Operations Research
07. Duality

Lagrangian

- **standard form problem** (not necessarily convex)

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leqslant 0, \quad i = 1, 2, \ldots, m \\ & h_i(x) = 0, \quad i = 1, 2, \ldots, p \end{array}$$

  variable $x \in \mathbb{R}^n$, domain $\mathcal{D}$, optimal value $p^\star$

- **Lagrangian** $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$, with $\operatorname{dom} f = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

  - weighted sum of objective and constraints
  - λ_i is Lagrange multiplier associated with $f_i(x) \leqslant 0$
  - ν_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrange Dual Function

- **Lagrange dual function**: $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$,

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

- g is concave, can be $-\infty$ for some λ, ν

- **lower bound property**: $g(\lambda, \nu) \leqslant p^\star$ if $\lambda \succcurlyeq 0$

- proof: if $\tilde{x}$ is feasible and $\lambda \succcurlyeq 0$, then

$$f_0(\tilde{x}) \geqslant L(\tilde{x}, \lambda, \nu) \geqslant \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = g(\lambda, \nu)$$

- minimizing over all feasible $\tilde{x}$ gives $p^\star \geqslant g(\lambda, \nu)$

Least-Norm Solution of Linear Equations

$$\begin{array}{ll} \text{minimize} & x^\top x \\ \text{subject to} & A\,x = b \end{array}$$

- Lagrangian is $L(x,\, \nu) = x^\top x + \nu^\top (A\,x - b)$
- to minimize L over x, set gradient equal to zero:

$$\nabla_x L(x,\, \nu) = 2x + A^\top \nu = 0 \implies x = -\frac{1}{2} A^\top \nu$$

- plug x into L to obtain

$$g(\nu) = L\Big(-\frac{1}{2} A^\top \nu,\, \nu \Big) = -\frac{1}{4} \nu^\top A A^\top \nu - b^\top \nu$$

- lower bound property: $p^\star \geqslant -\frac{1}{4} \nu^\top A A^\top \nu - b^\top \nu,\ \forall\, \nu$

Least-Norm Solution of Linear Equations

$$\begin{array}{ll} \text{minimize} & x^\top x \\ \text{subject to} & A\,x = b \end{array}$$

- Lagrangian is $L(x,\,\nu) = x^\top x + \nu^\top (A\,x - b)$
- to minimize L over x, set gradient equal to zero:

$$\nabla_x L(x,\,\nu) = 2x + A^\top \nu = 0 \implies x = -\frac{1}{2}A^\top \nu$$

- plug x into L to obtain

$$g(\nu) = L\Big(-\frac{1}{2}A^\top \nu,\,\nu\Big) = -\frac{1}{4}\nu^\top A A^\top \nu - b^\top \nu$$

- lower bound property: $p^\star \geqslant -\frac{1}{4}\nu^\top A A^\top \nu - b^\top \nu,\ \forall\, \nu$

Standard Form LP

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & A\,x = b, \quad x \succcurlyeq 0 \end{array}$$

- Lagrangian is

$$L(x, \lambda, \nu) = c^\top x - \lambda^\top x + \nu^\top (A\,x - b) = -b^\top \nu + (c + A^\top \nu - \lambda)^\top x$$

- L is affine in x, so

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^\top \nu & A^\top \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

- g is linear on affine domain $\{(\lambda, \nu) \mid A^\top \nu - \lambda + c = 0\}$, hence concave

- lower bound property: $p^\star \geqslant -b^\top \nu$ if $A^\top \nu + c \succcurlyeq 0$

Lagrange Dual and Conjugate Function

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & A\,x \preceq b, \quad C\,x = d \end{array}$$

- dual function

$$\begin{aligned} g(\lambda,\nu) &= \inf_{x\in\mathrm{dom}\, f_0} \left(f_0(x) + (A^\top\lambda + C^\top\nu)^\top x - b^\top\lambda - d^\top\nu \right) \\ &= -f_0^\star(-A^\top\lambda - C^\top\nu) - b^\top\lambda - d^\top\nu \end{aligned}$$

  where $f_0^\star(y) \equiv \sup_{x\in\mathrm{dom}\, f_0} y^\top x - f_0(x)$ is **conjugate** of f_0

- simplifies derivation of dual if conjugate of f_0 is known

- **example: entropy maxmization**

$$f_0(x) = \sum_{i=1}^{n} x_i \log x_i, \qquad f_0^\star(y) = \sum_{i=1}^{n} e^{y_i - 1}$$

The Lagrange Dual Problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

- find best lower bound on $p^\star$, obtained from Lagrange dual function
- a convex optimization problem, even if original **primal** problem is not
- dual optimal value denoted by $d^\star$
- λ, ν are dual feasible if $\lambda \succeq 0$, $(\lambda, \nu) \in \operatorname{dom} g$
- often simplified by making implicit constraint $(\lambda, \nu) \in \operatorname{dom} g$ explicit

Example: Standard Form LP

- primal standard form LP

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & A\,x = b, \quad x \succcurlyeq 0 \end{array}$$

- dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda,\, \nu) \\ \text{subject to} & \lambda \succcurlyeq 0 \end{array}$$

with $g(\lambda,\, \nu) = \begin{cases} -b^\top \nu & \text{if } A^\top \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$

- make implicit constraint explicit: eliminate λ to obtain transformed dual problem

$$\begin{array}{ll} \text{maximize} & -\,b^\top \nu \\ \text{subject to} & A^\top \nu + c \succcurlyeq 0 \end{array}$$

Equality Constrained Norm Minimization

$$
\begin{array}{ll}
\text{minimize} & \|x\| \\
\text{subject to} & A\,x = b
\end{array}
$$

- dual function is

$$
g(\nu) = \inf_x \left(\|x\| - \nu^\top A\,x + b^\top \nu \right) = \begin{cases} b^\top \nu & \|A^\top \nu\|_\star \leqslant 1 \\ -\infty & \text{otherwise} \end{cases}
$$

  where $\|\nu\|_\star = \sup_{\|u\| \leqslant 1} u^\top \nu$ is dual norm of $\| \cdot \|$

- lower bound property: $p^\star \geqslant b^\top \nu$ if $\|A^\top \nu\|_\star \leqslant 1$

Two-Way Partitioning

$$\begin{array}{ll} \text{minimize} & x^\top W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, 2, \ldots, n \end{array}$$

- a nonconvex problem; feasible set contains 2^n discrete points
- interpretation: partition $\{1, 2, \ldots, n\}$ into two sets encoded as $x_i = 1$ and $x_i = -1$
- W_{ij}: cost of assigning i, j to the same set
  $-W_{ij}$: cost of assigning i, j to different sets
- dual function is

$$\begin{aligned} g(\nu) &= \inf_x \Big(x^\top W x + \sum_i \nu_i (x_i^2 - 1) \Big) = \inf_x \left(x^\top (W + \operatorname{diag} \nu) x - \mathbf{1}^\top \nu \right) \\ &= \begin{cases} -\mathbf{1}^\top \nu & \text{if } W + \operatorname{diag} \nu \succcurlyeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

- lower bound property: $p^\star \geqslant -\mathbf{1}^\top \nu$ if $W + \operatorname{diag} \nu \succcurlyeq 0$

Weak and Strong Duality

- **weak duality**: $d^\star \leqslant p^\star$
  - always holds for convex and nonconvex problems
  - can be used to find nontrivial lower bounds for hard problems, e.g. solving the SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{1}^\top \nu \\ \text{subject to} & W + \mathrm{diag}(\nu) \succcurlyeq 0 \end{array}$$

    gives a lower bound for two-way partitioning problem

- **strong duality**: $d^\star = p^\star$
  - does not hold in general
  - (usually) holds for convex problems
  - conditions that guarantee strong duality in convex problems are called **constraint qualifications**

Slater's Constraint Qualification

strong duality holds for a convex problem

$$\begin{array}{rl} \text{maximize} & f_0(x) \\ \text{subject to} & f_i(x) \leqslant 0, \quad i = 1, 2, \ldots, m \\ & A\,x = b \end{array}$$

if it is **strictly feasible**, i.e. $\exists\, x \in \operatorname{int} \mathcal{D}$ with $f_i(x) < 0$, $i = 1, 2, \ldots, m$, $A\,x = b$

- also guarantees that the dual optimum is attained (if $p^\star > -\infty$)
- can be sharpened, e.g.
  - can replace $\operatorname{int} \mathcal{D}$ with $\operatorname{relint} \mathcal{D}$ (interior relative to affine hull)
  - linear inequalities do not need to hold with strict inequality
- there are many other types of constraint qualifications

Inequality Form LP

- **primal problem**

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & A\,x \preccurlyeq b \end{array}$$

- **dual function**

$$g(\lambda,\, \nu) = \inf_x \left((c + A^\top \lambda)^\top x - b^\top \lambda\right) = \begin{cases} -b^\top \lambda & \text{if } A^\top \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

- **dual problem**

$$\begin{array}{ll} \text{maximize} & -\,b^\top \lambda \\ \text{subject to} & A^\top \lambda + c = 0, \quad \lambda \succcurlyeq 0 \end{array}$$

- from the sharpened Slater's condition: $p^\star = d^\star$ if the primal problem is feasible
- in fact $p^\star = d^\star$ except when primal and dual are both infeasible

Quadratic Program

- **primal problem** (assume $P \in \mathsf{S}^n_{++}$)

$$\begin{array}{ll} \text{minimize} & x^\top P\, x \\ \text{subject to} & A\, x \preccurlyeq b \end{array}$$

- **dual function**

$$g(\lambda) = \inf_x \left(x^\top P\, x + \lambda^\top (A\, x - b)\right) = -\frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda - b^\top \lambda$$

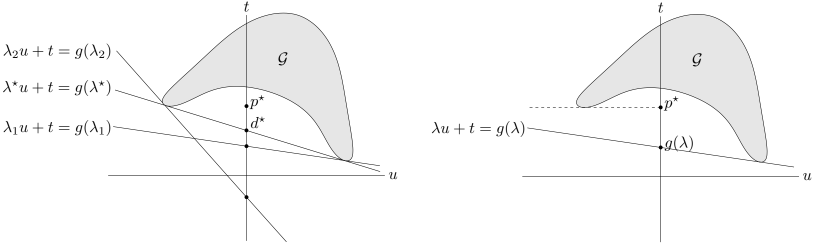
- **dual problem**

$$\begin{array}{ll} \text{maximize} & -\frac{1}{4}\lambda^\top A P^{-1} A^\top \lambda - b^\top \lambda \\ \text{subject to} & \lambda \succcurlyeq 0 \end{array}$$

- from the sharpened Slater's condition: $p^\star = d^\star$ if the primal problem is feasible
- in fact $p^\star = d^\star$ always

Geometric Interpretation

- for simplicity, consider problem with one constraint $f_1(x) \leqslant 0$
- $\mathcal{G} = \{(f_1(x), f_0(x)) \mid x \in \mathcal{D}\}$ is set of achievable (constraint, objective) values
- **interpretation of dual function**: $g(\lambda) = \inf_{(u,t)\in\mathcal{G}} (\lambda u + t)$

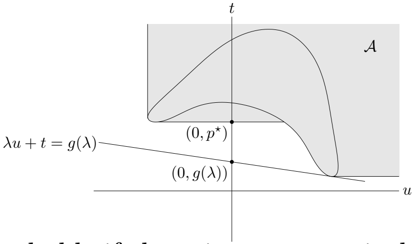


- $\lambda u + t = g(\lambda)$ is (non-vertical) supporting hyperplane to $\mathcal{G}$
- hyperplane intersects t-axis at $t = g(\lambda)$

Epigraph Variation

- same with $\mathcal{G}$ replaced with

$$\mathcal{A} = \{(u, t) \mid f_1(x) \leqslant u,\ f_0(x) \leqslant t \text{ for some } x \in \mathcal{D}\}$$



- strong duality holds if there is a non-vertical supporting hyperplane to $\mathcal{A}$ at $(0, p^\star)$
- for convex problem, $\mathcal{A}$ is convex, hence has supporting hyperplane at $(0, p^\star)$
- Slater's condition: if there exist $(\tilde{u}, \tilde{t}) \in \mathcal{A}$ with $\tilde{u} < 0$, then supporting hyperplane at $(0, p^\star)$ must be non-vertical

Complementary Slackness

- assume strong duality holds, $x^\star$ is primal optimal, $(\lambda^\star, \nu^\star)$ is dual optimal

$$\begin{aligned} f_0(x^\star) = g(\lambda^\star, \nu^\star) &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^\star f_i(x) + \sum_{i=1}^p \nu_i^\star h_i(x) \right) \\ &\leqslant f_0(x^\star) + \sum_{i=1}^m \lambda_i^\star f_i(x^\star) + \sum_{i=1}^p \nu_i^\star h_i(x^\star) \\ &\leqslant f_0(x^\star) \end{aligned}$$

- hence the two inequalities hold with equality
- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$
- $\lambda_i^\star f_i(x^\star) = 0$, $i = 1, 2, \ldots, m$: (known as **complementary slackness**)

$$\lambda_i^\star > 0 \implies f_i(x^\star) = 0, \qquad f_i(x^\star) < 0 \implies \lambda_i^\star = 0$$

Karush-Kuhn-Tucker (KKT) Conditions

KKT conditions (for a problem with differentiable f_i, h_i) are

1. primal constraints:

$$f_i(x) \leqslant 0, \quad i = 1, 2, \ldots, m$$
$$h_i(x) = 0, \quad i = 1, 2, \ldots, p$$

2. dual constraints: $\lambda \succcurlyeq 0$
3. complementary slackness:

$$\lambda_i f_i(x) = 0, \quad i = 1, 2, \ldots, m$$

4. gradient of Lagrangian with respect to x vanishes:

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0$$

if strong duality holds and x, λ, ν are optimal, they satisfy KKT conditions

KKT Conditions for Convex Problems

- if $\tilde{x}$, $\tilde{\lambda}$, $\tilde{\nu}$ satisfy KKT conditions for a convex problem, then they are optimal:

  – from complementary slackness: $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

  – from 4th condition (and convexity): $g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

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

- if Slater's condition is satisfied, then x is optimal $\iff$ $\exists\, \lambda$, ν that satisfy KKT conditions

  – recall that Slater implies strong duality, and dual optimum is attained

  – generalizes optimality condition $\nabla f_0(x) = 0$ for unconstrained problem

Perturbation and Sensitivity Analysis

unperturbed optimization and its dual

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leqslant 0, \quad i = 1, 2, \dots, m \\
& h_i(x) = 0, \quad i = 1, 2, \dots, p
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & g(\lambda, \nu) \\
\text{subject to} & \lambda \succcurlyeq 0
\end{array}
$$

perturbed optimization and its dual

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \leqslant u_i, \quad i = 1, 2, \dots, m \\
& h_i(x) = v_i, \quad i = 1, 2, \dots, p
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & g(\lambda, \nu) - u^\top \lambda - v^\top \nu \\
\text{subject to} & \lambda \succcurlyeq 0
\end{array}
$$

- x is primal variable; u, v are parameters
- $p^\star(u, v)$ is optimal value as a function of u, v
- $p^\star(0, 0)$ is optimal value of unperturbed problem

Global Sensitivity via Duality

assume strong duality holds for unperturbed problem, with $\lambda^\star$, $\nu^\star$ dual optimal

apply weak duality to perturbed problem:

$$p^\star(u, v) \geqslant g(\lambda^\star, \nu^\star) - u^\top \lambda^\star - v^\top \nu^\star = p^\star(0, 0) - u^\top \lambda^\star - v^\top \nu^\star$$

implications

- if $\lambda_i^\star$ large, $p^\star$ increases greatly if $u_i < 0$
- if $\lambda_i^\star$ small, $p^\star$ does not decrease much if $u_i > 0$
- if $\nu_i^\star$ large and positive, $p^\star$ increases greatly if $v_i < 0$
- if $\nu_i^\star$ large and negative, $p^\star$ increases greatly if $v_i > 0$
- if $\nu_i^\star$ small and positive, $p^\star$ does not decrease much if $v_i > 0$
- if $\nu_i^\star$ small and negative, $p^\star$ does not decrease much if $v_i < 0$

Local Sensitivity via Duality

if (in addition) $p^\star(u,v)$ is differentiable at $(0,0)$, then

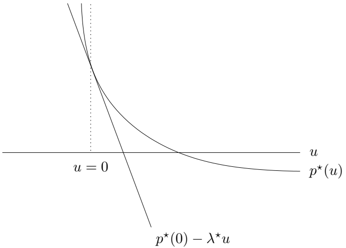
$$\lambda_i = -\frac{\partial p^\star(0,0)}{\partial u_i}, \quad \nu_i = -\frac{\partial p^\star(0,0)}{\partial v_i}$$

proof (for $\lambda_i^\star$): from global sensitivity result,

$$\frac{\partial p^\star(0,0)}{\partial u_i} = \lim_{t\to 0+} \frac{p^\star(te_i,0) - p^\star(0,0)}{t} \geqslant -\lambda_i^\star,$$

$$\frac{\partial p^\star(0,0)}{\partial u_i} = \lim_{t\to 0-} \frac{p^\star(te_i,0) - p^\star(0,0)}{t} \leqslant -\lambda_i^\star$$

$p^\star(u)$ for a problem with one (inequality) constraint:



Duality and Problem Reformulations

- equivalent formulations of a problem can lead to very different duals
- reformulating primal problem can be useful when dual is difficult to derive, or uninteresting

common reformulations

- introduce new variables and equality constraints
- make explicit constraints implicit or vice-versa
- transform objective or constraint functions, e.g. replace $f_0(x)$ by $\varphi(f_0(x))$ with φ convex, increasing

Introducing New Variables and Equality Constraints

- unconstrained problem: minimize $f_0(A\,x + b)$
- dual function is a constant: $g = \inf_x L(x) = \inf_x f_0(A\,x + b) = p^\star$
- we have strong duality, but dual is quite useless
- introduce new variable y and equality constraints $y = A\,x + b$

$$\begin{array}{rl} \text{minimize} & f_0(y) \\ \text{subject to} & A\,x + b - y = 0 \end{array}$$

- dual of reformulated problem is

$$\begin{array}{rl} \text{minimize} & b^\top \nu - f_0^\star(\nu) \\ \text{subject to} & A^\top \nu = 0 \end{array}$$

- a nontrivial, useful dual (providing the conjugate $f_0^\star$ is easy to express)

Example: Norm Approximation

- minimize $\|A\,x - b\|$
- introduce new variable y and equality constraints $y = A\,x - b$

$$\begin{array}{rl} \text{minimize} & \|y\| \\ \text{subject to} & A\,x - b - y = 0 \end{array}$$

- recall conjugate of general norm:

$$\|z\|^{\star} \equiv \begin{cases} 0 & \|z\|_{\star} \leqslant 1 \\ \infty & \text{otherwise} \end{cases}$$

- dual of reformulated norm approximation problem is

$$\begin{array}{rl} \text{minimize} & b^{\top}\nu \\ \text{subject to} & A^{\top}\nu = 0, \quad \|v\|_{\star} \leqslant 1 \end{array}$$

Theorems of Alternatives

- consider two systems of inequality and equality constraints
- called **weak alternatives** if no more than one system is feasible
- called **strong alternatives** if exactly one of them is feasible
- examples: for any $a \in \mathbb{R}$ with variable $x \in \mathbb{R}$,
  - $x > a$ and $x \leqslant a - 1$ are weak alternatives
  - $x > a$ and $x \leqslant a$ are strong alternatives
- a **theorem of alternatives** states that two inequality systems are (weak or strong) alternatives
- can be considered the extension of duality to feasibility problems

Feasibility Problems

- consider a system of (not necessarily convex) inequalities and equalities

$$
\begin{aligned}
f_i(x) &\leqslant 0, \quad i = 1, 2, \dots, m \\
h_i(x) &= 0, \quad i = 1, 2, \dots, p
\end{aligned}
$$

- express as **feasibility problem**

$$
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & f_i(x) \leqslant 0, \quad i = 1, 2, \dots, m \\
 & h_i(x) = 0, \quad i = 1, 2, \dots, p
\end{array}
$$

- if system is feasible, $p^\star = 0$; if not, $p^\star = \infty$

Duality for Feasibility Problems

- dual function of feasibility problem is

$$g(\lambda, \nu) = \inf_x \left(\sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

- for $\lambda \succcurlyeq 0$ we have $g(\lambda, \nu) \leqslant p^\star$
- it follows that feasibility of the inequality system

$$\lambda \succcurlyeq 0, \quad g(\lambda, \nu) > 0$$

  implies the original system is infeasible
- so this is a weak alternative to original system
- it is strong if f_i convex, h_i affine, and a constraint qualification holds
- g is positive homogeneous so we can write alternative system as

$$\lambda \succcurlyeq 0, \quad g(\lambda, \nu) \geqslant 1$$

Example: Nonnegative Solution of Linear Equations

- consider system

$$A\,x = b, \quad x \succcurlyeq 0$$

- dual function is

$$g(\lambda, \nu) = \begin{cases} -\nu^\top b & \text{if } A^\top \nu = \lambda \\ -\infty & \text{otherwise} \end{cases}$$

- can express strong alternative of $A\,x = b$, $x \succcurlyeq 0$ as

$$A^\top \nu \succcurlyeq 0, \quad \nu^\top b \leqslant -1$$

(we can replace $\nu^\top b \leqslant -1$ with $\nu^\top b = -1$)

Farkas Lemma

- Farkas lemma:

$$A\,x \preceq 0, \quad c^\top x < 0$$

and

$$A^\top y + c = 0, \quad y \succeq 0$$

are strong alternatives

- proof: use (strong) duality for (feasible) LP

$$\begin{array}{ll} \text{minimize} & c^\top x \\ \text{subject to} & A\,x \preceq 0 \end{array}$$