Optimization Lecture 5

Qingfu Zhang

Dept of CS , CityU

2024

Outline

Convex optimization problems

Some standard convex problems

Transforming problems

Convex optimization problems

Optimization problem in standard form

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

- $x \in \mathbf{R}^n$ is the optimization variable
- $f_0 : \mathbf{R}^n \to \mathbf{R}$ is the objective or cost function
- $f_i : \mathbf{R}^n \to \mathbf{R}$, $i = 1, \dots, m$, are the inequality constraint functions
- $h_i : \mathbf{R}^n \to \mathbf{R}$ are the equality constraint functions

why don't consider " $<$ " and " $>$ "?

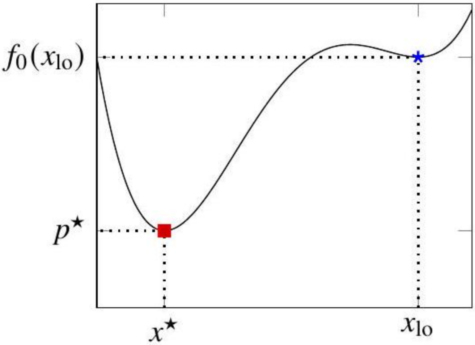
Feasible and optimal points

- $x \in \mathbf{R}^n$ is feasible if $x \in \operatorname{dom} f_0$ and it satisfies the constraints
- optimal value is
  $p^\star =$
  $\inf \{f_0(x) \mid f_i(x) \leq 0, i = 1, \ldots, m, h_k(x) = 0, k = 1, \ldots, p\}$
- $p^\star = \infty$ if problem is infeasible
- $p^\star = -\infty$ if problem is unbounded below
- a feasible x is optimal if $f_0(x) = p^\star$
- X_{opt}: the set of optimal points

Locally optimal points

x is **locally optimal** if there is an $R > 0$ such that x is optimal for

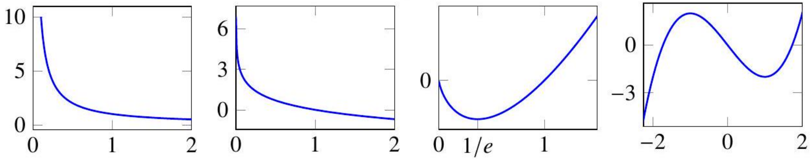
$$\begin{array}{ll} \text{minimize(over } z) & f_0(z) \\ \text{subject to} & f_i(z) \leq 0, \quad i = 1, \dots, m, \\ & h_i(z) = 0, \quad i = 1, \dots, p \\ & \|z - x\|_2 \leq R \end{array}$$



Examples

examples with $n = 1, m = p = 0$

- $f_0(x) = 1/x, \mathbf{dom}\, f_0 = \mathbf{R}_{++} : p^\star = 0$, no optimal point
- $f_0(x) = -\log x, \mathbf{dom}\, f_0 = \mathbf{R}_{++} : p^\star = -\infty$
- $f_0(x) = x \log x, \mathbf{dom}\, f_0 = \mathbf{R}_{++} : p^\star = -1/e, x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x : p^\star = -\infty, x = 1$ is locally optimal



$f_0(x) = 1/x$ $\quad$ $f_0(x) = -\log x$ $\quad$ $f_0(x) = x \log x$ $\quad$ $f_0(x) = x^3 - 3x$

Implicit and explicit constraints

standard form optimization problem has implicit constraint

$$x \in \mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom} f_i \cap \bigcap_{i=1}^{p} \operatorname{dom} h_i$$

- we call $\mathcal{D}$ the domain of the problem
- the constraints $f_i(x) \leq 0, h_i(x) = 0$ are the explicit constraints
- a problem is unconstrained if it has no explicit constraints ($m = p = 0$)

example:

$$\text{minimize } f_0(x) = -\sum_{i=1}^{k} \log\left(b_i - a_i^T x\right)$$

is an unconstrained problem with implicit constraints $a_i^T x < b_i$.

Feasibility problem

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

can be considered a special case of the general problem with $f_0(x) = 0$:

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

- $p^\star = 0$ if constraints are feasible; any feasible x is optimal.
- $p^\star = \infty$ if constraints are infeasible.

Standard form convex optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & a_i^T x = b_i, \quad i = 1, \ldots, p \end{array}$$

- objective and inequality constraints $f_0, f_1, \ldots, f_m$ are convex
- equality constraints are affine, often written as $Ax = b$
- feasible and optimal sets of a convex optimization problem are convex
- problem is quasiconvex if f_0 is quasiconvex, $f_1, \ldots, f_m$ are convex, $h_1, \ldots, h_p$ are affine

Example

- standard form problem

$$\begin{array}{ll} \text{minimize} & f_0(x) = x_1^2 + x_2^2 \\ \text{subject to} & f_1(x) = x_1/\left(1 + x_2^2\right) \leq 0 \\ & h_1(x) = (x_1 + x_2)^2 = 0 \end{array}$$

- f_0 is convex; feasible set $\{(x_1, x_2) \mid x_1 = -x_2 \leq 0\}$ is convex
- not a convex problem (by our definition) since f_1 is not convex, h_1 is not affine
- equivalent (but not identical) to the convex problem

$$\begin{array}{ll} \text{minimize} & x_1^2 + x_2^2 \\ \text{subject to} & x_1 \leq 0 \\ & x_1 + x_2 = 0 \end{array}$$

Local and global optima

Any locally optimal point of a convex problem is (globally) optimal

- suppose x is locally optimal, but there exists a feasible y with $f_0(y) < f_0(x)$
- x locally optimal means there is an $R > 0$ such that

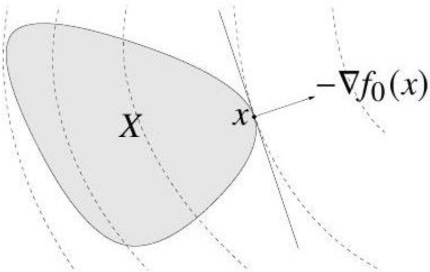
$$z \text{ feasible}, \quad \|z - x\|_2 \le R \quad \Longrightarrow \quad f_0(z) \ge f_0(x)$$

- consider $z = \theta y + (1 - \theta)x$ with $\theta = R/\left(2\|y - x\|_2\right)$
- $\|y - x\|_2 > R$, so $0 < \theta < 1/2$
- z is a convex combination of two feasible points, hence also feasible
- $\|z - x\|_2 = R/2$ and $f_0(z) \le \theta f_0(y) + (1 - \theta) f_0(x) < f_0(x)$, which contradicts our assumption that x is locally optimal

Optimality criterion for differentiable f_0

- x is optimal for a convex problem if and only if it is feasible and

$$\nabla f_0(x)^T(y - x) \geq 0 \text{ for all feasible } y$$



How to prove it?

- if nonzero, $\nabla f_0(x)$ defines a supporting hyperplane to feasible set X at x.

Examples:

- unconstrained problem (no implicit nor explicit): x minimizes $f_0(x)$ if and only if $\nabla f_0(x) = 0$
- equality constrained problem: x minimizes $f_0(x)$ subject to $Ax = b$ if and only if there exists a v such that

$$Ax = b, \quad \nabla f_0(x) + A^T v = 0$$

- minimization over nonnegative orthant: x minimizes $f_0(x)$ over $\mathbf{R}^n_+$ if and only if

$$x \geq 0, \quad \begin{cases} \nabla f_0(x)_i \geq 0 & x_i = 0 \\ \nabla f_0(x)_i = 0 & x_i > 0 \end{cases}$$

Some standard convex problems

Linear program (LP)

$$\begin{array}{ll} \text{minimize} & c^T x + d \\ \text{subject to} & Gx \leq h \\ & Ax = b \end{array}$$

- convex problem with affine objective and constraint functions
- feasible set is a polyhedron

Example: Diet problem

- Choose nonnegative quantities $x_1, \ldots, x_n$ of n foods
- one unit of food j costs c_j and contains amount A_{ij} of nutrient i
- healthy diet requires nutrient i in quantity at least b_i
- to find cheapest healthy diet, solve:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \geq b, \quad x \geq 0. \end{array}$$

- express in standard LP form as

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \begin{bmatrix} -A \\ -I \end{bmatrix} x \leq \begin{bmatrix} -b \\ 0 \end{bmatrix} \end{array}$$

Example: Piecewise-linear minimization

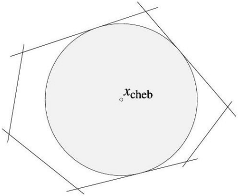
- minimize convex piecewise-linear function $f_0(x) = \max_{i=1,\ldots,m}\left(a_i^T x + b_i\right), x \in \mathbf{R}^n$
- equivalent to LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x + b_i \leq t, \quad i = 1, \ldots, m \end{array}$$

with variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$

Example: Chebyshev center of a polyhedron

Chebyshev center of $\mathcal{P} = \left\{x \mid a_i^T x \leq b_i, i = 1, \ldots, m\right\}$ is center of largest inscribed ball $\mathcal{B} = \{x_c + u \mid \|u\|_2 \leq r\}$



- $a_i^T x \leq b_i$ for all $x \in \mathcal{B}$ if and only if

$$\sup\left\{a_i^T (x_c + u) \mid \|u\|_2 \leq r\right\} = a_i^T x_c + r \left\|a_i\right\|_2 \leq b_i$$

- x_c, r can be determined by solving LP with variables x_c, r

$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x_c + r \left\|a_i\right\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array}$$

Quadratic program (QP)

$$
\begin{array}{ll}
\text{minimize} & (1/2)x^T P x + q^T x + r \\
\text{subject to} & Gx \preceq h \\
& Ax = b
\end{array}
$$

- $P \in \mathbf{S}^n_+$, so objective is convex quadratic
- minimize a convex quadratic function over a polyhedron

Example: Least squares

- least squares problem: minimize $\|Ax - b\|_2^2$
- analytical solution $x^\star = A^\dagger b$ $\left(A^\dagger\right.$ is pseudo-inverse $\left.\right)$
- can add linear constraints, e.g.,
- $x \geq 0$ (nonnegative least squares)
- $x_1 \leq x_2 \leq \cdots \leq x_n$ (isotonic regression)

Example: Linear program with random cost

- LP with random cost c, with mean $\bar{c}$ and covariance Σ
- hence, LP objective $c^T x$ is random variable with mean $\bar{c}^T x$ and variance $x^T \Sigma x$
- risk-averse problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{E} c^T x + \gamma \operatorname{var}\left(c^T x\right) \\ \text{subject to} & Gx \leq h, \quad Ax = b \end{array}$$

- $\gamma > 0$ is risk aversion parameter; controls the trade-off between expected cost and variance (risk)
- express as QP

$$\begin{array}{ll} \text{minimize} & \bar{c}^T x + \gamma x^T \Sigma x \\ \text{subject to} & Gx \leq h, \quad Ax = b \end{array}$$

Quadratically constrained quadratic program (QCQP)

$$
\begin{array}{ll}
\text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 \\
\text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \dots, m \\
& Ax = b
\end{array}
$$

- $P_i \in \mathbf{S}^n_+$; objective and constraints are convex quadratic
- if $P_1, \dots, P_m \in \mathbf{S}^n_{++}$, feasible region is intersection of m ellipsoids and an affine set

Second-order cone programming

$$
\begin{array}{ll}
\text{minimize} & f^T x \\
\text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m \\
& Fx = g
\end{array}
$$

($A_i \in \mathbf{R}^{n_i \times n}$, $F \in \mathbf{R}^{p \times n}$)

- inequalities are called second-order cone (SOC) constraints:

$$
\left(A_i x + b_i, c_i^T x + d_i\right) \in \text{ second-order cone in } \mathbf{R}^{n_i+1}
$$

- for $n_i = 0$, reduces to an LP; if $c_i = 0$, reduces to a QCQP
- more general than QCQP and LP

Example: Robust linear programming

suppose constraint vectors a_i are uncertain in the LP

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

two common approaches to handling uncertainty

- deterministic worst-case: constraints must hold for all $a_i \in \mathcal{E}_i$ (uncertainty ellipsoids)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \text{ for all } a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m \end{array}$$

- stochastic: a_i is random variable; constraints must hold with probability η

minimize $c^T x$
subject to $\text{prob}\left(a_i^T x \leq b_i\right) \geq \eta, \quad i = 1, \ldots, m$

Deterministic worst-case approach

- uncertainty ellipsoids are
  $\mathcal{E}_i = \{\bar{a}_i + P_i u \mid \|u\|_2 \leq 1\}, (\bar{a}_i \in \mathbf{R}^n, P_i \in \mathbf{R}^{n \times n})$
- center of $\mathcal{E}_i$ is $\bar{a}_i$; semi-axes determined by singular values/vectors of P_i
- robust LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \quad \forall a_i \in \mathcal{E}_i, \quad i = 1, \ldots, m \end{array}$$

- equivalent to SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \left\|P_i^T x\right\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array}$$

(follows from $\sup_{\|u\|_2 \leq 1} (\bar{a}_i + P_i u)^T x = \bar{a}_i^T x + \left\|P_i^T x\right\|_2$)

Stochastic approach

- assume $a_i \sim \mathcal{N}(\bar{a}_i, \Sigma_i)$
- $a_i^T x \sim \mathcal{N}\left(\bar{a}_i^T x, x^T \Sigma_i x\right)$, so

$$\operatorname{prob}\left(a_i^T x \leq b_i\right) = \Phi\left(\frac{b_i - \bar{a}_i^T x}{\left\|\Sigma_i^{1/2} x\right\|_2}\right)$$

where $\Phi(u) = (1/\sqrt{2\pi}) \int_{-\infty}^{u} e^{-t^2/2} dt$ is $\mathcal{N}(0,1)$ CDF

- $\textbf{\textit{prob}}\left(a_i^T x \leq b_i\right) \geq \eta$ can be expressed as $\bar{a}_i^T x + \Phi^{-1}(\eta)\left\|\Sigma_i^{1/2} x\right\|_2 \leq b_i$
- for $\eta \geq 1/2$, robust LP equivalent to SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{a}_i^T x + \Phi^{-1}(\eta)\left\|\Sigma_i^{1/2} x\right\|_2 \leq b_i, \quad i = 1, \ldots, m \end{array}$$

Conic form problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Fx + g \preceq_K 0 \\ & Ax = b \end{array}$$

- constraint $Fx + g \preceq_K 0$ involves a generalized inequality with respect to a proper cone K
- linear programming is a conic form problem with $K = \mathbf{R}^m_+$
- as with standard convex problem
- feasible and optimal sets are convex
- any local optimum is global

Semidefinite program (SDP)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b \end{array}$$

with $F_i, G \in \mathbf{S}^k$

- inequality constraint is called linear matrix inequality (LMI)
- includes problems with multiple LMI constraints: for example,

$$x_1 \hat{F}_1 + \cdots + x_n \hat{F}_n + \hat{G} \preceq 0, \quad x_1 \tilde{F}_1 + \cdots + x_n \tilde{F}_n + \tilde{G} \preceq 0$$

is equivalent to single LMI

$$x_1 \begin{bmatrix} \hat{F}_1 & 0 \\ 0 & \tilde{F}_1 \end{bmatrix} + x_2 \begin{bmatrix} \hat{F}_2 & 0 \\ 0 & \tilde{F}_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \hat{F}_n & 0 \\ 0 & \tilde{F}_n \end{bmatrix} + \begin{bmatrix} \hat{G} & 0 \\ 0 & \tilde{G} \end{bmatrix} \preceq 0$$

Example: Matrix norm minimization

$$\text{minimize } \|A(x)\|_2 = \left(\lambda_{\max}\left(A(x)^T A(x)\right)\right)^{1/2}$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_n A_n$ (with given $A_i \in \mathbf{R}^{p\times q}$)
equivalent SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} t\mathcal{I} & A(x) \\ A(x)^T & t\mathcal{I} \end{array}\right] \succeq 0 \end{array}$$

variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$

constraint follows from

$$\|A\|_2 \leq t \iff A^T A \preceq t^2\mathcal{I}, \quad t \geq 0$$
$$\iff \left[\begin{array}{cc} t\mathcal{I} & A \\ A^T & t\mathcal{I} \end{array}\right] \succeq 0$$

Transforming problems

Change of variables

- $\phi : \mathbf{R}^n \to \mathbf{R}^n$ is one-to-one with $\phi(\mathbf{dom}\,\phi) \supseteq \mathcal{D}$
- consider (possibly non-convex) problem

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

- change variables to z with $x = \phi(z)$
- can solve equivalent problem

$$
\begin{array}{ll}
\text{minimize} & \tilde{f}_0(z) \\
\text{subject to} & \tilde{f}_i(z) \leq 0, \quad i = 1, \ldots, m \\
& \tilde{h}_i(z) = 0, \quad i = 1, \ldots, p
\end{array}
$$

where $\tilde{f}_i(z) = f_i(\phi(z))$ and $\tilde{h}_i(z) = h_i(\phi(z))$

- recover original optimal point as $x^\star = \phi(z^\star)$

Example

- non-convex problem

$$\begin{array}{ll} \text{minimize} & x_1/x_2 + x_3/x_1 \\ \text{subject to} & x_2/x_3 + x_1 \leq 1 \\ & x > 0 \end{array}$$

- change variables using $x = \phi(z) = \exp z$ to get

$$\begin{array}{ll} \text{minimize} & \exp(z_1 - z_2) + \exp(z_3 - z_1) \\ \text{subject to} & \exp(z_2 - z_3) + \exp(z_1) \leq 1 \end{array}$$

which is convex.

Transformation of objective and constraint functions

suppose

- ϕ_0 is monotone increasing
- $\psi_i(u) \leq 0$ if and only if $u \leq 0, i = 1, \ldots, m$
- $\varphi_i(u) = 0$ if and only if $u = 0, i = 1, \ldots, p$

standard form optimization problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \phi_0\left(f_0(x)\right) \\ \text{subject to} & \psi_i\left(f_i(x)\right) \leq 0, \quad i = 1, \ldots, m \\ & \varphi_i\left(h_i(x)\right) = 0, \quad i = 1, \ldots, p \end{array}$$

example: minimizing $\|Ax - b\|$ is equivalent to minimizing $\|Ax - b\|^2$

Converting maximization to minimization

- suppose ϕ_0 is monotone decreasing
- the maximization problem

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

is equivalent to the minimization problem

$$\begin{array}{ll} \text{minimize} & \phi_0\left(f_0(x)\right) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \dots, m \\ & h_i(x) = 0, \quad i = 1, \dots, p \end{array}$$

examples:

- $\phi_0(u) = -u$ transforms maximizing a concave function to minimizing a convex function
- $\phi_0(u) = 1/u$ transforms maximizing a concave positive function to minimizing a convex function

Eliminating equality constraints

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

is equivalent to

$$\begin{array}{ll} \text{minimize(over } z) & f_0\left(Fz + x_0\right) \\ \text{subject to} & f_i\left(Fz + x_0\right) \leq 0, \quad i = 1, \dots, m \end{array}$$

where F and x_0 are such that $Ax = b \Longleftrightarrow x = Fz + x_0$ for some z

Introducing equality constraints

$$\begin{array}{ll} \text{minimize} & f_0\left(A_0 x + b_0\right) \\ \text{subject to} & f_i\left(A_i x + b_i\right) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize}\left(\text{ over } x, y_i\right) & f_0\left(y_0\right) \\ \text{subject to} & f_i\left(y_i\right) \leq 0, \quad i = 1, \ldots, m \\ & y_i = A_i x + b_i, \quad i = 0, 1, \ldots, m \end{array}$$

Introducing slack variables for linear inequalities

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize(over } x, s) & f_0(x) \\ \text{subject to} & a_i^T x + s_i = b_i, \quad i = 1, \ldots, m \\ & s_i \geq 0, \quad i = 1, \ldots m \end{array}$$

Epigraph form

standard form convex problem is equivalent to

$$\begin{array}{ll} \text{minimize(over } x, t) & t \\ \text{subject to} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

Minimizing over some variables

$$\begin{array}{ll} \text{minimize} & f_0(x_1, x_2) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x_1) \\ \text{subject to} & f_i(x_1) \leq 0, \quad i = 1, \ldots, m \end{array}$$

where $\tilde{f}_0(x_1) = \inf_{x_2} f_0(x_1, x_2)$