LECTURE 13

LECTURE OUTLINE

- Subgradients
- Fenchel inequality
- Sensitivity in constrained optimization
- Subdifferential calculus
- Optimality conditions

SUBGRADIENTS

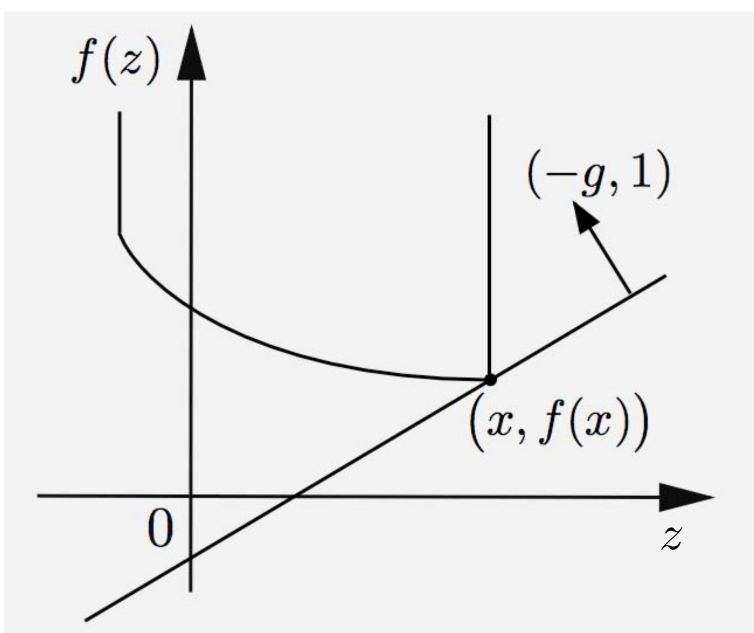
- Let $f : \Re^n \mapsto (-\infty, \infty]$ be a convex function. A vector $g \in \Re^n$ is a *subgradient* of f at a point $x \in \text{dom}(f)$ if

$$f(z) \geq f(x) + (z - x)'g, \qquad \forall\, z \in \Re^n$$

- g is a subgradient if and only if

$$f(z) - z'g \geq f(x) - x'g, \qquad \forall\, z \in \Re^n$$

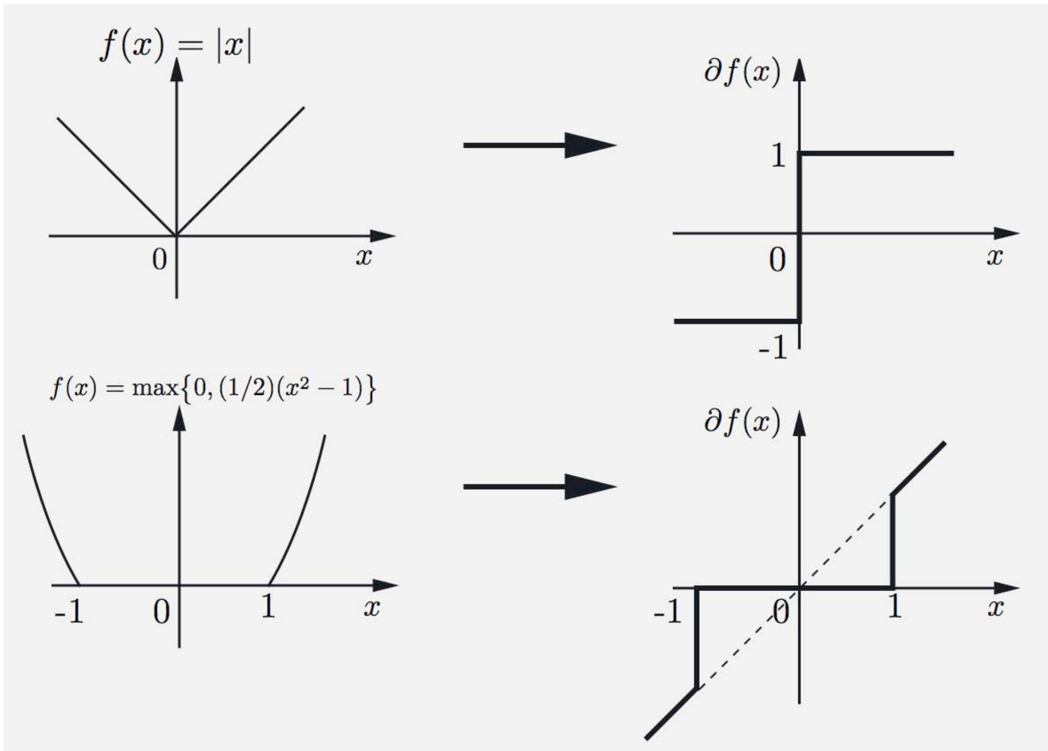
so g is a subgradient at x if and only if the hyperplane in $\Re^{n+1}$ that has normal $(-g, 1)$ and passes through $\big(x, f(x)\big)$ supports the epigraph of f.



- The set of all subgradients at x is the *subdifferential of f at x*, denoted $\partial f(x)$.

EXAMPLES OF SUBDIFFERENTIALS

- Some examples:



- If f is differentiable, then $\partial f(x) = \{\nabla f(x)\}$.

Proof: If $g \in \partial f(x)$, then

$$f(x+z) \geq f(x) + g'z, \qquad \forall\ z \in \Re^n.$$

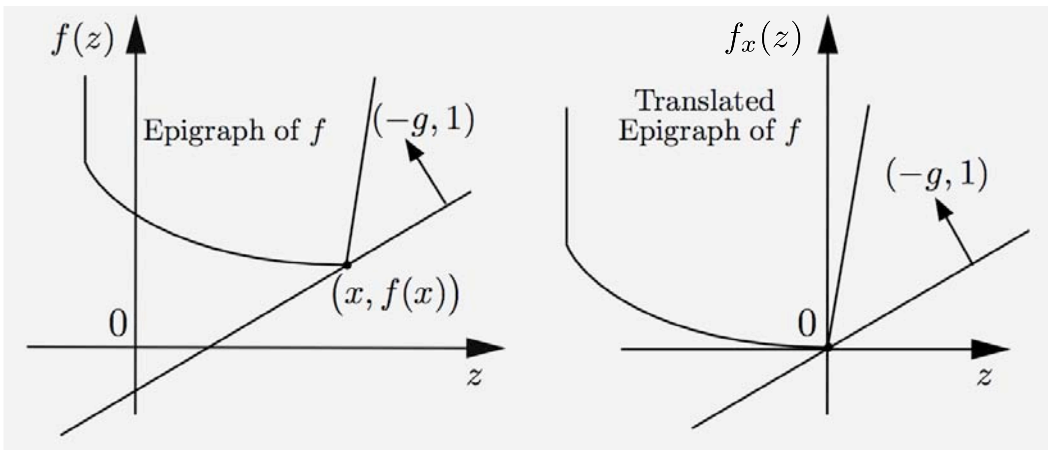
Apply this with $z = \gamma\big(\nabla f(x) - g\big)$, $\gamma \in \Re$, and use 1st order Taylor series expansion to obtain

$$\gamma \|\nabla f(x) - g\|^2 \geq o(\gamma), \qquad \forall\ \gamma \in \Re$$

EXISTENCE OF SUBGRADIENTS

- Note the connection with MC/MC

$$M = \mathrm{epi}(f_x), \qquad f_x(z) = f(x+z) - f(x)$$



- Let $f : \Re^n \mapsto (-\infty, \infty]$ be a proper convex function. For every $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$,

$$\partial f(x) = S^{\perp} + G,$$

where:

  – S is the subspace that is parallel to the affine hull of $\mathrm{dom}(f)$

  – G is a nonempty and compact set.

- Furthermore, $\partial f(x)$ is nonempty and compact if and only if x is in the interior of $\mathrm{dom}(f)$.

EXAMPLE: SUBDIFFERENTIAL OF INDICATOR

- Let C be a convex set, and δ_C be its indicator function.
- For $x \notin C$, $\partial \delta_C(x) = \emptyset$, by convention.
- For $x \in C$, we have $g \in \partial \delta_C(x)$ iff

$$\delta_C(z) \geq \delta_C(x) + g'(z - x), \qquad \forall\, z \in C,$$

or equivalently $g'(z - x) \leq 0$ for all $z \in C$. Thus $\partial \delta_C(x)$ is the *normal cone of C at x*, denoted $N_C(x)$:

$$N_C(x) = \big\{g \mid g'(z - x) \leq 0,\ \forall\, z \in C\big\}.$$

- **Example:** For the case of a polyhedral set

$$P = \{x \mid a_i'x \leq b_i,\ i = 1, \ldots, m\},$$

we have

$$N_P(x) = \begin{cases} \{0\} & \text{if } x \in \text{int}(P), \\ \text{cone}\big(\{a_i \mid a_i'x = b_i\}\big) & \text{if } x \notin \text{int}(P). \end{cases}$$

FENCHEL INEQUALITY

- Let $f : \Re^n \mapsto (-\infty, \infty]$ be proper convex and let $f^\star$ be its conjugate. Using the definition of conjugacy, we have *Fenchel's inequality*:

$$x'y \le f(x) + f^\star(y), \qquad \forall\, x \in \Re^n,\ y \in \Re^n.$$

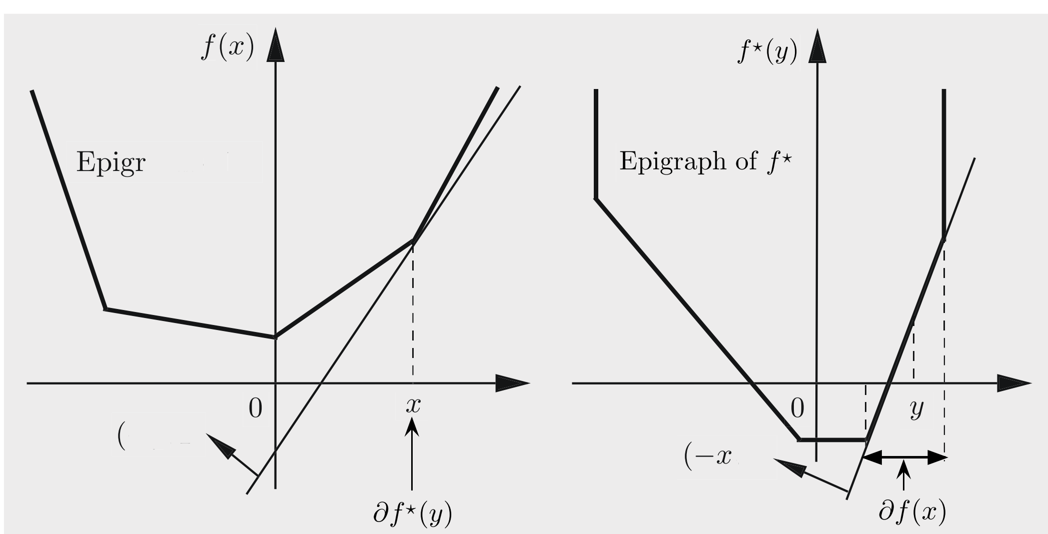
- **Conjugate Subgradient Theorem:** The following two relations are equivalent for a pair of vectors (x, y):

  (i) $x'y = f(x) + f^\star(y)$.

  (ii) $y \in \partial f(x)$.

If f is closed, (i) and (ii) are equivalent to

(iii) $x \in \partial f^\star(y)$.



MINIMA OF CONVEX FUNCTIONS

- **Application:** Let f be closed proper convex and let X^* be the set of minima of f over $\Re^n$. Then:

  (a) $X^* = \partial f^\star(0)$.

  (b) X^* is nonempty if $0 \in \mathrm{ri}\big(\mathrm{dom}(f^\star)\big)$.

  (c) X^* is nonempty and compact if and only if $0 \in \mathrm{int}\big(\mathrm{dom}(f^\star)\big)$.

Proof: (a) From the subgradient inequality,

$$x^* \text{ minimizes } f \quad \text{iff} \quad 0 \in \partial f(x^*),$$

and since

$$0 \in \partial f(x^*) \quad \text{iff} \quad x^* \in \partial f^\star(0),$$

we have

$$x^* \text{ minimizes } f \quad \text{iff} \quad x^* \in \partial f^\star(0),$$

(b) $\partial f^\star(0)$ is nonempty if $0 \in \mathrm{ri}\big(\mathrm{dom}(f^\star)\big)$.

(c) $\partial f^\star(0)$ is nonempty and compact if and only if $0 \in \mathrm{int}\big(\mathrm{dom}(f^\star)\big)$. **Q.E.D.**

SENSITIVITY INTERPRETATION

- Consider MC/MC for the case $M = \text{epi}(p)$.
- Dual function is

$$q(\mu) = \inf_{u \in \Re^m} \big\{ p(u) + \mu' u \big\} = -p^\star(-\mu),$$

where $p^\star$ is the conjugate of p.

- Assume p is proper convex and strong duality holds, so $p(0) = w^* = q^* = \sup_{\mu \in \Re^m} \big\{ -p^\star(-\mu) \big\}$. Let Q^* be the set of dual optimal solutions,

$$Q^* = \big\{ \mu^* \mid p(0) + p^\star(-\mu^*) = 0 \big\}.$$

From Conjugate Subgradient Theorem, $\mu^* \in Q^*$ if and only if $-\mu^* \in \partial p(0)$, i.e., $Q^* = -\partial p(0)$.

- If p is convex and differentiable at 0, $-\nabla p(0)$ is equal to the unique dual optimal solution μ^*.
- Constrained optimization example

$$p(u) = \inf_{x \in X,\, g(x) \le u} f(x),$$

If p is convex and differentiable,

$$\mu_j^* = -\frac{\partial p(0)}{\partial u_j}, \qquad j = 1, \ldots, r.$$

EXAMPLE: SUBDIFF. OF SUPPORT FUNCTION

- Consider the support function $\sigma_X(y)$ of a set X. To calculate $\partial\sigma_X(y)$ at some y, we introduce

$$r(y) = \sigma_X(y + y), \qquad y \in \Re^n.$$

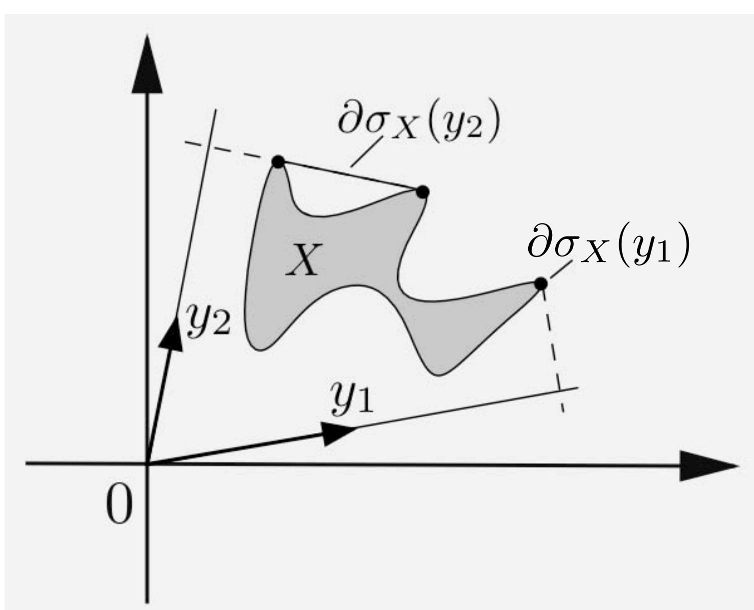
- We have $\partial\sigma_X(y) = \partial r(0) = \arg\min_{x\in\Re^n} r^\star(x)$.
- We have $r^\star(x) = \sup_{y\in\Re^n}\{y'x - r(y)\}$, or

$$r^\star(x) = \sup_{y\in\Re^n}\{y'x - \sigma_X(y + y)\} = \delta(x) - y'x,$$

where δ is the indicator function of $\mathrm{cl}\big(\mathrm{conv}(X)\big)$.

- Hence $\partial\sigma_X(y) = \arg\min_{x\in\Re^n} \delta(x) - y'x$, or

$$\partial\sigma_X(y) = \arg\max_{x\in\mathrm{cl}\left(\mathrm{conv}(X)\right)} y'x$$



EXAMPLE: SUBDIFF. OF POLYHEDRAL FN

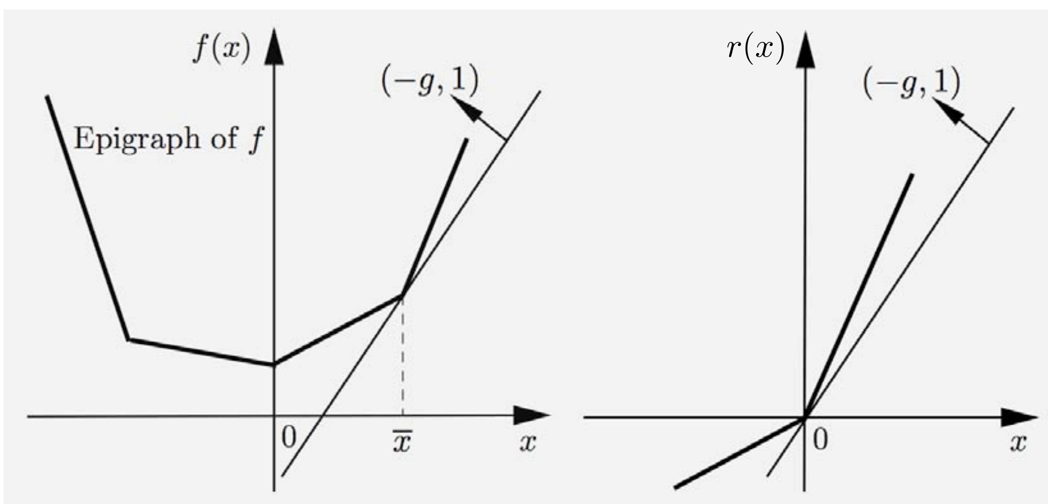
- Let

$$f(x) = \max\{a_1'x + b_1, \ldots, a_r'x + b_r\}.$$

- For a fixed $x \in \Re^n$, consider

$$A_x = \big\{j \mid a_j'x + b_j = f(x)\big\}$$

and the function $r(x) = \max\big\{a_j'x \mid j \in A_x\big\}$.



- It can be seen that $\partial f(x) = \partial r(0)$.

- Since r is the support function of the finite set $\{a_j \mid j \in A_x\}$, we see that

$$\partial f(x) = \partial r(0) = \operatorname{conv}\big(\{a_j \mid j \in A_x\}\big)$$

CHAIN RULE

- Let $f : \Re^m \mapsto (-\infty, \infty]$ be convex, and A be a matrix. Consider $F(x) = f(Ax)$ and assume that F is proper. If either f is polyhedral or else the range of $R(A) \cap \mathrm{ri}(\mathrm{dom}(f)) \neq \emptyset$, we have

$$\partial F(x) = A'\partial f(Ax), \qquad \forall\, x \in \Re^n.$$

Proof: Showing $\partial F(x) \supset A'\partial f(Ax)$ is simple and does not require the relative interior assumption. For the reverse inclusion, let $d \in \partial F(x)$ so $F(z) \geq F(x) + (z-x)'d \geq 0$ or $f(Az) - z'd \geq f(Ax) - x'd$ for all z, so (Ax, x) solves

$$\begin{aligned} &\text{minimize} \quad f(y) - z'd \\ &\text{subject to} \;\; y \in \mathrm{dom}(f), \quad Az = y. \end{aligned}$$

If $R(A) \cap \mathrm{ri}(\mathrm{dom}(f)) \neq \emptyset$, by strong duality theorem, there is a dual optimal solution λ, such that

$$(Ax, x) \in \arg \min_{y \in \Re^m,\, z \in \Re^n} \big\{ f(y) - z'd + \lambda'(Az - y) \big\}$$

Since the min over z is unconstrained, we have $d = A'\lambda$, so $Ax \in \arg\min_{y \in \Re^m} \big\{ f(y) - \lambda' y \big\}$, or

$$f(y) \geq f(Ax) + \lambda'(y - Ax), \qquad \forall\, y \in \Re^m.$$

Hence $\lambda \in \partial f(Ax)$, so that $d = A'\lambda \in A'\partial f(Ax)$. It follows that $\partial F(x) \subset A'\partial f(Ax)$. In the polyhedral case, $\mathrm{dom}(f)$ is polyhedral. **Q.E.D.**

SUM OF FUNCTIONS

- Let $f_i : \Re^n \mapsto (-\infty, \infty]$, $i = 1, \ldots, m$, be proper convex functions, and let

$$F = f_1 + \cdots + f_m.$$

- Assume that $\cap_{1=1}^m \text{ri}\big(\text{dom}(f_i)\big) \neq \emptyset$.
- Then

$$\partial F(x) = \partial f_1(x) + \cdots + \partial f_m(x), \qquad \forall\, x \in \Re^n.$$

Proof: We can write F in the form $F(x) = f(Ax)$, where A is the matrix defined by $Ax = (x, \ldots, x)$, and $f : \Re^{mn} \mapsto (-\infty, \infty]$ is the function

$$f(x_1, \ldots, x_m) = f_1(x_1) + \cdots + f_m(x_m).$$

Use the proof of the chain rule.

- **Extension:** If for some k, the functions f_i, $i = 1, \ldots, k$, are polyhedral, it is sufficient to assume

$$\Big(\cap_{i=1}^k \text{dom}(f_i)\Big) \cap \Big(\cap_{i=k+1}^m \text{ri}\big(\text{dom}(f_i)\big)\Big) \neq \emptyset.$$

CONSTRAINED OPTIMALITY CONDITION

- Let $f : \Re^n \mapsto (-\infty, \infty]$ be proper convex, let X be a convex subset of $\Re^n$, and assume that one of the following four conditions holds:

  (i) $\text{ri}\big(\text{dom}(f)\big) \cap \text{ri}(X) \neq \emptyset$.

  (ii) f is polyhedral and $\text{dom}(f) \cap \text{ri}(X) \neq \emptyset$.

  (iii) X is polyhedral and $\text{ri}\big(\text{dom}(f)\big) \cap X \neq \emptyset$.

  (iv) f and X are polyhedral, and $\text{dom}(f) \cap X \neq \emptyset$.

  Then, a vector x^* minimizes f over X iff there exists $g \in \partial f(x^*)$ such that $-g$ belongs to the normal cone $N_X(x^*)$, i.e.,

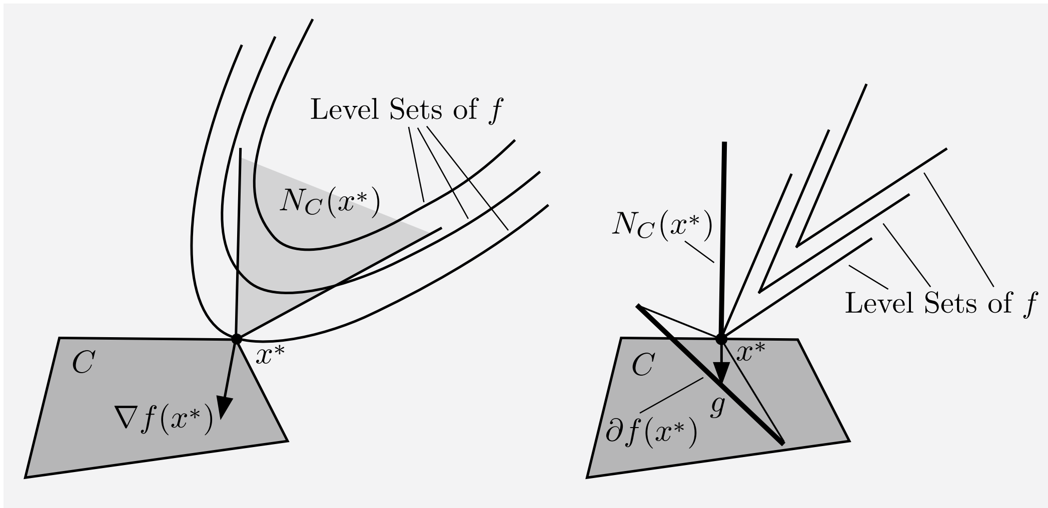
$$g'(x - x^*) \geq 0, \qquad \forall\, x \in X.$$

Proof: x^* minimizes

$$F(x) = f(x) + \delta_X(x)$$

if and only if $0 \in \partial F(x^*)$. Use the formula for subdifferential of sum. **Q.E.D.**

ILLUSTRATION OF OPTIMALITY CONDITION



- In the figure on the left, f is differentiable and the condition is that

$$-\nabla f(x^*) \in N_C(x^*),$$

which is equivalent to

$$\nabla f(x^*)'(x - x^*) \geq 0, \qquad \forall\, x \in X.$$

- In the figure on the right, f is nondifferentiable, and the condition is that

$$-g \in N_C(x^*) \quad \text{for some } g \in \partial f(x^*).$$