

Duality and Multiparametric Programming

Prof. Davide M. Raimondo

Dipartimento di Ingegneria Industriale e dell'Informazione

Università degli Studi di Pavia

davide.raimondo@unipv.it

Duality

standard form problem (without equality constraints)

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \end{array}$$

- optimal value p^* , domain D
- called **primal problem** (in context of duality)

(for now) we **don't** assume convexity

Duality

Lagrangian $L : \mathbf{R}^{n+m} \rightarrow \mathbf{R}$

$$L(x, \lambda) = f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)$$

- λ_i called *Lagrange multipliers* or *dual variables*
- objective is *augmented* with weighted sum of constraint functions

Lagrange dual function

(Lagrange) dual function $g : \mathbf{R}^m \rightarrow \mathbf{R} \cup \{-\infty\}$

$$\begin{aligned} g(\lambda) &= \inf_x L(x, \lambda) \\ &= \inf_x (f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)) \end{aligned}$$

- minimum of augmented cost as function of weights
- can be $-\infty$ for some λ
- g is concave (even if f_i not convex!)

Not easy to understand
(we will check it for QPs and LPs)

Lagrange dual function: LP example

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x - b_i \leq 0, \quad i = 1, \dots, m \end{array}$$

$$\begin{aligned} L(x, \lambda) &= c^T x + \sum_{i=1}^m \lambda_i (a_i^T x - b_i) \\ &= -b^T \lambda + (A^T \lambda + c)^T x \end{aligned}$$

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

Lower bound property

if $\lambda \succeq 0$ and x is primal feasible, then

$$g(\lambda) \leq f_0(x)$$

$f_0(x) - g(\lambda)$ is called the **duality gap** of (primal feasible) x and $\lambda \succeq 0$

minimize over primal feasible x to get, for any $\lambda \succeq 0$,

$$g(\lambda) \leq p^*$$

$\lambda \in \mathbf{R}^m$ is **dual feasible** if $\lambda \succeq 0$ and $g(\lambda) > -\infty$

dual feasible points yield lower bounds on optimal value!

proof: if $f_i(x) \leq 0$ and $\lambda_i \geq 0$,

$$\begin{aligned} f_0(x) &\geq f_0(x) + \sum_i \lambda_i f_i(x) \\ &\geq \inf_z \left(f_0(z) + \sum_i \lambda_i f_i(z) \right) \\ &= g(\lambda) \end{aligned}$$

Lagrangian dual problem

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

- called **(Lagrange) dual problem** (associated with primal problem)
- always a convex problem, even if primal isn't!
- optimal value denoted d^*
- we always have $d^* \leq p^*$ (called *weak duality*)
- $p^* - d^*$ is *optimal duality gap*

Strong duality

for convex problems, we (usually) have *strong duality*:

$$d^* = p^*$$

when strong duality holds, dual optimal λ^* serves as **certificate of optimality** for primal optimal point x^*

Strong duality

many conditions or *constraint qualifications* guarantee strong duality for convex problems

Slater's condition: if primal problem is strictly feasible (and convex), *i.e.*, there exists $x \in \mathbf{relint} D$ with

$$f_i(x) < 0, \quad i = 1, \dots, m$$

then we have $p^* = d^*$

Dual of linear program

(primal) LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$

- n variables, m inequality constraints

dual of LP is (after making implicit equality constraints explicit)

$$\begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0 \\ & \lambda \succeq 0 \end{array}$$

- dual of LP is also an LP (indeed, in std LP format)
- m variables, n equality constraints, m nonnegativity constraints

for LP we have strong duality except in one (pathological) case: primal and dual *both* infeasible ($p^* = +\infty, d^* = -\infty$)

Dual of quadratic program

(primal) QP

$$\begin{array}{ll} \text{minimize} & x^T P x \\ \text{subject to} & Ax \preceq b \end{array}$$

we assume $P \succ 0$ for simplicity

Lagrangian is $L(x, \lambda) = x^T P x + \lambda^T (Ax - b)$

$\nabla_x L(x, \lambda) = 0$ yields $x = -(1/2)P^{-1}A^T\lambda$, hence dual function is

$$g(\lambda) = -(1/4)\lambda^T A P^{-1} A^T \lambda - b^T \lambda$$

- concave quadratic function
- all $\lambda \succeq 0$ are dual feasible

dual of QP is

$$\begin{array}{ll} \text{maximize} & -(1/4)\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

... another QP

Duality in algorithms

many algorithms produce at iteration k

- a primal feasible $x^{(k)}$
- and a dual feasible $\lambda^{(k)}$

with $f_0(x^{(k)}) - g(\lambda^{(k)}) \rightarrow 0$ as $k \rightarrow \infty$

hence at iteration k we **know** $p^* \in [g(\lambda^{(k)}), f_0(x^{(k)})]$

- useful for stopping criteria
- algorithms that use dual solution are often more efficient (*e.g.*, LP)

Stopping criteria

$$\text{absolute error} = f_0(x^{(k)}) - p^* \leq \epsilon$$

stopping criterion:

$$\text{until } (f_0(x^{(k)}) - g(\lambda^{(k)}) \leq \epsilon)$$

$$\text{relative error} = \frac{f_0(x^{(k)}) - p^*}{|p^*|} \leq \epsilon$$

stopping criterion:

$$\text{until } (g(\lambda^{(k)}) > 0 \ \& \ \frac{f_0(x^{(k)}) - g(\lambda^{(k)})}{g(\lambda^{(k)})} \leq \epsilon)$$

$$\text{or } (f_0(x^{(k)}) < 0 \ \& \ \frac{f_0(x^{(k)}) - g(\lambda^{(k)})}{-f_0(x^{(k)})} \leq \epsilon)$$

achieve **target value** ℓ or, prove ℓ is unachievable (*i.e.*, determine either $p^* \leq \ell$ or $p^* > \ell$)

stopping criterion:

$$\text{until } (f_0(x^{(k)}) \leq \ell \ \text{or} \ g(\lambda^{(k)}) > \ell)$$

Complementary slackness condition

suppose x^* , λ^* are primal, dual feasible with zero duality gap (hence, they are primal, dual optimal)

$$\begin{aligned} f_0(x^*) &= g(\lambda^*) \\ &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) \right) \\ &\leq f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) \end{aligned}$$

hence we have $\sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$, and so

$$\lambda_i^* f_i(x^*) = 0, \quad i = 1, \dots, m$$

- i th constraint inactive at optimum $\implies \lambda_i = 0$
- $\lambda_i^* > 0$ at optimum $\implies i$ th constraint active at optimum

KKT optimality conditions

suppose

- f_i are differentiable
- x^* , λ^* are (primal, dual) optimal, with zero duality gap

by complementary slackness we have

$$f_0(x^*) + \sum_i \lambda_i^* f_i(x^*) = \inf_x \left(f_0(x) + \sum_i \lambda_i^* f_i(x) \right) \implies \nabla f_0(x^*) + \sum_i \lambda_i^* \nabla f_i(x^*) = 0$$

i.e., x^* minimizes $L(x, \lambda^*)$

KKT optimality conditions

if x^* , λ^* are (primal, dual) optimal, with zero duality gap, they satisfy

$$f_i(x^*) \leq 0$$

$$\lambda_i^* \geq 0$$

$$\lambda_i^* f_i(x^*) = 0$$

$$\nabla f_0(x^*) + \sum_i \lambda_i^* \nabla f_i(x^*) = 0$$

the **Karush-Kuhn-Tucker** (KKT) optimality conditions

conversely, if the problem is convex and x^* , λ^* satisfy KKT, then they are (primal, dual) optimal

Equality constraints

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

- optimal value p^*
- again assume (for now) not necessarily convex

define **Lagrangian** $L : \mathbf{R}^{n+m+p} \rightarrow \mathbf{R}$ as

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

dual function is $g(\lambda, \nu) = \inf_x L(x, \lambda, \nu)$

(λ, ν) is dual feasible if $\lambda \succeq 0$ and $g(\lambda, \nu) > -\infty$

(no sign condition on ν)

Dual of LP in standard form

- Primal LP

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subj. to} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$



$$\begin{aligned} \min_x \quad & c^\top x \\ \text{subj. to} \quad & b - Ax = 0 \\ & -x \leq 0 \end{aligned}$$

- Lagrangian

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

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



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

Resulting dual LP

$$\begin{aligned} \max_{\lambda, \nu} \quad & b^\top \nu \\ \text{subj. to} \quad & A^\top \nu + \lambda = c \\ & \lambda \geq 0 \end{aligned}$$

Properties

lower bound property: if x is primal feasible and (λ, ν) is dual feasible, then $g(\lambda, \nu) \leq f_0(x)$, hence

$$g(\lambda, \nu) \leq p^*$$

dual problem: find best lower bound

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

(note ν unconstrained), optimal value d^*

weak duality: $d^* \leq p^*$ always

strong duality: if primal is convex then (usually) $d^* = p^*$

Slater condition: if primal is convex (*i.e.*, f_i convex, h_i affine) and strictly feasible, *i.e.*, there exists $x \in \text{relint } D$ s.t.

$$f_i(x) < 0, \quad h_i(x) = 0,$$

then $d^* = p^*$

KKT optimality conditions with equality constraints

assume f_i, h_i differentiable

if x^*, λ^*, ν^* are optimal, with zero duality gap, then they satisfy KKT conditions

$$f_i(x^*) \leq 0, h_i(x^*) = 0$$

$$\lambda_i^* \geq 0$$

$$\lambda_i^* f_i(x^*) = 0$$

$$\nabla f_0(x^*) + \sum_i \lambda_i^* \nabla f_i(x^*) + \sum_i \nu_i^* \nabla h_i(x^*) = 0$$

conversely, if they satisfy KKT and the problem is convex, then x^*, λ^*, ν^* are optimal

KKT of LP in standard form

$$\begin{aligned}Ax &= b \\x &\geq 0 \\A^\top \nu + \lambda &= c \\ \lambda &\geq 0 \\x^\top \lambda &= 0\end{aligned}$$

$$\lambda = \nu$$

$$s = \lambda$$



change of notation

(the second one is used in some scripts)

$$\begin{aligned}Ax &= b \\x &\geq 0 \\A^\top \lambda + s &= c \\s &\geq 0 \\x^\top s &= 0\end{aligned}$$

Duality problem: example

Consider the following LP

$$\max 5x_1 + 12x_2 + 4x_3$$

$$x_1 + 2x_2 + x_3 \leq 10$$

$$2x_1 - x_2 + 3x_3 = 8$$

$$x_1, x_2, x_3 \geq 0$$

Find the dual problem (first rewrite the primal LP in standard form)

Parametric programming

- General formulation

$$f^*(x) = \inf_z f(z, x)$$
$$g(z, x) \leq 0$$

- $z \in Z \subseteq R^m$ is the vector of optimization variables
- $x \in X \subseteq R^n$ is the vector of parameters (i.e. initial state conditions)
 - When $x \in R^n$ with $n > 1$, then we talk about multi-parametric programming
- **Parametric program:** solution for the full range of parameters x
- **We consider**
 1. **Multi-parametric LPs (mp-LPs)**
 2. **Multi-parametric QPs (mp-QPs)**

Multiparametric LP (mp-LP)

Formulation: given $c \in R^m, G \in R^{q \times m}, S \in R^{q \times n}, W \in R^q$

Primal problem

$$J^*(x) = \min_z J(z, x) = c^T z,$$

subj. to $Gz \leq W + Sx$

Dual problem

$$\max_{\pi} (W + Sx)^T \pi,$$

subj. to $G^T \pi = c,$
 $\pi \leq 0$

KKT Conditions

- $Gz - Sx - W \leq 0$
- $\pi \leq 0$
- $(Gz - Sx - W)^T \pi = 0$
- $G^T \pi = c$

Multiparametric LP (mp-LP)

Feasible set: $X^* \subseteq X$ is the set of all points for which exist a solution to the primal (P.P.) and the dual problem (D.P.).

Denoting with $I = \{1, \dots, q\}$ the constraint index set, define

- **Set of Active constraints** $\mathcal{A}(x) = \{i \in I \mid \forall z : J(z, x) = J^*(x) \Rightarrow G_i z - S_i x - W_i = 0\}$
- **Set of Inactive constraints** $\mathcal{N}(x) = \{i \in I \mid \exists z : J(z, x) = J^*(x) \wedge G_i z - S_i x - W_i < 0\}$
- **Critical region** for a given set $\mathcal{A}^* \subseteq I$

$$\mathcal{CR}_{\mathcal{A}^*} = \{x \in \mathcal{X} \mid \mathcal{A}(x) = \mathcal{A}^*\}$$

mp-LP: geometric approach

- exploration of the parameter space X using geometric methods
- requires an LP solver

Step 1: solve an LP: for an initial parameter vector $x_0 \in \mathcal{X}$:

1. Solve P.P. & D.P. $\Rightarrow z^*(x_0)$ and $\pi^*(x_0)$
2. Obtain $\mathcal{A}_0 = \mathcal{A}(x_0)$ and $\mathcal{N}_0 = \mathcal{N}(x_0)$ and matrices
$$\{G_{\mathcal{A}_0}, S_{\mathcal{A}_0}, W_{\mathcal{A}_0}\} = \{G_i, S_i, W_i \mid i \in \mathcal{A}(x_0)\}$$
$$\{G_{\mathcal{N}_0}, S_{\mathcal{N}_0}, W_{\mathcal{N}_0}\} = \{G_i, S_i, W_i \mid i \in \mathcal{N}(x_0)\}$$

mp-LP: geometric approach

- exploration of the parameter space X using geometric methods
- requires an LP solver

Step 2: determine : \mathcal{CR}_{A_0} , $z_0^*(x)$ and $J_0^*(x)$

1. From the primal feasibility conditions

$$\begin{aligned} G_{A_0} z_0^*(x) &= W_{A_0} + S_{A_0} x \\ G_{N_0} z_0^*(x) &< W_{N_0} + S_{N_0} x \end{aligned}$$



2. Compute optimizer (as a function of x)

$$z_0^*(x) = G_{A_0}^{-1} S_{A_0} x + G_{A_0}^{-1} W_{A_0} = F_0 x + g_0$$

3. Critical region

$$\mathcal{CR}_{A_0} = \{x \mid (G_{N_0} F_0 - S_{N_0})x < W_{N_0} - G_{N_0} g_0\}$$

4. Compute the optimal value function using the D.P. (strong duality holds)

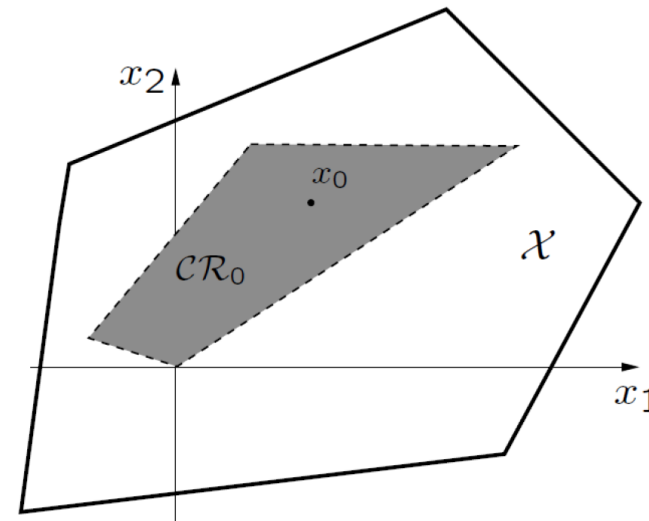
$$J_0^*(x) = (S_{A_0} x + W_{A_0})^T \pi_{A_0}^*$$

mp-LP: geometric approach

- Critical region \mathcal{CR}_{A_0} (\mathcal{CR}_0 from now on) is defined by strict inequalities \rightarrow an **open polyhedral set**
- $\mathcal{A} = \mathcal{A}(x)$ uniquely determines $\mathcal{CR}_A \rightarrow$ **the regions do not overlap!**
- the **optimizer** $z^*(x)$ is **affine** over \mathcal{CR}_0

Step 3: explore the rest of X

- Replace \mathcal{CR}_0 by its closure $\bar{\mathcal{CR}}_0$
- For $x \in X \setminus \mathcal{CR}_0$ find optimality conditions and corresponding critical regions covering the entire feasible set X^*
- *Different exploration strategies*



mp-LP: geometric approach

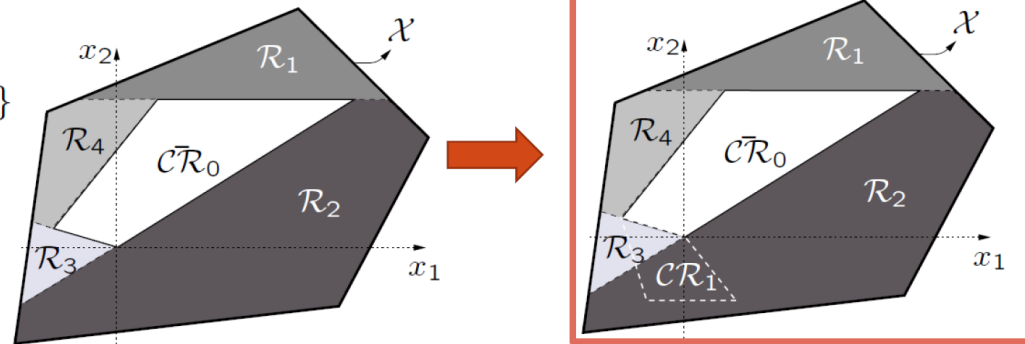
1st approach: reversing inequalities

- Note: regions R_i **are not** critical regions
- Proceed recursively: repeat the whole procedure for each R_i
- The entire set X is explored in finite number of iterations
- Problem: **critical regions can be artificially divided among different R_i**

$$\bar{C}\mathcal{R}_0 = \{x \in \mathcal{X} \mid Hx \leq K\}$$

$$\mathcal{R}_i = \{x \in \mathcal{X} \mid H_i x \geq K_i \wedge H_j x \leq K_j, \forall j < i\}$$

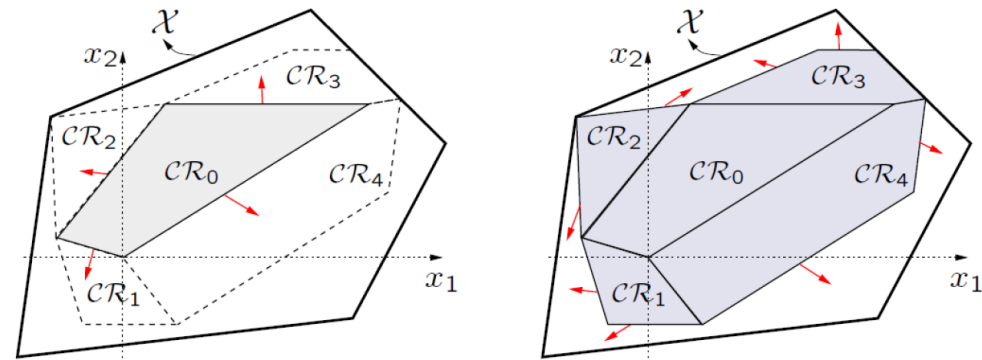
$$\mathcal{X} \setminus \bar{C}\mathcal{R}_0 = \bigcup_i \mathcal{R}_i$$



mp-LP: geometric approach

2nd approach: crossing the facets

- For each of the facets of CR_0 a point outside the region but close to the facet is selected and the procedure is repeated
- Critical regions are computed “in one piece”, no artificial splitting
- No formal proof that whole X is covered
 - In practice usually outperforms the strategy based on reversing inequalities



mp-LP: properties

1. Feasible set X^* is closed and convex
2. If the optimal solution z^* is unique for all x in X^* , the optimizer function $z(x) : X^* \rightarrow \mathbb{R}^m$ is
 - Continuous
 - Polyhedral piecewise affine (PPWA) over X^*
 - Affine in each CR_i

If the solution is not unique, it is always possible to choose a continuous and PPWA optimizer function $z^*(x)$

3. The value function $J^*(x) : X^* \rightarrow \mathbb{R}$ is
 - Convex
 - PPWA over X^* , affine in each CR_i

mp-QP

Formulation: multi-parametric quadratic programs of the form (*Assume $H > 0$*)

$$J^*(x) = \min_z \left\{ J(z, x) = \frac{1}{2} z^T H z \right\}$$

subj. to $Gz \leq W + Sx$

where $z \in \mathcal{Z} \subseteq \mathbb{R}^m$, and $x \in \mathcal{X} \subseteq \mathbb{R}^n$, $G \in \mathbb{R}^{q \times m}$.

KKT Conditions

$$\begin{aligned} Hz^* + G^T \lambda^* &= 0, \\ \lambda_i^* (G_i z^* - S_i x - W_i) &= 0, \quad i = 1, \dots, q \\ \lambda^* &\geq 0, \\ Gz^* &\leq W + Sx \end{aligned}$$

mp-QP: geometric algorithm

- exploration of the parameter space X using geometric methods
- requires an QP solver

Step 1: solve an QP: for an initial parameter vector $x_0 \in \mathcal{X}$:

1. Solve P.P. & D.P. $\Rightarrow z^*(x_0)$ and $\lambda^*(x_0)$
2. Obtain $\mathcal{A}_0 = \mathcal{A}(x_0)$ and $\mathcal{N}_0 = \mathcal{N}(x_0)$ and matrices
$$\{G_{\mathcal{A}_0}, S_{\mathcal{A}_0}, W_{\mathcal{A}_0}\} = \{G_i, S_i, W_i \mid i \in \mathcal{A}(x_0)\}$$
$$\{G_{\mathcal{N}_0}, S_{\mathcal{N}_0}, W_{\mathcal{N}_0}\} = \{G_i, S_i, W_i \mid i \in \mathcal{N}(x_0)\}$$

mp-QP: geometric algorithm

- exploration of the parameter space X using geometric methods
- requires an QP solver

Step 2: determine : $\mathcal{CR}_{\mathcal{A}_0}$, $z_0^*(x)$ and $J_0^*(x)$

1. From the primal feasibility conditions

$$\begin{aligned} G_{\mathcal{A}_0} z_0^*(x) &= W_{\mathcal{A}_0} + S_{\mathcal{A}_0} x \\ G_{\mathcal{N}_0} z_0^*(x) &< W_{\mathcal{N}_0} + S_{\mathcal{N}_0} x \end{aligned}$$



2. Since $z^* = -H^{-1}G^T\lambda^*$ and $\lambda_{\mathcal{A}}^*(x) = -(G_{\mathcal{A}}H^{-1}G_{\mathcal{A}}^T)^{-1}(W_{\mathcal{A}} + S_{\mathcal{A}}x)$
the optimizer function is $z^*(x) = H^{-1}G_{\mathcal{A}}^T(G_{\mathcal{A}}H^{-1}G_{\mathcal{A}}^T)^{-1}(W_{\mathcal{A}} + S_{\mathcal{A}}x)$

3. Critical region

$$\mathcal{CR}_{\mathcal{A}} = \{x \mid Ax \leq b\}$$

$$\begin{aligned} A &= \begin{bmatrix} GH^{-1}G_{\mathcal{A}}^T(G_{\mathcal{A}}H^{-1}G_{\mathcal{A}}^T)^{-1}S_{\mathcal{A}} - S \\ (G_{\mathcal{A}}H^{-1}G_{\mathcal{A}}^T)^{-1}S_{\mathcal{A}} \end{bmatrix} \\ b &= \begin{bmatrix} W + GH^{-1}G_{\mathcal{A}}^T(G_{\mathcal{A}}H^{-1}G_{\mathcal{A}}^T)^{-1}W_{\mathcal{A}} \\ (G_{\mathcal{A}}H^{-1}G_{\mathcal{A}}^T)^{-1}W_{\mathcal{A}} \end{bmatrix} \end{aligned}$$

mp-QP: geometric algorithm

- Note that the function $z^*(\mathbf{x})$ is a **uniquely defined affine function** over the critical region CR_A
- Moreover, the critical region is a **polyhedral set** in the \mathbf{x} -space

Step 3: explore the rest of X (*as before..*)

mp-QP: properties

1. Feasible set X^* is closed and convex
2. The optimizer function $z(\mathbf{x}) : X^* \rightarrow \mathbb{R}^m$ is
 - Continuous
 - Polyhedral piecewise affine (PPWA) over X^*
 - Affine in each CR_i
3. The value function $J^*(\mathbf{x}) : X^* \rightarrow \mathbb{R}$ is
 - Continuous
 - Convex
 - polyhedral piecewise quadratic (PPWQ) (piecewise quadratic over the critical regions CR_i)