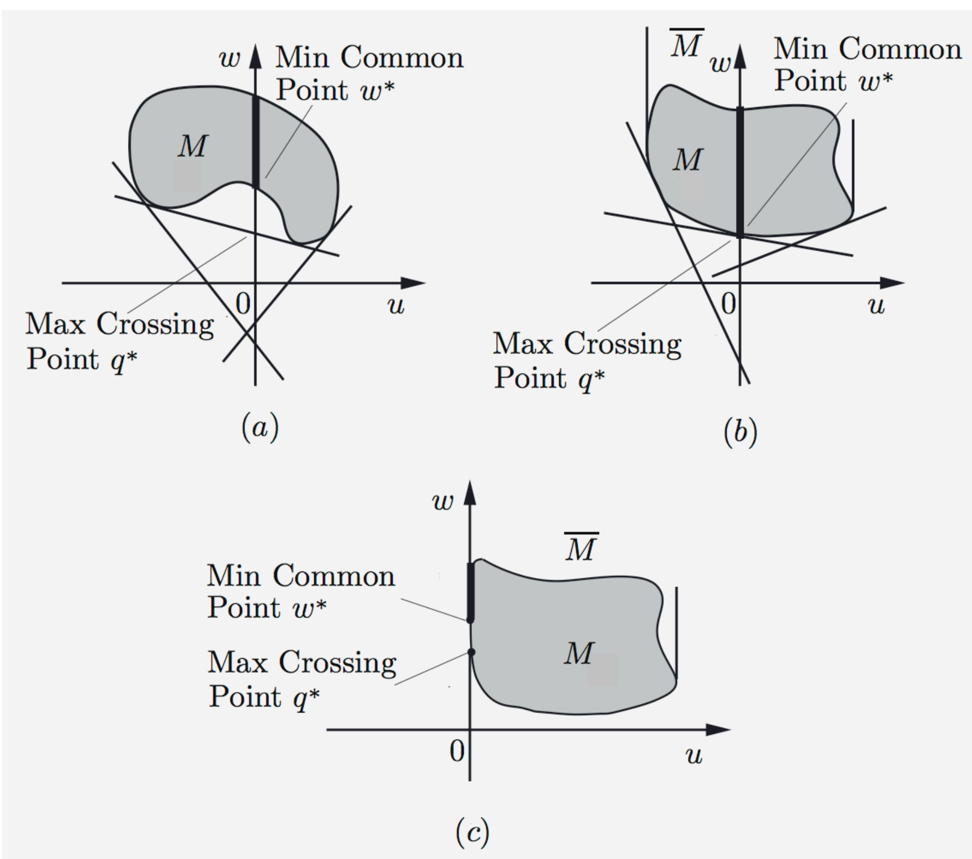
LECTURE 9

LECTURE OUTLINE

- Minimax problems and zero-sum games
- Min Common / Max Crossing duality for minimax and zero-sum games
- Min Common / Max Crossing duality theorems
- Strong duality conditions
- Existence of dual optimal solutions

Reading: Sections 3.4, 4.3, 4.4, 5.1



REVIEW OF THE MC/MC FRAMEWORK

- Given set $M \subset \Re^{n+1}$,

$$w^* = \inf_{(0,w)\in M} w, \quad q^* = \sup_{\mu\in\Re^n} q(\mu) \stackrel{\triangle}{=} \inf_{(u,w)\in M} \{w+\mu' u\}$$

- **Weak Duality:** $q^* \leq w^*$

- **Important special case:** $M = \text{epi}(p)$. Then $w^* = p(0)$, $q^* = p^{\star\star}(0)$, so we have $w^* = q^*$ if p is closed, proper, convex.

- Some applications:
  - Constrained optimization: $\min_{x\in X,\, g(x)\leq 0} f(x)$, with $p(u) = \inf_{x\in X,\, g(x)\leq u} f(x)$
  - Other optimization problems: Fenchel and conic optimization
  - Useful theorems related to optimization: Farkas' lemma, theorems of the alternative
  - Subgradient theory
  - Minimax problems, 0-sum games

- **Strong Duality:** $q^* = w^*$. Requires that M have some convexity structure, among other conditions

MINIMAX PROBLEMS

Given $\phi : X \times Z \mapsto \Re$, where $X \subset \Re^n$, $Z \subset \Re^m$ consider

$$\begin{aligned} &\text{minimize} && \sup_{z \in Z} \phi(x, z) \\ &\text{subject to} && x \in X \end{aligned}$$

or

$$\begin{aligned} &\text{maximize} && \inf_{x \in X} \phi(x, z) \\ &\text{subject to} && z \in Z. \end{aligned}$$

- Some important contexts:
  - Constrained optimization duality theory
  - Zero sum game theory
- We always have

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) \leq \inf_{x \in X} \sup_{z \in Z} \phi(x, z)$$

- **Key question:** When does equality hold?

CONSTRAINED OPTIMIZATION DUALITY

- For the problem

$$\begin{aligned} &\text{minimize} \;\; f(x) \\ &\text{subject to} \;\; x \in X, \qquad g(x) \leq 0 \end{aligned}$$

introduce the Lagrangian function

$$L(x, \mu) = f(x) + \mu' g(x)$$

- Primal problem (equivalent to the original)

$$\min_{x \in X} \sup_{\mu \geq 0} L(x, \mu) = \begin{cases} f(x) & \text{if } g(x) \leq 0, \\ \infty & \text{otherwise,} \end{cases}$$

- Dual problem

$$\max_{\mu \geq 0} \inf_{x \in X} L(x, \mu)$$

- Key duality question: Is it true that

$$\inf_{x \in \Re^n} \sup_{\mu \geq 0} L(x, \mu) = w^* \overset{?}{=} q^* = \sup_{\mu \geq 0} \inf_{x \in \Re^n} L(x, \mu)$$

ZERO SUM GAMES

- Two players: 1st chooses $i \in \{1, \ldots, n\}$, 2nd chooses $j \in \{1, \ldots, m\}$.

- If i and j are selected, the 1st player gives a_{ij} to the 2nd.

- Mixed strategies are allowed: The two players select probability distributions

$$x = (x_1, \ldots, x_n), \qquad z = (z_1, \ldots, z_m)$$

over their possible choices.

- Probability of (i, j) is $x_i z_j$, so the expected amount to be paid by the 1st player

$$x' A z = \sum_{i,j} a_{ij} x_i z_j$$

where A is the $n \times m$ matrix with elements a_{ij}.

- Each player optimizes his choice against the worst possible selection by the other player. So
  - 1st player minimizes $\max_z x' A z$
  - 2nd player maximizes $\min_x x' A z$

SADDLE POINTS

Definition: (x^*, z^*) is called a *saddle point* of ϕ if

$$\phi(x^*, z) \leq \phi(x^*, z^*) \leq \phi(x, z^*), \quad \forall\, x \in X,\, \forall\, z \in Z$$

Proposition: (x^*, z^*) is a saddle point if and only if the minimax equality holds and

$$x^* \in \arg\min_{x \in X} \sup_{z \in Z} \phi(x, z), \quad z^* \in \arg\max_{z \in Z} \inf_{x \in X} \phi(x, z) \quad (*)$$

Proof: If (x^*, z^*) is a saddle point, then

$$\begin{aligned} \inf_{x \in X} \sup_{z \in Z} \phi(x, z) &\leq \sup_{z \in Z} \phi(x^*, z) = \phi(x^*, z^*) \\ &= \inf_{x \in X} \phi(x, z^*) \leq \sup_{z \in Z} \inf_{x \in X} \phi(x, z) \end{aligned}$$

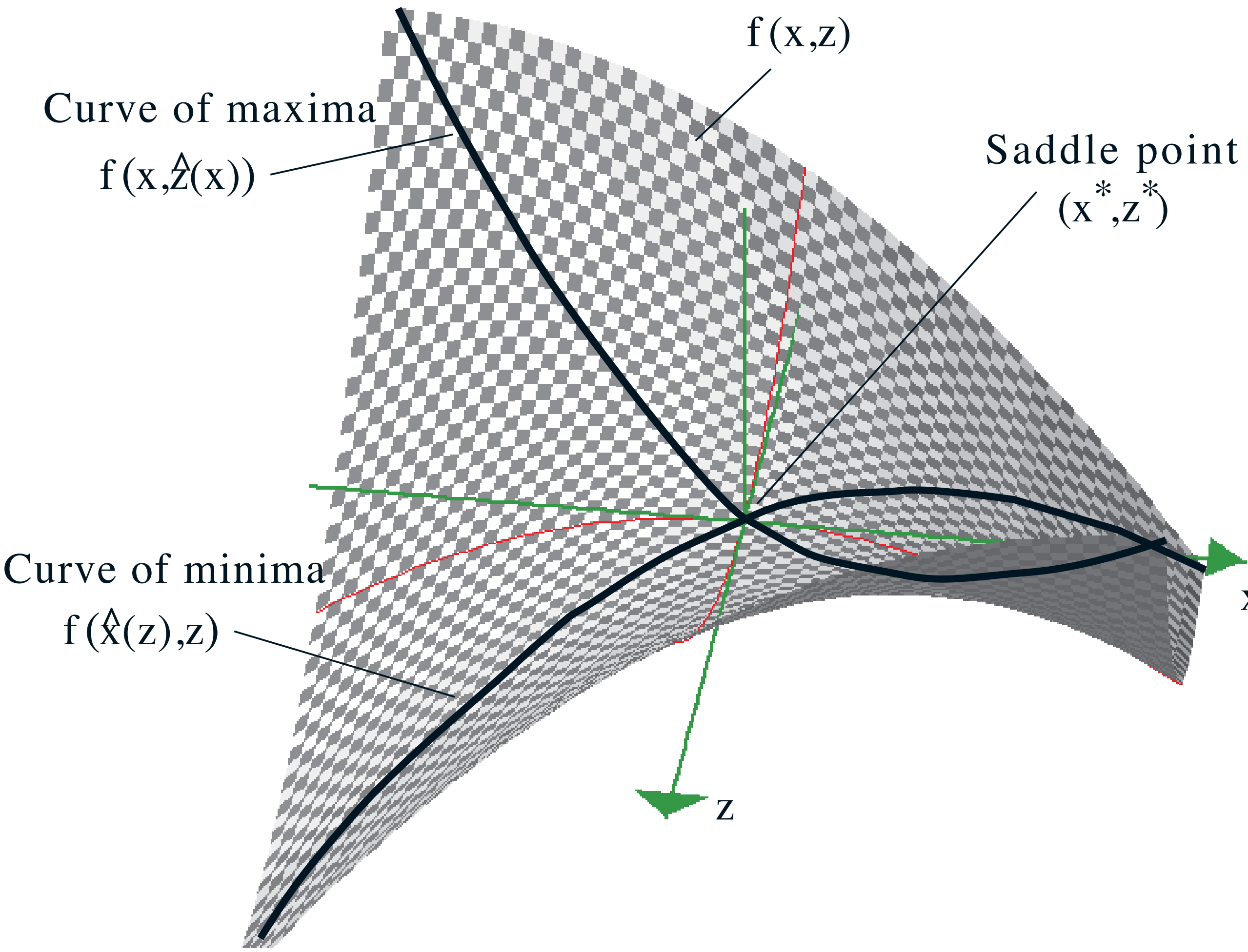
By the minimax inequality, the above holds as an equality throughout, so the minimax equality and Eq. (*) hold.

Conversely, if Eq. (*) holds, then

$$\begin{aligned} \sup_{z \in Z} \inf_{x \in X} \phi(x, z) &= \inf_{x \in X} \phi(x, z^*) \leq \phi(x^*, z^*) \\ &\leq \sup_{z \in Z} \phi(x^*, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z) \end{aligned}$$

Using the minimax equ., (x^*, z^*) is a saddle point.

VISUALIZATION



The curve of maxima $f(x, \hat{z}(x))$ lies above the curve of minima $f(\hat{x}(z), z)$, where

$$\hat{z}(x) = \arg\max_{z} f(x, z), \qquad \hat{x}(z) = \arg\min_{x} f(x, z)$$

Saddle points correspond to points where these two curves meet.

MINIMAX MC/MC FRAMEWORK

- Introduce perturbation function $p : \Re^m \mapsto [-\infty, \infty]$

$$p(u) = \inf_{x \in X} \sup_{z \in Z} \big\{ \phi(x,z) - u'z \big\}, \qquad u \in \Re^m$$

- Apply the MC/MC framework with $M = \text{epi}(p)$. If p is convex, closed, and proper, no duality gap.
- Introduce $\hat{\text{cl}}\, \phi$, the *concave closure of* ϕ viewed as a function of z for fixed x
- We have

$$\sup_{z \in Z} \phi(x,z) = \sup_{z \in \Re^m} (\hat{\text{cl}}\, \phi)(x,z),$$

so

$$w^* = p(0) = \inf_{x \in X} \sup_{z \in \Re^m} (\hat{\text{cl}}\, \phi)(x,z).$$

- The dual function can be shown to be

$$q(\mu) = \inf_{x \in X} (\hat{\text{cl}}\, \phi)(x,\mu), \qquad \forall\ \mu \in \Re^m$$

so if $\phi(x,\cdot)$ is concave and closed,

$$w^* = \inf_{x \in X} \sup_{z \in \Re^m} \phi(x,z), \qquad q^* = \sup_{z \in \Re^m} \inf_{x \in X} \phi(x,z)$$

PROOF OF FORM OF DUAL FUNCTION

- Write $p(u) = \inf_{x\in X} p_x(u)$, where

$$p_x(u) = \sup_{z\in Z}\big\{\phi(x,z) - u'z\big\}, \qquad x \in X,$$

and note that

$$\inf_{u\in\Re^m}\big\{p_x(u)+u'\mu\big\} = -\sup_{u\in\Re^m}\big\{u'(-\mu)-p_x(u)\big\} = -p_x^\star(-\mu)$$

Except for a sign change, p_x is the conjugate of $(-\phi)(x,\cdot)$ [assuming $(-\hat{\text{cl}}\,\phi)(x,\cdot)$ is proper], so

$$p_x^\star(-\mu) = -(\hat{\text{cl}}\,\phi)(x,\mu).$$

Hence, for all $\mu \in \Re^m$,

$$\begin{aligned} q(\mu) &= \inf_{u\in\Re^m}\big\{p(u) + u'\mu\big\} \\ &= \inf_{u\in\Re^m}\inf_{x\in X}\big\{p_x(u) + u'\mu\big\} \\ &= \inf_{x\in X}\inf_{u\in\Re^m}\big\{p_x(u) + u'\mu\big\} \\ &= \inf_{x\in X}\big\{-p_x^\star(-\mu)\big\} \\ &= \inf_{x\in X}(\hat{\text{cl}}\,\phi)(x,\mu) \end{aligned}$$

DUALITY THEOREMS

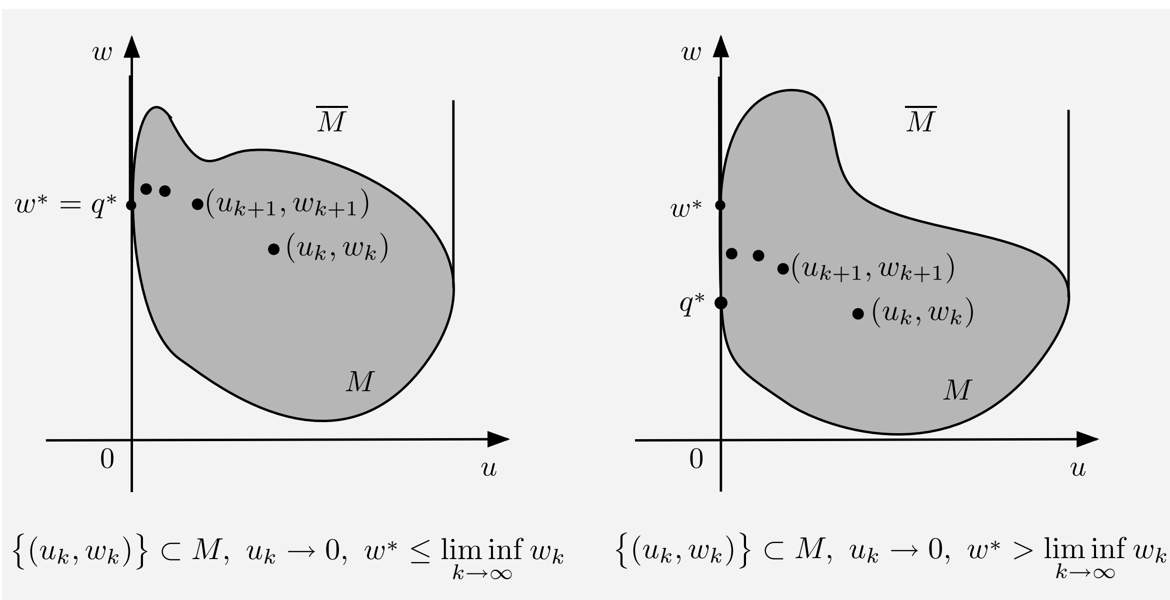
- Assume that $w^* < \infty$ and that the set

$$\overline{M} = \Big\{(u, w) \mid \text{there exists } \overline{w} \text{ with } \overline{w} \leq w \text{ and } (u, \overline{w}) \in M\Big\}$$

is convex.

- **Min Common/Max Crossing Theorem I:** We have $q^* = w^*$ if and only if for every sequence $\big\{(u_k, w_k)\big\} \subset M$ with $u_k \to 0$, there holds

$$w^* \leq \liminf_{k\to\infty} w_k.$$



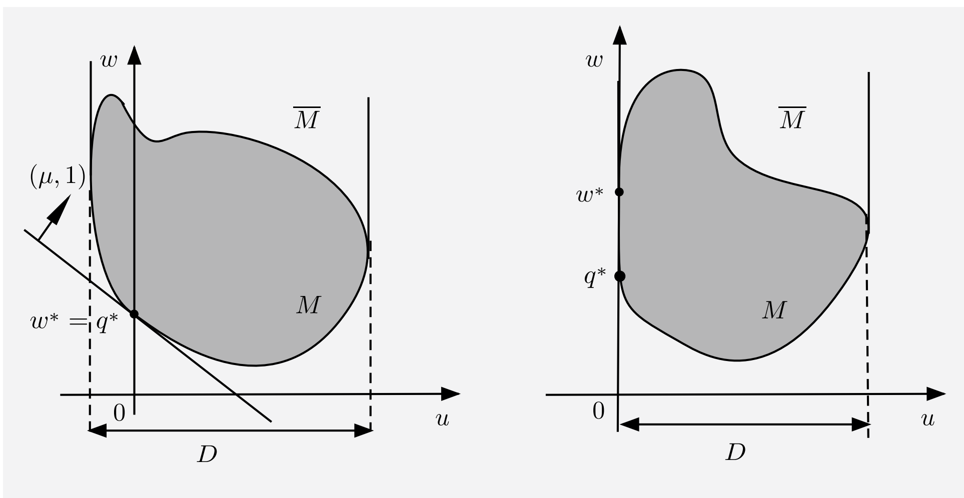
- **Corollary:** If $M = \text{epi}(p)$ where p is closed proper convex and $p(0) < \infty$, then $q^* = w^*$.

DUALITY THEOREMS (CONTINUED)

- **Min Common/Max Crossing Theorem II:** Assume in addition that $-\infty < w^*$ and that

$$D = \{u \mid \text{there exists } w \in \Re \text{ with } (u, w) \in \overline{M}\}$$

contains the origin in its relative interior. Then $q^* = w^*$ and there exists μ such that $q(\mu) = q^*$.



- Furthermore, the set $\{\mu \mid q(\mu) = q^*\}$ is nonempty and compact if and only if D contains the origin in its interior.

- **Min Common/Max Crossing Theorem III:** Involves polyhedral assumptions, and will be developed later.

PROOF OF THEOREM I

- Assume that $q^* = w^*$. Let $\{(u_k, w_k)\} \subset M$ be such that $u_k \to 0$. Then,

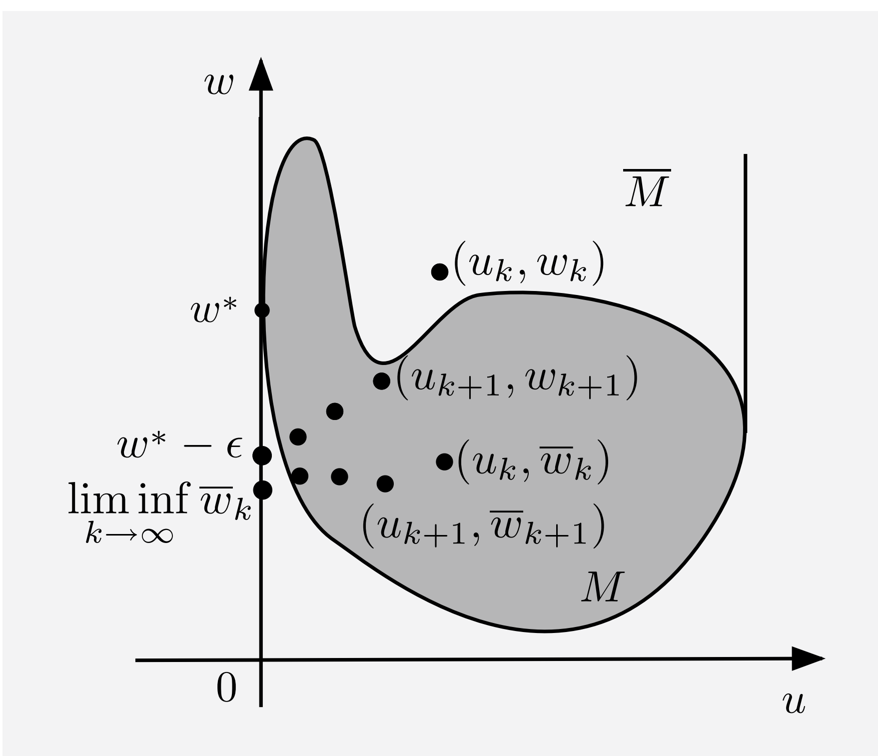
$$q(\mu) = \inf_{(u,w)\in M} \{w + \mu' u\} \le w_k + \mu' u_k, \quad \forall\, k,\ \forall\, \mu \in \Re^n$$

Taking the limit as $k \to \infty$, we obtain $q(\mu) \le \liminf_{k\to\infty} w_k$, for all $\mu \in \Re^n$, implying that

$$w^* = q^* = \sup_{\mu\in\Re^n} q(\mu) \le \liminf_{k\to\infty} w_k$$

Conversely, assume that for every sequence $\{(u_k, w_k)\} \subset M$ with $u_k \to 0$, there holds $w^* \le \liminf_{k\to\infty} w_k$. If $w^* = -\infty$, then $q^* = -\infty$, by weak duality, so assume that $-\infty < w^*$. Steps:

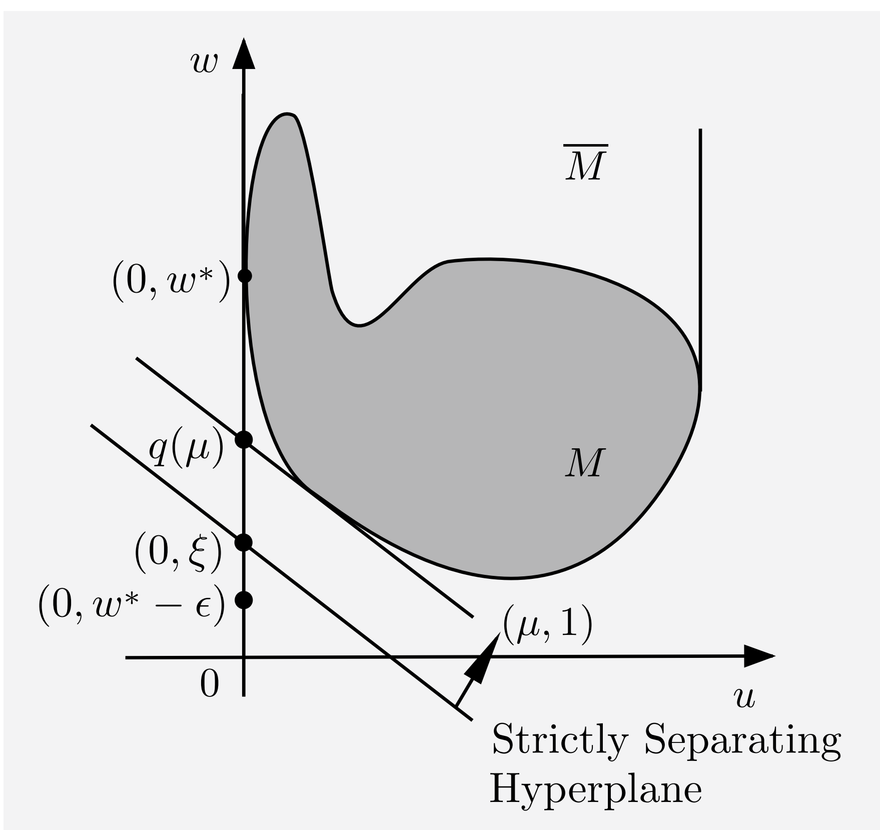
- **Step 1:** $(0, w^* - \epsilon) \notin \mathrm{cl}(\overline{M})$ for any $\epsilon > 0$.



PROOF OF THEOREM I (CONTINUED)

- **Step 2:** $\overline{M}$ does not contain any vertical lines. If this were not so, $(0, -1)$ would be a direction of recession of $\mathrm{cl}(\overline{M})$. Because $(0, w^*) \in \mathrm{cl}(\overline{M})$, the entire halfline $\{(0, w^* - \epsilon) \mid \epsilon \geq 0\}$ belongs to $\mathrm{cl}(\overline{M})$, contradicting Step 1.

- **Step 3:** For any $\epsilon > 0$, since $(0, w^* - \epsilon) \notin \mathrm{cl}(\overline{M})$, there exists a nonvertical hyperplane strictly separating $(0, w^* - \epsilon)$ and $\overline{M}$. This hyperplane crosses the $(n+1)$st axis at a vector $(0, \xi)$ with $w^* - \epsilon \leq \xi \leq w^*$, so $w^* - \epsilon \leq q^* \leq w^*$. Since ϵ can be arbitrarily small, it follows that $q^* = w^*$.



PROOF OF THEOREM II

• Note that $(0, w^*)$ is not a relative interior point of $\overline{M}$. Therefore, by the Proper Separation Theorem, there is a hyperplane that passes through $(0, w^*)$, contains $\overline{M}$ in one of its closed halfspaces, but does not fully contain $\overline{M}$, i.e., for some $(\mu, \beta) \neq (0, 0)$

$$\beta w^* \leq \mu' u + \beta w, \qquad \forall\, (u, w) \in \overline{M},$$

$$\beta w^* < \sup_{(u,w) \in \overline{M}} \{\mu' u + \beta w\}$$

Will show that the hyperplane is nonvertical.

• Since for any $(\overline{u}, \overline{w}) \in M$, the set $\overline{M}$ contains the halfline $\big\{(\overline{u}, w) \mid \overline{w} \leq w\big\}$, it follows that $\beta \geq 0$. If $\beta = 0$, then $0 \leq \mu' u$ for all $u \in D$. Since $0 \in \mathrm{ri}(D)$ by assumption, we must have $\mu' u = 0$ for all $u \in D$ a contradiction. Therefore, $\beta > 0$, and we can assume that $\beta = 1$. It follows that

$$w^* \leq \inf_{(u,w) \in \overline{M}} \{\mu' u + w\} = q(\mu) \leq q^*$$

Since the inequality $q^* \leq w^*$ holds always, we must have $q(\mu) = q^* = w^*$.

NONLINEAR FARKAS' LEMMA

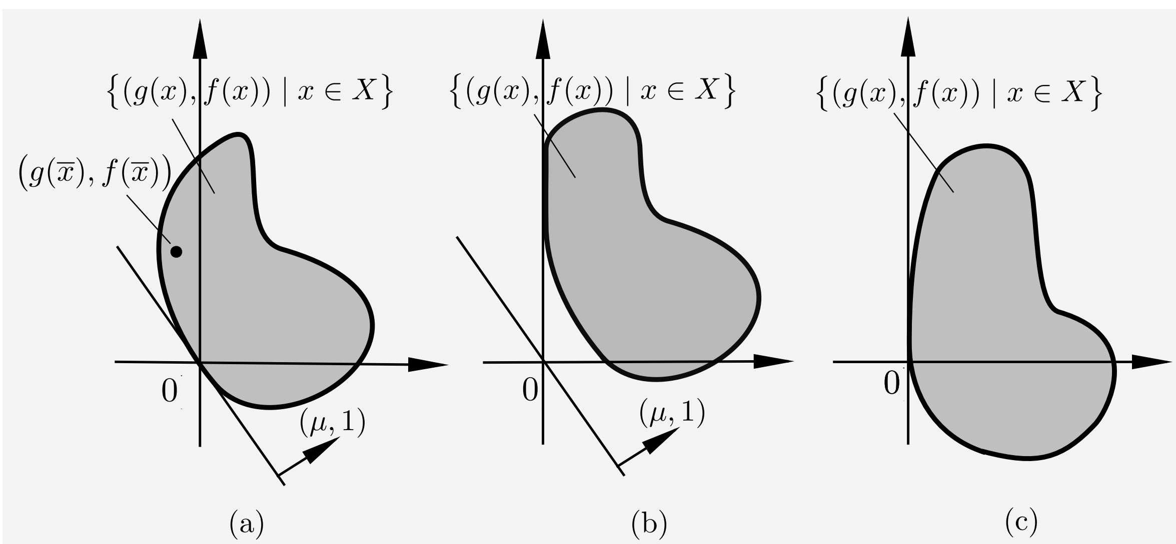
- Let $X \subset \Re^n$, $f : X \mapsto \Re$, and $g_j : X \mapsto \Re$, $j = 1, \ldots, r$, be convex. Assume that

$$f(x) \geq 0, \qquad \forall\, x \in X \text{ with } g(x) \leq 0$$

Let

$$Q^* = \big\{\mu \mid \mu \geq 0,\ f(x) + \mu' g(x) \geq 0,\ \forall\, x \in X\big\}.$$

Then Q^* is nonempty and compact if and only if there exists a vector $\overline{x} \in X$ such that $g_j(\overline{x}) < 0$ for all $j = 1, \ldots, r$.



- The lemma asserts the existence of a nonvertical hyperplane in $\Re^{r+1}$, with normal $(\mu, 1)$, that passes through the origin and contains the set

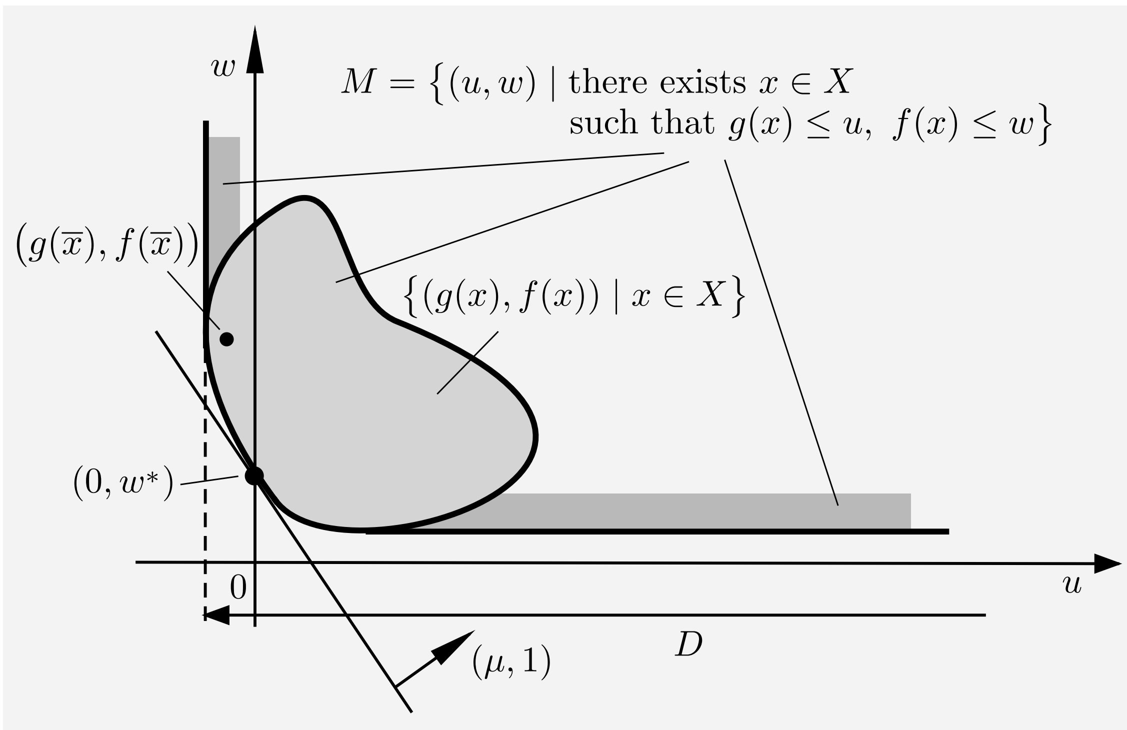
$$\big\{\big(g(x), f(x)\big) \mid x \in X\big\}$$

in its positive halfspace.

PROOF OF NONLINEAR FARKAS' LEMMA

- Apply MC/MC to

$$M = \big\{(u, w) \mid \text{there is } x \in X \text{ s. t. } g(x) \le u,\ f(x) \le w\big\}$$



- M is equal to $\overline{M}$ and is formed as the union of positive orthants translated to points $\big(g(x), f(x)\big)$, $x \in X$.

- The convexity of X, f, and g_j implies convexity of M.

- MC/MC Theorem II applies: we have

$$D = \big\{u \mid \text{there exists } w \in \Re \text{ with } (u, w) \in \overline{M}\big\}$$

and $0 \in \text{int}(D)$, because $\big(g(\overline{x}), f(\overline{x})\big) \in M$.