① $\tilde{x}$ is minimizer of $L(x, \tilde{\lambda}, \tilde{\nu})$

↳ convex

↳ $\nabla_x L \big|_{x=\tilde{x}} = 0$

② $\underbrace{g(\tilde{\lambda}, \tilde{\nu})}_{\text{dual feasible}} = \min_x L(x, \tilde{\lambda}, \tilde{\nu})$

$= L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

$= \underbrace{f_0(\tilde{x})}_{\text{primal feasible}}$

$p^* \geq g(\tilde{\lambda}, \tilde{x})$ } holds by weak duality

↳ $\forall g(\lambda, \nu)$ is a lower bound

"every dual feasible point gives you a lower bound on the primal)

$f_0(\tilde{x}) \geq p^*$

↳ $p^* = \min_{\vec{x} \in \text{domain}} f_0(\vec{x})$

→ p^* is 5 or smaller

↳ For $\forall$ points, $f_0(\tilde{x}) \geq p^*$ $\forall \tilde{x}$

$$p^* \leq g(\tilde{\lambda}, \tilde{\nu}) \quad \forall \tilde{\lambda}, \tilde{\nu}$$

$$g(\tilde{\lambda}, \tilde{\nu}) \leq p^* \leq f_0(\tilde{x}) = g(\tilde{\lambda}, \hat{\nu})$$

$$p^* = f_0(\hat{x}) = g(\hat{\lambda}, \hat{\nu})$$

↳ if duality gap is 0, can never do better than this point