problems:

$$p^* = f(x) + g(Ax) \to \min_{x \in E \cap A^{-1}(G)}$$

$$d^* = f^*(-A^*\lambda) + g^*(\lambda) \to \min_{\lambda \in G_* \cap (-A^*)^{-1}(E_*)},$$

Then we have weak duality: $p^* \geq d^*$. Furthermore, if the functions f and g are convex and $A(\mathbf{relint}(E)) \cap \mathbf{relint}(G) \neq \varnothing$, then we have strong duality: $p^* = d^*$. While points $x^* \in E \cap A^{-1}(G)$ and $\lambda^* \in G_* \cap (-A^*)^{-1}(E_*)$ are optimal values for primal and dual problem if and only if:

$$\begin{aligned} -A^*\lambda^* &\in \partial f(x^*) \\ \lambda^* &\in \partial g(Ax^*) \end{aligned}$$

Convex case is especially important since if we have Fenchel - Rockafellar problem with parameters (f, g, A), than the dual problem has the form $(f^*, g^*, -A^*)$.

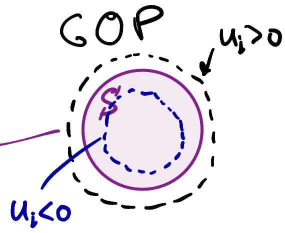
5.2 Sensitivity analysis

$p^* = d^*$

Let us switch from the original optimization problem

p^*

$$\boxed{\begin{aligned} &f_0(x) \to \min_{x \in \mathbb{R}^n} \\ \text{s.t. } &f_i(x) \leq 0, \; i = 1, \dots, m \\ &h_i(x) = 0, \; i = 1, \dots, p \end{aligned}} \tag{P}$$



Пусть есть сильная двойств.

To the perturbed version of it:

возмущенная задача

$p^*(u,v)$

$p^*(0,0) = p^*$

$$\boxed{\begin{aligned} &f_0(x) \to \min_{x \in \mathbb{R}^n} \\ \text{s.t. } &f_i(x) \leq u_i, \; i = 1, \dots, m \\ &h_i(x) = v_i, \; i = 1, \dots, p \end{aligned}} \tag{Per}$$

Note, that we still have the only variable $x \in \mathbb{R}^n$, while treating $u \in \mathbb{R}^m, v \in \mathbb{R}^p$ as parameters. It is obvious, that $\text{Per}(u, v) \to \text{P}$ if $u = 0, v = 0$. We will denote the optimal value of Per as $p^*(u, v)$, while the optimal value of the original problem P is just p^*. One can immediately say, that $p^*(u, v) = p^*$.

Speaking of the value of some i-th constraint we can say, that

- $\boxed{u_i = 0}$ leaves the original problem
- $\boxed{u_i > 0}$ means that we have relaxed the inequality
- $\boxed{u_i < 0}$ means that we have tightened the constraint

$f_i(x) \leq u_i$
$h_i(x) = v_i$

One can even show, that when P is convex optimization problem, $p^*(u, v)$ is a convex function.

Suppose, that strong duality holds for the orriginal problem and suppose, that x is any feasible point for the perturbed problem:

$$\begin{aligned} p^*(0,0) = p^* = \boxed{d^* = g(\lambda^*, \nu^*)} &\leq \quad \text{по опр.} \\ \leq L(x, \lambda^*, \nu^*) &= \\ = f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) &\leq \\ \leq f_0(x) + \sum_{i=1}^m \lambda_i^* u_i + \sum_{i=1}^p \nu_i^* v_i &. \end{aligned}$$

Which means

$$\boxed{f_0(x) \geq p^*(0,0) - {\lambda^*}^T u - {\nu^*}^T v}$$

And taking the optimal x for the perturbed problem, we have:

$$\boxed{p^*(u,v) \geq p^*(0,0) - \lambda^{*T}u - \nu^{*T}v} \quad (S) \tag{4}$$

In scenarios where strong duality holds, we can draw several insights about the sensitivity of optimal solutions in relation to the Lagrange multipliers. These insights are derived from the inequality expressed in equation above:

$\lambda_i^* \geq 0$ $\quad \lambda_i^* >> 0$ $\quad$ ухудшим задачку $u_i < 0$

1. **Impact of Tightening a Constraint (Large $\lambda_i^\star$):**
   When the ith constraint's Lagrange multiplier, $\lambda_i^\star$, holds a substantial value, and if this constraint is tightened (choosing $u_i < 0$), there is a guarantee that the optimal value, denoted by $p^\star(u,v)$, will significantly increase.

2. **Effect of Adjusting Constraints with Large Positive or Negative $\nu_i^\star$:**
   - If $\nu_i^\star$ is large and positive and $v_i < 0$ is chosen, or ← сожмём огр
   - If $\nu_i^\star$ is large and negative and $v_i > 0$ is selected, ← ослабить огр.

   then in either scenario, the optimal value $p^\star(u,v)$ is expected to increase greatly.

$$\min f_0(x)$$
$$f_i(x) \leq 0$$

$h_i(x) = 0 \iff h_i(x) \leq 0$, $-h_i(x) \leq 0$

3. **Consequences of Loosening a Constraint (Small $\lambda_i^\star$):**
   If the Lagrange multiplier $\lambda_i^\star$ for the ith constraint is relatively small, and the constraint is loosened (choosing $u_i > 0$), it is anticipated that the optimal value $p^\star(u,v)$ will not significantly decrease.

4. **Outcomes of Tiny Adjustments in Constraints with Small $\nu_i^\star$:**
   - When $\nu_i^\star$ is small and positive, and $v_i > 0$ is chosen, or
   - When $\nu_i^\star$ is small and negative, and $v_i < 0$ is opted for,

   in both cases, the optimal value $p^\star(u,v)$ will not significantly decrease.

These interpretations provide a framework for understanding how changes in constraints, reflected through their corresponding Lagrange multipliers, impact the optimal solution in problems where strong duality holds.

5.3 Local sensitivity

Suppose now that $p^*(u,v)$ is differentiable at $u=0, v=0$.

$$\boxed{\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i} \qquad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}} \tag{5}$$

To show this result we consider the directional derivative of $p^*(u,v)$ along the direction of some i-th basis vector e_i:

производная по напр e_i

$$\boxed{\lim_{t\to 0} \frac{p^*(te_i,0) - p^*(0,0)}{t} = \frac{\partial p^*(0,0)}{\partial u_i}}$$

$\lambda^* = -\nabla_u p^*$

From the inequality Equation 4 and taking the limit $t \to 0$ with $t > 0$ we have

i-ый базисн

$$\boxed{\frac{p^*(te_i,0) - p^*}{t} \geq -\lambda_i^*} \to \frac{\partial p^*(0,0)}{\partial u_i} \geq -\lambda_i^* \to \quad \nu^* = -\nabla_v p^*$$

For the negative $t < 0$ we have:

$$\boxed{\frac{p^*(te_i,0) - p^*}{t} \leq -\lambda_i^*} \to \frac{\partial p^*(0,0)}{\partial u_i} \leq -\lambda_i^*$$

The same idea can be used to establish the fact about v_i.

The local sensitivity result Equation 5 provides a way to understand the impact of constraints on the optimal solution x^* of an optimization problem. If a constraint $f_i(x^*)$ is negative at x^*, it's not affecting the optimal solution, meaning small changes to this constraint won't alter the optimal value. In this case, the corresponding optimal Lagrange multiplier will be zero, as per the principle of complementary slackness.

However, if $f_i(x^*) = 0$, meaning the constraint is precisely met at the optimum, then the situation is different. The value of the i-th optimal Lagrange multiplier, $\lambda_i^\star$, gives us insight into how 'sensitive' or 'active' this constraint is. A small $\lambda_i^\star$ indicates that slight adjustments to the constraint won't significantly affect the optimal value. Conversely, a large $\lambda_i^\star$ implies that even minor changes to the constraint can have a significant impact on the optimal solution.

5.4 Shadow prices or tax interpretation

$f_k(x) \leq 0$ – НОЛОГ

$L(x, \lambda) = f_0 + \lambda^T f(x) \leq f_0 + \lambda^T \cdot$?

$t \leq u$

PROFIT $\boxed{\lambda^T \cdot u}$

Consider an enterprise where x represents its operational strategy and $f_0(x)$ is the operating cost. Therefore, $-f_0(x)$ denotes the profit in dollars. Each constraint $f_i(x) \leq 0$ signifies a resource or regulatory limit. The goal is to maximize profit while adhering to these limits, which is equivalent to solving:

БИЗНЕС

Функция Лагр.

$L(x, \lambda) = f_0(x) + \lambda^T f(x)$

если нарушаем стоимость работы с учётом линейных штрафов

ничего не нарушено $x \in S$

$f_0(x)$

$$
\begin{aligned}
& f_0(x) \rightarrow \min_{x \in \mathbb{R}^n} \\
& \text{s.t. } f_i(x) \leq 0, \ i = 1, \ldots, m
\end{aligned}
$$

$f_0(x)$ – стоимость работы предприят

$-f_0(x)$ – ПРИБЫЛЬ

$f_i(x)$ – ОГРАНИЧЕНИЯ

The optimal profit here is $-p^*$.

λ_i^* – оптимальная стоимость нарушения i-ого

Now, imagine a scenario where exceeding limits is allowed, but at a cost. This cost is linear to the extent of violation, quantified by f_i. The charge for breaching the i^{th} constraint is $\lambda_i f_i(x)$. If $f_i(x) < 0$, meaning the constraint is not fully utilized, $\lambda_i f_i(x)$ represents income for the firm. Here, λ_i is the cost (in dollars) per unit of violation for $f_i(x)$.

$g(\lambda) = ?$ – лучшее, что можно сделать при текущих штрафах

For instance, if $f_1(x) \leq 0$ limits warehouse space, the firm can rent out extra space at λ_1 dollars per square meter or rent out unused space for the same rate.

$p^* \geq d^*$

$g(\lambda) \rightarrow \max_\lambda$ | d^* – лучшая стоимость функц. предпр в худших условиях

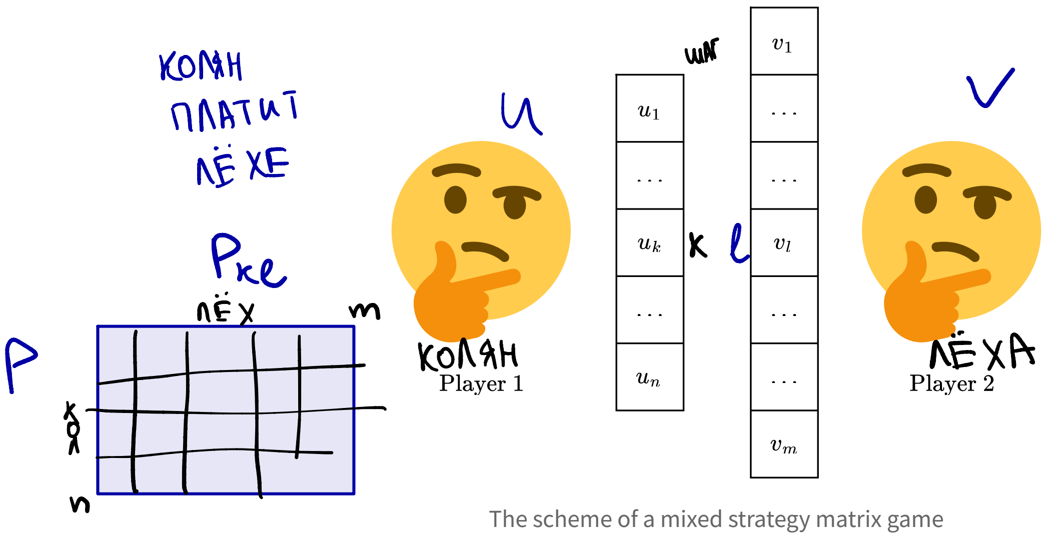
The firm's total cost, considering operational and constraint costs, is $L(x, \lambda) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x)$. The firm aims to minimize $L(x, \lambda)$, resulting in an optimal cost $g(\lambda)$. The dual function $g(\lambda)$ represents the best possible cost for the firm based on the prices of constraints λ, and the optimal dual value d^* is this cost under the most unfavorable price conditions.

$d^* = p^*$ если в бизнесе сильная двойственность

Weak duality implies that the cost in this flexible scenario (where the firm can trade constraint violations) is always less than or equal to the cost in the strict original scenario. This is because any optimal operation x^* from the original scenario will cost less in the flexible scenario, as the firm can earn from underused constraints.

If strong duality holds and the dual optimum is reached, the optimal λ^* represents prices where the firm gains no extra advantage from trading constraint violations. These optimal λ^* values are often termed 'shadow prices' for the original problem, indicating the hypothetical cost of constraint flexibility.

5.5 Mixed strategies for matrix games



The scheme of a mixed strategy matrix game

In zero-sum matrix games, players 1 and 2 choose actions from sets $\{1, \ldots, n\}$ and $\{1, \ldots, m\}$, respectively. The outcome is a payment from player 1 to player 2, determined by a payoff matrix $P \in \mathbb{R}^{n \times m}$. Each player aims to use mixed strategies, choosing actions according to a probability distribution: player 1 uses probabilities u_k for each action i, and player 2 uses v_l.

The expected payoff from player 1 to player 2 is given by $\boxed{\sum_{k=1}^n \sum_{l=1}^m u_k v_l P_{kl}} = \boxed{u^T P v}$. Player 1 seeks to minimize this expected payoff, while player 2 aims to maximize it.

$\mathbb{E}(\text{КОЛЯН} \rightarrow \text{ЛЁХЕ})$

вер-ть выбор k строку

вер-ть выбор l столбц

5.5.1 КОЛЯН [redacted] Perspective

Assuming player 2 knows player 1's strategy u, player 2 will choose v to maximize $u^T P v$. The worst-case expected payoff is thus:

3 4 7

$3v_1 + 4v_2 + 7v_3$

$$\boxed{\max_{v \geq 0, 1^T v = 1} u^T P v} = \boxed{\max_{i=1,\ldots,m} (P^T u)_i}$$

ВЫБОР КОЛЯНА

Player 1's optimal strategy minimizes this worst-case payoff, leading to the optimization problem:

КОЛЯН PROBLEM

КОЛЯН ХОДИТ

$$\begin{aligned} &\min \left(\max_{i=1,\ldots,m} (P^T u)_i \right) \\ &\text{s.t. } u \geq 0 \\ &1^T u = 1 \end{aligned}$$

рациональная стратегия ЛЁХИ

$$\begin{aligned} &\min t \\ &P^T u \leq t \cdot 1 \\ &u \geq 0 \\ &1^T u = 1 \end{aligned} \tag{6}$$

This forms a convex optimization problem with the optimal value denoted as $\boxed{p_1^*}$.

5.5.2 Player 2's Perspective

Conversely, if player 1 knows player 2's strategy v, the goal is to minimize $u^T P v$. This leads to:

РАЦ. СТРАТ. КОЛЯНА

$$\left(\min_{u \geq 0, 1^T u = 1} u^T P v = \min_{i=1,\ldots,n} (P v)_i \right)$$

Player 2 then maximizes this to get the largest guaranteed payoff, solving the optimization problem:

ЛЁХА PROBLEM

$$\begin{aligned} &\max \left[\min_{i=1,\ldots,n} (P v)_i \right] \\ &\text{s.t. } v \geq 0 \\ &1^T v = 1 \end{aligned} \tag{7}$$

РАЦ. СТРАТ. КОЛЯНА

The optimal value here is p_2^*.

5.5.3 Duality and Equivalence

It's generally advantageous to know the opponent's strategy, but surprisingly, in mixed strategy matrix games, this advantage disappears. The key lies in duality: the problems above are Lagrange duals. By formulating player 1's problem as a linear program and introducing Lagrange multipliers, we find that the dual problem matches player 2's problem. Due to strong duality in feasible linear programs, $\boxed{p_1^* = p_2^*}$, showing no advantage in knowing the opponent's strategy.

5.5.4 Formulating and Solving the Lagrange Dual

We approach problem Equation 6 by setting it up as a linear programming (LP) problem. The goal is to minimize a variable t, subject to certain constraints:

1. $u \geq 0$,
2. The sum of elements in u equals 1 ($1^T u = 1$),
3. $P^T u$ is less than or equal to t times a vector of ones ($P^T u \leq t\mathbf{1}$).

$$\begin{aligned} &\min t \\ &P^T u \leq t \cdot 1 \\ &u \geq 0 \\ &1^T u = 1 \end{aligned}$$

$$\max_i (P^T u) = t$$

Here, t is an additional variable in the real numbers ($t \in \mathbb{R}$).

5.5.5 Constructing the Lagrangian

We introduce multipliers for the constraints: λ for $\boxed{P^T u \leq t\mathbf{1}}$, μ for $u \geq 0$, and ν for $1^T u = 1$. The Lagrangian is then formed as:

$$\boxed{L = t + \lambda^T (P^T u - t\mathbf{1}) - \mu^T u + \nu(1 - 1^T u)} = \nu + (1 - 1^T \lambda) t + (P\lambda - \nu\mathbf{1} - \mu)^T u$$

5.5.6 Defining the Dual Function

The dual function $g(\lambda, \mu, \nu)$ is defined as:

$$\boxed{g(\lambda, \mu, \nu) = \begin{cases} \nu & \text{if } 1^T \lambda = 1 \text{ and } P\lambda - \nu\mathbf{1} = \mu \\ -\infty & \text{otherwise} \end{cases}}$$

5.5.7 Solving the Dual Problem

$\nu \to \max$

The dual problem seeks to maximize ν under the following conditions:

$1^T \lambda$

$P_1^* = P_2^*$

1. $\lambda \geq 0$,
2. The sum of elements in λ equals 1 ($1^T \lambda = 1$),
3. $\mu \geq 0$,
4. $P\lambda - \nu \mathbf{1} = \mu$.

Upon eliminating μ, we obtain the Lagrange dual of Equation 6:

$$\boxed{\begin{aligned} &\max \nu \\ &\text{s.t. } \lambda \geq 0 \\ &1\lambda^T = 1 \\ &P\lambda \geq \nu \mathbf{1} \end{aligned}}$$

= ΛЁXA PROBLEM

5.5.8 Conclusion

This formulation shows that the Lagrange dual problem is equivalent to problem Equation 7. Given the feasibility of these linear programs, strong duality holds, meaning the optimal values of Equation 6 and Equation 7 are equal.

6 References

- Convex Optimization — Boyd & Vandenberghe @ Stanford
- Course Notes for EE227C. Lecture 13
- Course Notes for EE227C. Lecture 14
- Optimality conditions
- Seminar 7 @ CMC MSU
- Seminar 8 @ CMC MSU
- Convex Optimization @ Berkeley - 10th lecture