# Conic Duality Theorems and Applications

Yinyu Ye

Department of Management Science and Engineering

Stanford University

Stanford, CA 94305, U.S.A.

http://www.stanford.edu/~yyye

Chapters 4.1-4.2 and 6.1-6.4

## Primal and Dual of Conic LP

Recall the pair of

$$
\begin{array}{lll}
(CLP) & \text{minimize} & \mathbf{c} \bullet \mathbf{x} \\
& \text{subject to} & \mathbf{a}_i \bullet \mathbf{x} = b_i, i = 1, 2, ..., m, \ \mathbf{x} \in K;
\end{array}
$$

and it dual problem

$$
\begin{array}{lll}
(CLD) & \text{maximize} & \mathbf{b}^T \mathbf{y} \\
& \text{subject to} & \sum_i^m y_i \mathbf{a}_i + \mathbf{s} = \mathbf{c}, \ \mathbf{s} \in K^*,
\end{array}
$$

where $\mathbf{y} \in \mathcal{R}^m$, $\mathbf{s}$ is called the dual slack vector/matrix, and $K^*$ is the dual cone of $K$.

**Theorem 1** *(Weak duality theorem)*

$$\mathbf{c} \bullet \mathbf{x} - \mathbf{b}^T \mathbf{y} = \mathbf{x} \bullet \mathbf{s} \geq 0$$

*for any feasible $\mathbf{x}$ of (CLP) and $(\mathbf{y}, \mathbf{s})$ of (CLD).*

## CLP Duality Theorems

The weak duality theorem shows that a feasible solution to either problem yields a bound on the value of the other problem. We call $\mathbf{c} \bullet \mathbf{x} - \mathbf{b}^T\mathbf{y}$ the duality gap.

**Corollary 1** *Let $\mathbf{x}^* \in \mathcal{F}_p$ and $(\mathbf{y}^*, \mathbf{s}^*) \in \mathcal{F}_d$. Then, $\mathbf{c} \bullet x^* = \mathbf{b}^T\mathbf{y}^*$ implies that $\mathbf{x}^*$ is optimal for (CLP) and $(\mathbf{y}^*, \mathbf{s}^*)$ is optimal for (CLD).*

Is the reverse also true? That is, given $\mathbf{x}^*$ optimal for (CLP), then there is $(\mathbf{y}^*, \mathbf{s}^*)$ feasible for (CLD) and $\mathbf{c} \bullet \mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$?

This is called the Strong Duality Theorem.

“True” when $K = \mathcal{R}^n_+$, that is, the polyhedral cone case.

## Proof of Strong Duality Theorem for LP

Let (LP) have a minimizer $\mathbf{x}^* \in \mathcal{F}_p$. Then, the system

$$A\mathbf{x}' - \mathbf{b}\tau = \mathbf{0}, \quad (\mathbf{x}';\tau) \geq \mathbf{0}, \quad \mathbf{c}^T\mathbf{x}' - (\mathbf{c}^T\mathbf{x}^*)\tau = -1 < 0$$

must have no feasible solution $(\mathbf{x}';\tau)$. This is because otherwise, if $\tau > 0$, $\mathbf{x}'/\tau$ is feasible for (LP) and $\mathbf{c}^T\mathbf{x}'/\tau < \mathbf{c}^T\mathbf{x}^*$, which is a contradiction; and if $\tau = 0$, $\mathbf{x}^* + \mathbf{x}'$ is feasible for (LP) and $\mathbf{c}^T(\mathbf{x}^* + \mathbf{x}') = \mathbf{c}^T\mathbf{x}^* - 1 < \mathbf{c}^T\mathbf{x}^*$, which is also a contradiction. Thus, from the LP alternative system pair II, there is $\mathbf{y}^*$ feasible for

$$\mathbf{c} - A^T\mathbf{y}^* \geq \mathbf{0}, \quad -\mathbf{c}^T\mathbf{x}^* + \mathbf{b}^T\mathbf{y}^* \geq 0.$$

Then, $\mathbf{y}^*$ is feasible for (LD) from the first inequality; and from the weak duality theorem and the second inequality $\mathbf{c}^T\mathbf{x}^* - \mathbf{b}^T\mathbf{y}^* = 0$.

## LP and LD Cases

**Theorem 2** *The following statements hold for every pair of (LP) and (LD) :*

**i)** *If (LP) and (LD) are both feasible, then both problems have optimal solutions and the optimal objective values of the objective functions are equal, that is, optimal solutions for both (LP) and (LD) exist and there is no duality gap.*

**ii)** *If (LP) or (LD) is feasible and bounded, then the other is feasible and bounded.*

**iii)** *If (LP) or (LD) is feasible and unbounded, then the other has no feasible solution.*

**iv)** *If (LP) or (LD) is infeasible, then the other is either unbounded or has no feasible solution.*

A case that neither (LP) nor (LD) is feasible: $\mathbf{c} = (-1;\ 0), \quad A = (0,\ -1), \quad b = 1.$

The proofs follow the Farkas lemma and the Weak Duality Theorem.

## Optimality Conditions for LP

$$\left\{ (\mathbf{x},\mathbf{y},\mathbf{s}) \in (\mathcal{R}^n_+, \mathcal{R}^m, \mathcal{R}^n_+) : \begin{array}{rcl} \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y} & = & 0 \\ A\mathbf{x} & = & \mathbf{b} \\ -A^T\mathbf{y} - \mathbf{s} & = & -\mathbf{c} \end{array} \right\},$$

which is a system of linear inequalities and equations. Now it is easy to verify whether or not a pair $(\mathbf{x},\mathbf{y},\mathbf{s})$ is optimal.

## Complementarity Gap

For feasible $\mathbf{x}$ and $(\mathbf{y}, \mathbf{s})$, $\mathbf{x}^T\mathbf{s} = \mathbf{x}^T(\mathbf{c} - A^T\mathbf{y}) = \mathbf{c}^T\mathbf{x} - \mathbf{b}^T\mathbf{y}$ is called the complementarity gap.

If $\mathbf{x}^T\mathbf{s} = 0$, then we say $\mathbf{x}$ and $\mathbf{s}$ are complementary to each other.

Since both $\mathbf{x}$ and $\mathbf{s}$ are nonnegative, $\mathbf{x}^T\mathbf{s} = 0$ implies that $\mathbf{x}. * \mathbf{s} = \mathbf{0}$ or $x_j s_j = 0$ for all $j = 1, \ldots, n$.

$$
\begin{aligned}
\mathbf{x}. * \mathbf{s} &= \mathbf{0} \\
A\mathbf{x} &= \mathbf{b} \\
-A^T\mathbf{y} - \mathbf{s} &= -\mathbf{c}.
\end{aligned}
$$

This system has total $2n + m$ unknowns and $2n + m$ equations including $n$ nonlinear equations.

## General CLP: an SDP Example with a Duality Gap

The strong duality theorem may not hold for general convex cones:

$$\mathbf{c} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{a}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

## When Strong Duality Theorems Holds for CLP

**Theorem 3** *The following statements hold for every pair of (CLP) and (CLD):*

**i)** *If (CLP) and (CLD) both are feasible, and furthermore one of them have an interior, then there is no duality gap between (CLP) and (CLD). However, one of the optimal solution may not be attainable.*

**ii)** *If (CLP) and (CLD) both are feasible and have interior, then, then both have attainable optimal solutions with no duality gap.*

**iii)** *If (CLP) or (CLD) is feasible and unbounded, then the other has no feasible solution.*

**iv)** *If (CLP) or (CLD) is infeasible, and furthermore the other is feasible and has an interior, then the other is unbounded.*

In case i), one of the optimal solution may not be attainable although no gap.

## SDP Example with Zero-Duality Gap but not Attainable

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; A_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and} \quad b_1 = 2.$$

The primal has an interior, but the dual does not.

# Proof of CLP Strong Duality Theorem

i) Let $\mathcal{F}_p$ be feasible and have an interior, and let $z^*$ be its infimum. Then we consider the alternative system pair

$$\mathcal{A}\mathbf{x} - \mathbf{b}\tau = \mathbf{0},\ \mathbf{c} \bullet \mathbf{x} - z^*\tau < 0,\ (\mathbf{x}, \tau) \in K \times R_+,$$

and

$$\mathcal{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ -\mathbf{b}^T\mathbf{y} + s = -z^*,\ (\mathbf{s}, s) \in K^* \times R_+.$$

But the former is infeasible, so that we have a solution for the latter. From the Weak Duality theorem, we must have $s = 0$, that is, we have a solution $(\mathbf{y}, \mathbf{s})$ such that

$$\mathcal{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ \mathbf{b}^T\mathbf{y} = z^*,\ \mathbf{s} \in K^*.$$

ii) We only need to prove that there exist a solution $\mathbf{x} \in \mathcal{F}_p$ such that $\mathbf{c} \bullet \mathbf{x} = z^*$, that is, the infimum of (CLP) is attainable. But this is just the other side of the proof given that $\mathcal{F}_d$ is feasible and has an interior, and $z^*$ is also the supremum of (CLD).

iii) The proof by contradiction follows the Weak Duality Theorem.

iv) Suppose $\mathcal{F}_d$ is empty and $\mathcal{F}_p$ is feasible and have an interior. Then, we have $\bar{\mathbf{x}} \in \text{int}\, K$ and $\bar{\tau} > 0$ such that $\mathcal{A}\bar{\mathbf{x}} - \mathbf{b}\bar{\tau} = \mathbf{0},\ (\bar{\mathbf{x}}, \bar{\tau}) \in \text{int}(K \times R_+).$ Then, for any $z^*$, we again consider the alternative system pair

$$\mathcal{A}\mathbf{x} - \mathbf{b}\tau = \mathbf{0},\ \mathbf{c} \bullet \mathbf{x} - z^*\tau < 0,\ (\mathbf{x}, \tau) \in K \times R_+,$$

and

$$\mathcal{A}^T\mathbf{y} + \mathbf{s} = \mathbf{c},\ -\mathbf{b}^T\mathbf{y} + s = -z^*,\ (\mathbf{s}, s) \in K^* \times R_+.$$

But the latter is infeasible, so that the formal has a feasible solution for any $z^*$. At such an solution, if $\tau > 0$, we have $\mathbf{c} \bullet (\mathbf{x}/\tau) < z^*$; if $\tau = 0$, we have $\hat{\mathbf{x}} + \alpha\mathbf{x}$, where $\hat{\mathbf{x}}$ is any feasible solution for (CLP), being feasible for (CLP) and its objective value goes to $-\infty$ as $\alpha$ goes to $\infty$.

## Optimality and Complementarity Conditions for SDP

$$\begin{array}{rcl} \mathbf{c} \bullet X - \mathbf{b}^T \mathbf{y} & = & 0 \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T \mathbf{y} - S & = & -\mathbf{c} \\ X, S & \succeq & \mathbf{0} \end{array}, \tag{1}$$

$$\begin{array}{rcl} XS & = & \mathbf{0} \\ \mathcal{A}X & = & \mathbf{b} \\ -\mathcal{A}^T \mathbf{y} - S & = & -\mathbf{c} \\ X, S & \succeq & \mathbf{0} \end{array} \tag{2}$$

## LP, SOCP, and SDP Examples

$$\begin{array}{ll} \text{min} & 2x_1 + x_2 + x_3 \\ \text{s. t.} & x_1 + x_2 + x_3 = 1, \\ & (x_1; x_2; x_3) \geq \mathbf{0}. \end{array} \qquad \begin{array}{ll} \text{max} & y \\ \text{s.t.} & \mathbf{e} \cdot y + \mathbf{s} = (2;\ 1;\ 1), \\ & (s_1; s_2; s_3) \geq \mathbf{0}. \end{array}$$

$$\begin{array}{ll} \text{min} & 2x_1 + x_2 + x_3 \\ \text{s.t.} & x_1 + x_2 + x_3 = 1, \\ & x_1 - \sqrt{x_2^2 + x_3^2} \geq 0. \end{array} \qquad \begin{array}{ll} \text{max} & y \\ \text{s.t.} & \mathbf{e} \cdot y + \mathbf{s} = (2;\ 1;\ 1), \\ & s_1 - \sqrt{s_2^2 + s_3^2} \geq 0. \end{array}$$

For the SOCP case: $2 - y \geq \sqrt{2(1-y)^2}$. Since $y = 1$ is feasible for the dual, $y^* \geq 1$ so that the dual constraint becomes $2 - y \geq \sqrt{2}(y-1)$ or $y \leq \sqrt{2}$. Thus, $y^* = \sqrt{2}$, and there is no duality gap.

$$\begin{array}{ll} \text{minimize} & \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \\ \text{subject to} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} \bullet \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} = 1, \\ & \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq \mathbf{0}, \end{array}$$

$$\begin{array}{ll} \text{maximize} & y \\ \text{subject to} & \begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix} y + \mathbf{s} = \begin{pmatrix} 2 & .5 \\ .5 & 1 \end{pmatrix}, \\ & \mathbf{s} = \begin{pmatrix} s_1 & s_2 \\ s_2 & s_3 \end{pmatrix} \succeq \mathbf{0}. \end{array}$$

## Distributionally Robust Learning (DRL)

The simple sample average minimization

$$\text{mimimize}_{\mathbf{x}\in X} \quad \sum_{k=1}^N \hat{p}_k h(\mathbf{x},\xi_k) \tag{3}$$

where $\xi^k$ represents the $k$th sample data and $\hat{p}_k$ is its sample/empirical probability.

Suppose we like to "robustfy" the problem by considering

$$\begin{array}{lcl} \text{mimimize}_{\mathbf{x}\in X} & & \left[\max_{\mathbf{d}\in D} \sum_{k=1}^N (\hat{p}_k + d_k) h(\mathbf{x},\xi_k)\right] \\ & = & \sum_{k=1}^N \hat{p}_k h(\mathbf{x},\xi_k) + \left[\max_{\mathbf{d}\in D} \sum_{k=1}^N d_k h(\mathbf{x},\xi_k)\right] \end{array} \tag{4}$$

where $D$ is given by $D = \{\mathbf{d} :\ \sum_{k=1}^N d_k = 0,\ \|\mathbf{d}\|_2^2 \le 1/N\}$ which has a second-order cone representation

$$D = \{(d_0;\mathbf{d}) :\ d_0 = 1/\sqrt{N},\ \sum_{k=1}^N d_k = 0,\ \|\mathbf{d}\|_2 \le d_0\}.$$

## The Inner SOCP Porblem

$$\begin{array}{ll} \max_{(d_0;\mathbf{d})\in D} & \sum_{k=1}^{N} d_k h(\mathbf{x},\xi_k) \\ \text{s,t,} & d_0 = \frac{1}{\sqrt{N}}, \\ & \sum_{k=1}^{N} d_k = 0, \\ & (d_0;\mathbf{d}) \in SOC^{N+1}; \end{array}$$

where its dual is:

$$\begin{array}{ll} \min_{\{\lambda_0,\lambda_1\}} & \frac{1}{\sqrt{N}}\lambda_0 \\ \text{s,t,} & \lambda_0(1;\mathbf{0}) + \lambda_1(0;\mathbf{e}) - (s_0;\mathbf{s}) = (0;\mathbf{h}(\mathbf{x},\xi)), \\ & (s_0;\mathbf{s}) \in SOC^{N+1}; \end{array}$$

which can simplifies to

$$\begin{array}{ll} \min_{\{\lambda_0,\lambda_1\}} & \frac{1}{\sqrt{N}}\lambda_0 \\ \text{s,t,} & (\lambda_0;\lambda_1\mathbf{e} - \mathbf{h}(\mathbf{x},\xi)) \in SOC^{N+1}; \end{array}$$

or further simplified to

$$\min_{\lambda_1} \quad \frac{1}{\sqrt{N}}\|\lambda_1\mathbf{e} - \mathbf{h}(\mathbf{x},\xi)\| = \frac{1}{\sqrt{N}}\sqrt{\sum_{k=1}^{N}(\lambda_1 - h(\mathbf{x},\xi_k))^2}.$$

## Reformulation of the DRL Problem

$$\text{mimimize}_{\mathbf{x}\in X} \quad \sum_{k=1}^{N} \hat{p}_k h(\mathbf{x},\xi_k) + \quad \frac{1}{\sqrt{N}}\left[\min_{\lambda_1} \sqrt{\sum_{k=1}^{N}(\lambda_1 - h(\mathbf{x},\xi_k))^2}\right]$$

Or

$$\text{mimimize}_{\{\mathbf{x}\in X,\lambda_1\}} \quad \sum_{k=1}^{N} \hat{p}_k h(\mathbf{x},\xi_k) + \frac{1}{\sqrt{N}}\sqrt{\sum_{k=1}^{N}(\lambda_1 - h(\mathbf{x},\xi_k))^2}$$

One should have

$$\lambda_1 = \frac{1}{N}\sum_{k=1}^{N} h(\mathbf{x},\xi_k)$$

the mean value of $h(\mathbf{x},\xi_k),\ k = 1,...,N$.

Thus, the final DRL problem becomes

$$\text{mimimize}_{\mathbf{x}\in X} \quad \sum_{k=1}^{N} \hat{p}_k h(\mathbf{x},\xi_k) + \frac{1}{\sqrt{N}}\sqrt{\sum_{k=1}^{N}\left(\frac{1}{N}\sum_{k=1}^{N} h(\mathbf{x},\xi_k) - h(\mathbf{x},\xi_k)\right)^2}$$

This is the original sample average objective plus the standard deviation of the samples.

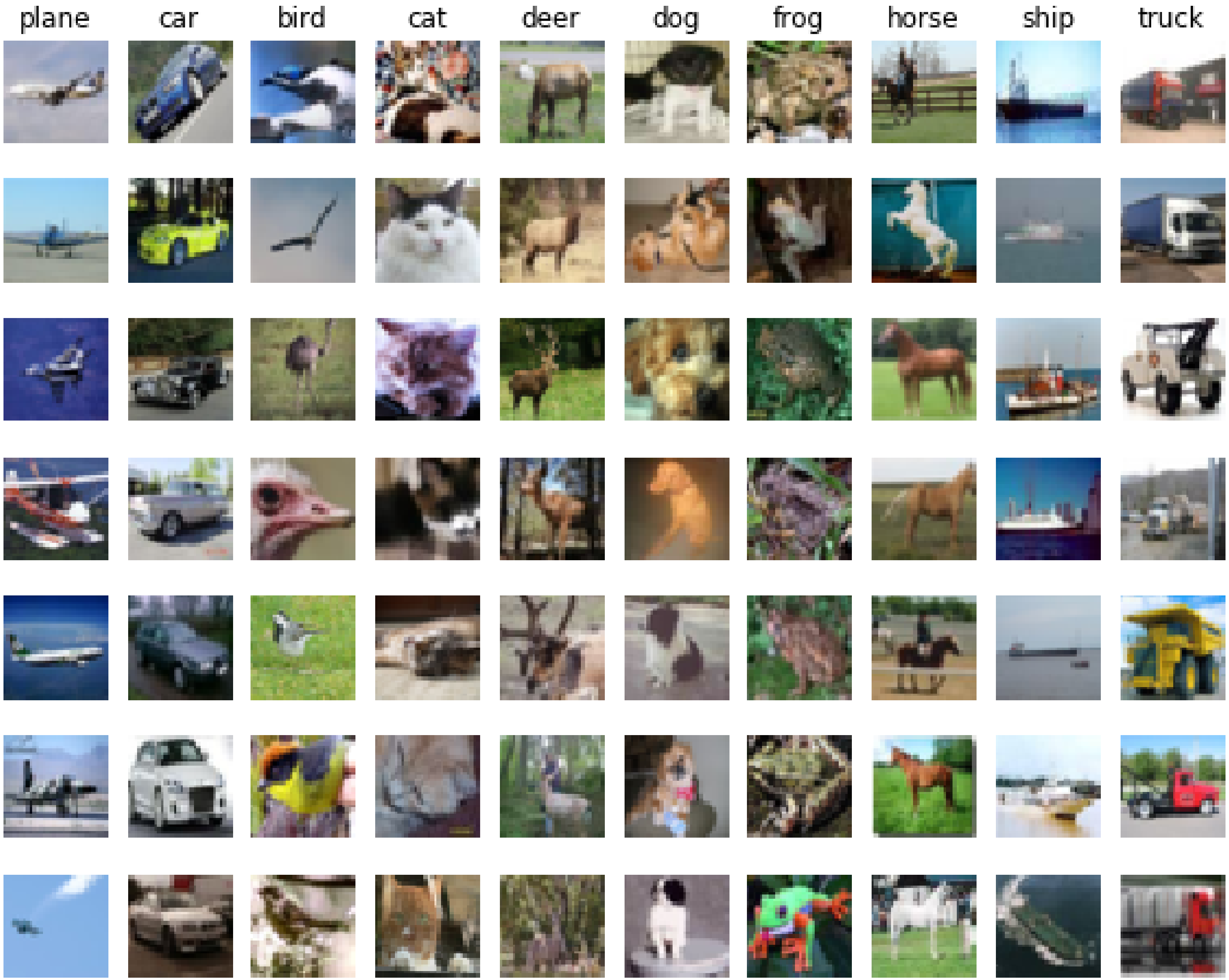


Figure 1: Result of the DRO Learning I: Original

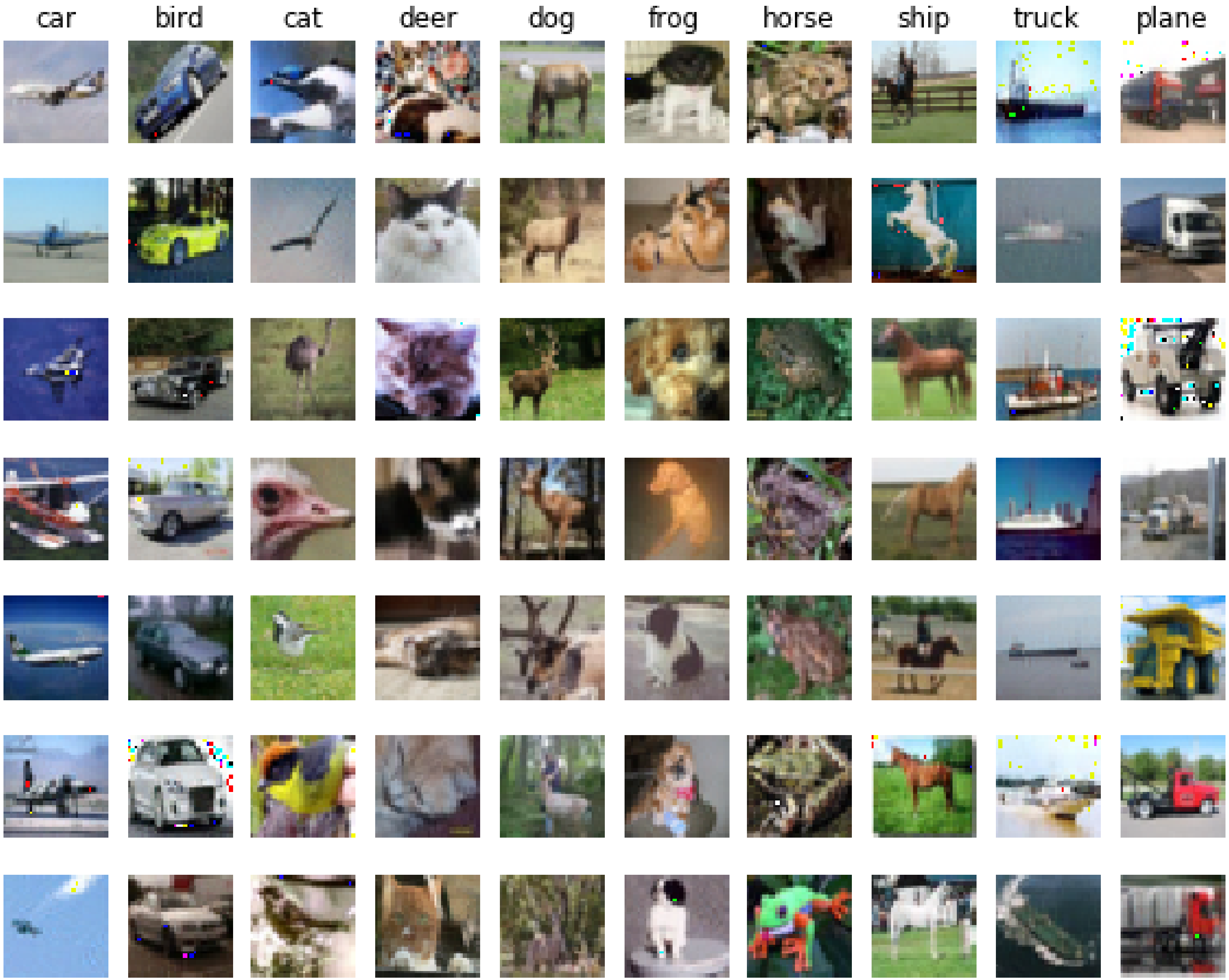


Figure 2: Result of the DRO Learning II: Nonrobust Result

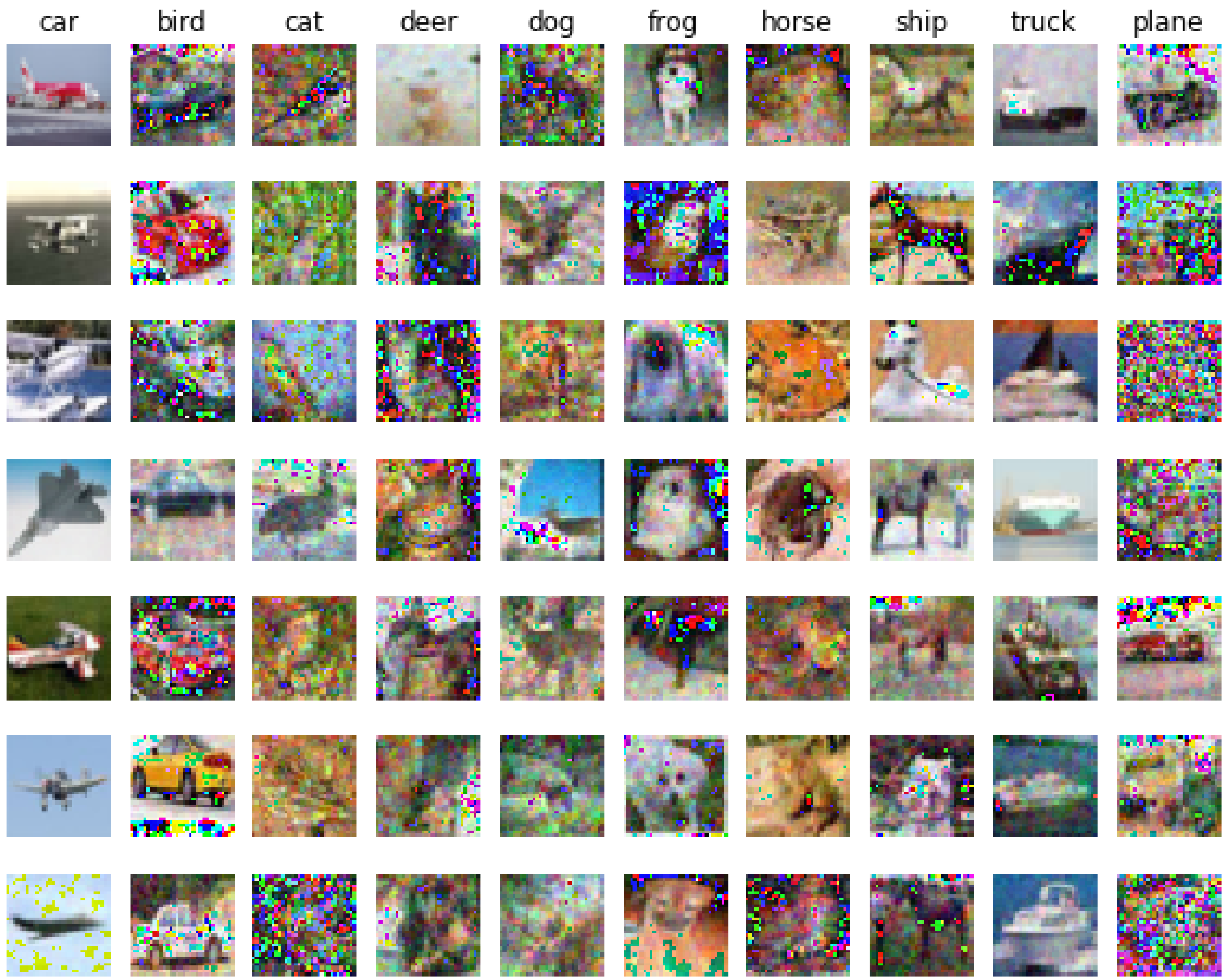


Figure 3: Result of the DRO Learning III: DRO Result