Optimization Methods and Software
Vol. 00, No. 00, January 2009, 1–16

RESEARCH ARTICLE

Robust Regression Analysis for Gene-Environment and Eco-Finance Networks under Polyhedral and Ellipsoidal Uncertainty

Erik Kropat[a*] and Gerhard-Wilhelm Weber[b]

[a] *Universität der Bundeswehr München, Department of Informatics, Institute for Theoretical Computer Science, Mathematics and Operations Research Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany*

[b] *Middle East Technical University, Institute of Applied Mathematics, Departments of Financial Mathematics, Scientific Computing and Actuarial Sciences, 06531 Ankara, Turkey*

(*v4.4 released October 2008*)

Target-environment networks provide a conceptual framework for the analysis of complex regulatory systems that are effected by noise and uncertainty. They occur in many disciplines and are often referred to as gene-environment networks in computational biology and eco-finance networks in financial sciences. Clustering and classification can be applied for an identification of functionally related groups of targets and environmental factors. A parameterized linear model is introduced that determines the intricate interactions and synergetic connections between target and environmental clusters. For an estimation of parameters, a crisp regression regression problem is considered. In case of uncertain system states, the clusters are represented in terms of polyhedrons and ellipsoids and we derive the corresponding set-theoretic robust counterpart programs.

Keywords: gene-environment networks, eco-finance networks, robust regression, uncertainty modelling, computational biology, clustering, classification

AMS Subject Classification: 90C30, 92C42, 91B76

1. Introduction

The modeling and prediction of *regulatory networks* is of considerable importance in many disciplines such as finance, biology, medicine and life sciences. The identification of the underlying network topology allows to gain deeper insights in the regulating effects and the hidden relationships between the variables under consideration. This is even more promising as the technical developments of the last decades have produced a huge amount of data that is still waiting for a deeper analysis. Although many theoretical contributions from various disciplines have focussed on the analysis of such systems, the identification of regulatory networks from real-world data is still challenging mathematics. In particular, the presence of noise and data uncertainty raises serious problems to be dealt with on both the

*Corresponding author. Email: erik.kropat@unibw.de

†Honorary positions: University of Aveiro, Portugal; University of Siegen, Germany; Universiti Teknologi Malaysia, Skudai, Malaysia.

ISSN: 1055-6788 print/ISSN 1029-4937 online

DOI: 10.1080/1055678xxxxxxxxxxxxx
http://www.informaworld.com

theoretical and computational side. Beside this, the regulatory system has often to be further extended and improved with regard to the unknown effects of additional parameters and factors which may exert a disturbing influence on the key variables under consideration. All these dynamical networks are affected

- by uncertainty in the data, both in their input and their output parts, or, in other words,
- by uncertainty in the scenarios and by random fluctuation,
- by the necessity to reduce the model complexity, i.e., to regularize, rarefy and stabilize.

In this regard, we are on the way between complete determinism in processes and the rich randomness as it can be investigated by stochastic calculus and, especially, Lévy processes. In 2002, we started our modelling of processes related with *genetic networks* in the deterministic case where, then, in the following years, we included the role of the environment which yielded us our *gene-environment* and *eco-finance networks*. Since in these kinds of dynamics, the impact of the environmental items became implied as additive "shift" terms which can also be called as perturbations, we arrived at our first implication of noise. Having once entered the domain of uncertainty, we went on working it out, firstly, by *interval uncertainty* where, however, the dependencies and correlations between the various items from biology, medicine, these sectors of ecology, education and finance were not taken into account yet [49, 52]. We treated those modelling tasks by the help of *Chebychev approximation* and *Generalized Semi-Infinite Optimization*. By turning to the case of *ellipsoidal uncertainty* [20–22] and, as far as splines were used for approximation, by applying *Multivariate Adaptive Regression Spline* instead of *Generalized Additive Models* [36–38], we could overcome that drawback and we included stochastic dependencies and interactions into our model. Here, the dimensions of the ellipsoids are motivated by additional information related to the model items and their similarities, i.e., on how much they are close to each other and how the distribution of such clusters expresses itself geometrically in ellipsoidal forms. In game theoretical contexts, we called these clusters *(sub) coalitions*. In this way, we arrived at a family or, in particular, sequence of ellipsoids which can be regarded as the bodies which contain our target or environmental variables at the corresponding times, i.e., the processes studied, by confidence levels of, e.g., 95%.

As it is clearly understood today, environmental factors constitute an essential group of regulating components and by including these additional variables the models performance can be significantly improved. The advantage of such an refinement has been demonstrated for example in [24], where it is shown that prediction and classification performances of supervised learning methods for the most complex genome-wide human disease classification can be greatly improved by considering environmental aspects. Many other examples from biology and life sciences refer to regulatory systems where environmental effects are strongly involved. Among them are, e.g., *metabolic networks* [2, 28, 49], *immunological networks* [14], *social-* and *ecological networks* [13]. We refer to [12, 16, 33–35, 55, 56] for applications, practical examples and numerical calculations.

Whenever we want to particularly address items to the financial sector among the target variables or the environmental variables which, in fact, maybe be regarded in a dual relationship mutually, then we arrive at *eco-finance networks* [19]. This interpretation and variety of our studies also represents that the identification of dynamics related with the Kyoto Protocol, where financial expenditures and emissions reduced interact in time (TEM model) [16, 18, 23, 29–32]. Financial negotiation processes, represented in the way of collaborative game theory [44, 45], and the identification and dynamics of financial processes given by stochastic differential equations and their time-discretized versions [50, 51], are an important part our research.

Recent studies on target-environment and gene-environment networks focussed on systems with functionally related groups of target and environmental factors. These groups are identified in a preprocessing step of clustering and classification and the corresponding uncertain multivariate states are represented by ellipsoids [20–22]. The interaction of clusters is determined by affine-linear equations based on ellipsoidal calculus. Various regression problems are introduced for an identification of unknown system parameters from (ellipsoidal) measurement data. In addition, problems of network rarefication and the corresponding mixed-integer regression problems as well as a further relaxation by means of continuous optimization have been addressed in [22]. For further details on the underlying set-theoretic regression theory and the solvability by semi-definite programming we refer to [20–22].

In this paper, we further extend this approach and offer a new perspective where clusters of targets and groups of environmental factors take influence on the states and values of *single targets* and *single environmental variables*.

The comparison of measurements and predictions of the model leads to a regression model for parameter estimation. Since the clusters can be affected by noise and errors, the uncertain multivariate states are represented by polyhedrons and ellipsoids what refers to the concept of *robustness* for mathematical programming problems.

The paper is organized as follows: In Section 2 some basic facts and notation about target and environmental variables as well as the partitioning of data in clusters are provided. Then, in Section 3, a time-discrete linear model is introduced that relates the single variables and the multivariate states of groups of target and environmental factors. The corresponding regression model for parameter estimation is addressed in Section 4. In a further step, data uncertainty becomes included into our modelling in Section 5 and Section 6, where the multivariate states of clusters are represented in terms of polyhedrons and ellipsoids. Hereby, the corresponding regression models can be reformulated in terms of robust counterpart programs.

2. Target-Environment Networks

In this study, we consider time-discrete target-environment regulatory systems with n targets and m environmental factors. The vector $\mathbb{X} = \big[\mathbb{X}_1, \ldots, \mathbb{X}_n\big]'$ denotes the expression values of the targets and $\mathbb{E} = \big[\mathbb{E}_1, \ldots, \mathbb{E}_m\big]'$ represents the environmental variables, where $[\,\cdot\,]'$ stands for the transposition of a matrix or vector. Often, functionally related groups of targets and environmental items can be identified in a preprocessing step of clustering and classification. These groups exert a more or less regulating influence on single targets or environmental factors. For this reason, the set of targets is divided in R disjoint clusters $C_r \subset \{1, \ldots, n\},\ r = 1, \ldots, R$. Similarly, the set of all environmental items is divided in S disjoint clusters $D_s \subset \{1, \ldots, m\},\ s = 1, \ldots, S$. Here, we focus on non-overlapping clusters and assume a strict sub-division of the variables, so that the relations $C_{r_1} \cap C_{r_2} = \emptyset$ for all $r_1 \neq r_2$ and $D_{s_1} \cap D_{s_2} = \emptyset$ for all $s_1 \neq s_2$ are fulfilled. According to the cluster structure, we introduce the sub-vector $X_r \in \mathbb{R}^{|C_r|}$ of $\mathbb{X}$ as the restriction of $\mathbb{X}$ given by elements of C_r. In the same way, the sub-vector $E_s \in \mathbb{R}^{|D_s|}$ is defined as the restriction of $\mathbb{E}$ given by elements of D_s.

3. The Linear Model

In this section, we introduce a time-discrete model for the states of the targets $\mathbb{X}_j$, $j = 1, \ldots, n$, and environmental factors $\mathbb{E}_i$, $i = 1, \ldots, m$. Four types of interactions and regulating effects are involved: (TT) target $\rightarrow$ target, (ET) environment $\rightarrow$ target, (TE) target $\rightarrow$ environment, (EE) environment $\rightarrow$ environment. A general linear model is given by

$$\begin{aligned} \mathbb{X}_j^{(\kappa+1)} &= \zeta_{j0}^T + \big[\mathbb{X}^{(\kappa)}\big]'\Theta_j^{TT} + \big[\mathbb{E}^{(\kappa)}\big]'\Theta_j^{ET}, \\ \mathbb{E}_i^{(\kappa+1)} &= \zeta_{i0}^E + \big[\mathbb{X}^{(\kappa)}\big]'\Theta_i^{TE} + \big[\mathbb{E}^{(\kappa)}\big]'\Theta_i^{EE}, \end{aligned}$$

with $\kappa \geq 0$, where $\Theta_j^{TT}, \Theta_i^{TE} \in \mathbb{R}^n$, $\Theta_j^{ET}, \Theta_i^{EE} \in \mathbb{R}^m$ denote the vectors of parameters, $\zeta_{j0}^T, \zeta_{i0}^E \in \mathbb{R}$ are intercepts. This model depends on $(n+m)(n+m+1)$ unknown parameters. The initial values $\mathbb{X}^{(0)}$ and $\mathbb{E}^{(0)}$ can be given by the first measurements of targets and environmental factors, i.e., $\mathbb{X}^{(0)} := \overline{\mathbb{X}}^{(0)}$ and $\mathbb{E}^{(0)} := \overline{\mathbb{E}}^{(0)}$.

Referring to the classification in terms of clusters, the model under consideration can be reformulated as follows:

$$\left.\begin{aligned} \mathbb{X}_j^{(\kappa+1)} &= \zeta_{j0}^T + \sum_{r=1}^{R}\big[X_r^{(\kappa)}\big]'\Theta_{jr}^{TT} + \sum_{s=1}^{S}\big[E_s^{(\kappa)}\big]'\Theta_{js}^{ET}, \\ \mathbb{E}_i^{(\kappa+1)} &= \zeta_{i0}^E + \sum_{r=1}^{R}\big[X_r^{(\kappa)}\big]'\Theta_{ir}^{TE} + \sum_{s=1}^{S}\big[E_s^{(\kappa)}\big]'\Theta_{is}^{EE} \end{aligned}\right\} \qquad \text{(LCM)}$$

with $\kappa \geq 0$, where (LCM) stands for *linear cluster model.* The vectors Θ_{jr}^{TT} and Θ_{ir}^{TE} are $|C_r|$-subvectors of the parameter vectors Θ_j^{TT} and Θ_i^{TE}, respectively. These subvectors are given by the indices of cluster C_r. Similarly, the vectors Θ_{js}^{ET} and Θ_{is}^{EE} are $|D_s|$-subvectors of the parameter vectors Θ_j^{ET} and Θ_i^{EE}. Since all clusters are disjoint, the aforementioned subvectors correspond to distinct parts of the parameter vectors.

4. The Regression Problem

We now turn to an estimation of parameters of the linear cluster model (LCM). For a regression analysis, the predictions of (LCM) have to be compared with the states of targets $\overline{\mathbb{X}}^{(\kappa)} = \big[\overline{\mathbb{X}}_1^{(\kappa)}, \ldots, \overline{\mathbb{X}}_n^{(\kappa)}\big]' \in \mathbb{R}^n$ and environmental observations $\overline{\mathbb{E}}^{(\kappa)} = \big[\overline{\mathbb{E}}_1^{(\kappa)}, \ldots, \overline{\mathbb{E}}_m^{(\kappa)}\big]' \in \mathbb{R}^m$, $\kappa = 0, 1, \ldots, T$, which are obtained from measurements taken at sampling times $t_0 < t_1 < \ldots < t_T$. By inserting these measurements in model (LCM) we obtain the following predictions:

$$\begin{aligned} \widehat{\mathbb{X}}_j^{(\kappa+1)} &= \zeta_{j0}^T + \sum_{r=1}^{R}\big[\overline{X}_r^{(\kappa)}\big]'\Theta_{jr}^{TT} + \sum_{s=1}^{S}\big[\overline{E}_s^{(\kappa)}\big]'\Theta_{js}^{ET}, \\ \widehat{\mathbb{E}}_i^{(\kappa+1)} &= \zeta_{i0}^E + \sum_{r=1}^{R}\big[\overline{X}_r^{(\kappa)}\big]'\Theta_{ir}^{TE} + \sum_{s=1}^{S}\big[\overline{E}_s^{(\kappa)}\big]'\Theta_{is}^{EE}, \end{aligned}$$

where $\kappa = 0, 1, \ldots, T-1$. We set $\widehat{\mathbb{X}}_j^{(0)} := \overline{\mathbb{X}}_j^{(0)}$ and $\widehat{\mathbb{E}}_i^{(0)} := \overline{\mathbb{E}}_i^{(0)}$ as well as $\widehat{\mathbb{X}}^{(\kappa)} = \big[\widehat{\mathbb{X}}_1^{(\kappa)}, \ldots, \widehat{\mathbb{X}}_n^{(\kappa)}\big]'$ and $\widehat{\mathbb{E}}^{(\kappa)} = \big[\widehat{\mathbb{E}}_1^{(\kappa)}, \ldots, \widehat{\mathbb{E}}_m^{(\kappa)}\big]'$, where $\kappa = 0, 1, \ldots, T$; $i = 1, \ldots, n$; $j = 1, \ldots, m$.

The comparison of measurements and predictions leads to the following regression problem:

$$\text{Minimize} \quad \sum_{\kappa=1}^{T} \Big\{ \Big\| \widehat{\mathbb{X}}^{(\kappa)} - \overline{\mathbb{X}}^{(\kappa)} \Big\|_* + \Big\| \widehat{\mathbb{E}}^{(\kappa)} - \overline{\mathbb{E}}^{(\kappa)} \Big\|_* \Big\}.$$

The norm $\|\cdot\|_*$ can be for example the 1-, 2-, ∞-norm or, in general, also the $\|\cdot\|_p$-norm. We note that we have analyzed the regression problem for all these norms with both polyhedral and ellipsoidal data and we will address these approaches in forthcoming papers.

In this paper, we will restrict ourselves to the 1-norm and we consider the following regression problem:

$$\text{Minimize} \quad \sum_{\kappa=1}^{T} \Big\{ \sum_{j=1}^{n} \Big| \widehat{\mathbb{X}}_j^{(\kappa)} - \overline{\mathbb{X}}_j^{(\kappa)} \Big| + \sum_{i=1}^{m} \Big| \widehat{\mathbb{E}}_i^{(\kappa)} - \overline{\mathbb{E}}_i^{(\kappa)} \Big| \Big\}. \tag{RP}$$

In the following sections, data uncertainty will become included. For this reason it is convenient to reformulate the regression problem (RP) as follows:

$$\begin{aligned}
&\text{Minimize} \quad \sum_{\kappa=1}^{T} \Big\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \Big\} \\
&\text{such that} \quad \Big| \widehat{\mathbb{X}}_j^{(\kappa)} - \overline{\mathbb{X}}_j^{(\kappa)} \Big| \le p_j^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ j = 1, \ldots, n), \\
&\qquad\qquad\quad \Big| \widehat{\mathbb{E}}_i^{(\kappa)} - \overline{\mathbb{E}}_i^{(\kappa)} \Big| \le q_i^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ i = 1, \ldots, m).
\end{aligned}$$

This problem can be equivalently written as

$$\begin{aligned}
&\text{Minimize} \quad \sum_{\kappa=1}^{T} \Big\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \Big\} \\
&\text{such that} \quad \Big| \zeta_{j0}^{T} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} - \overline{\mathbb{X}}_j^{(\kappa)} \Big| \le p_j^{(\kappa)}, \\
&\qquad\qquad\quad \Big| \zeta_{i0}^{E} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} - \overline{\mathbb{E}}_i^{(\kappa)} \Big| \le q_i^{(\kappa)} \\
&\qquad\qquad\quad (\kappa = 1, \ldots, T;\ j = 1, \ldots, n;\ i = 1, \ldots, m).
\end{aligned}$$

5. Robust Regression with Polyhedral Uncertainty

Measurements and observations of targets and environmental factors are usually effected by uncertainty. The regression problem (RP) depends on crisp (numerical) measurements and does not reflect the disturbing influence of unprecise data. For this reason, we now turn to *robust regression models* with regard to data sets with cluster partition. There are several ways to describe data uncertainty from a set-theoretic perspective. When an individual error can be assigned to each target and environmental factor, the corresponding states of variables are given by *intervals*, whereas the states of clusters are represented by *hyperrectangles*. When errors of clusters elements are correlated, non-paraxial sets have to be considered and the polyhedral uncertainty sets can be replaced by error ellipsoids. In this section, we consider the case of paraxial error sets and refer to polyhedral uncertainty. Correlation between cluster elements and ellipsoidal uncertainty will be addressed in Section 6.

In order to represent data uncertainty in terms of error intervals, we now assume that there is - a possibly time-dependent - maximal error for each target and environmental factor that is denoted by $(err^T)_j^{(\kappa)} \geq 0$ and $(err^E)_i^{(\kappa)} \geq 0$, respectively. The maximal errors are the elements of the vectors

$$
\begin{aligned}
(Err^T)^{(\kappa)} &= \Big[(err^T)_1^{(\kappa)}, \ldots, (err^T)_n^{(\kappa)}\Big]', \\
(Err^E)^{(\kappa)} &= \Big[(err^E)_1^{(\kappa)}, \ldots, (err^E)_m^{(\kappa)}\Big]',
\end{aligned}
$$

with $\kappa = 0, 1, \ldots, T$. Since an individual error is assigned to each single variable, we assume that the states of clusters of targets and environmental factors are subject to polyhedral uncertainty. The states of clusters are given in terms of hyperrectangles where the corresponding measurement value defines its center:

$$
\begin{aligned}
X_r^{(\kappa)} &\in \mathcal{Q}\Big(\overline{X}_r^{(\kappa)}, (Err^T)_r^{(\kappa)}\Big) \subset \mathbb{R}^{|C_r|}, \\
E_s^{(\kappa)} &\in \mathcal{Q}\Big(\overline{E}_s^{(\kappa)}, (Err^E)_s^{(\kappa)}\Big) \subset \mathbb{R}^{|D_s|},
\end{aligned}
$$

with

$$
\begin{aligned}
\mathcal{Q}\Big(\overline{X}_r^{(\kappa)}, (Err^T)_r^{(\kappa)}\Big) &= \Big\{\overline{X}_r^{(\kappa)} + \mathrm{diag}\big((Err^T)_r^{(\kappa)}\big)u_r \,\Big|\, u_r \in \mathbb{R}^{|C_r|},\ \|u_r\|_\infty \leq 1\Big\}, \\
\mathcal{Q}\Big(\overline{E}_s^{(\kappa)}, (Err^E)_s^{(\kappa)}\Big) &= \Big\{\overline{E}_s^{(\kappa)} + \mathrm{diag}\big((Err^E)_s^{(\kappa)}\big)v_s \,\Big|\, v_s \in \mathbb{R}^{|D_s|},\ \|v_s\|_\infty \leq 1\Big\},
\end{aligned}
$$

where $\kappa = 0, 1, \ldots, T$. The measurements $\overline{X}_r^{(\kappa)}$ and $\overline{E}_s^{(\kappa)}$ determine the centers of the polyhedrons and the corresponding diagonal matrices $\mathrm{diag}\big((Err^T)_r^{(\kappa)}\big)$ and $\mathrm{diag}\big((Err^E)_s^{(\kappa)}\big)$ are given by the errors of cluster elements.

In the following, we assume that the constraints are satisfied for all realizations of the states $X_r^{(\kappa)} \in \mathcal{Q}\Big(\overline{X}_r^{(\kappa)}, (Err^T)_r^{(\kappa)}\Big)$ and $E_s^{(\kappa)} \in \mathcal{Q}\Big(\overline{E}_s^{(\kappa)}, (Err^E)_s^{(\kappa)}\Big)$. This leads us to a robust regression problem with polyhedral uncertainty:

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{\kappa=1}^{T} \Big\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \Big\} \\
\text{such that} \quad & \Big| \zeta_{j0}^T + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} - \mathbb{X}_j^{(\kappa)} \Big| \le p_j^{(\kappa)}, \\
& \Big| \zeta_{i0}^E + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} - \mathbb{E}_i^{(\kappa)} \Big| \le q_i^{(\kappa)} \\
& (\kappa = 1, \ldots, T;\ j = 1, \ldots, n;\ i = 1, \ldots, m) \\
& \forall X_r^{(\kappa)} \in \mathcal{Q}\Big(\overline{X}_r^{(\kappa)}, (Err^T)_r^{(\kappa)}\Big) \ \ (\kappa = 0, \ldots, T-1;\ r = 1, \ldots, R), \\
& \forall E_s^{(\kappa)} \in \mathcal{Q}\Big(\overline{E}_s^{(\kappa)}, (Err^E)_s^{(\kappa)}\Big) \ \ (\kappa = 0, \ldots, T-1;\ s = 1, \ldots, S).
\end{aligned}
$$

The next problem results from a reformulation of the absolute values:

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{\kappa=1}^{T} \Big\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \Big\} \\
\text{such that} \quad & \zeta_{j0}^T + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} - \mathbb{X}_j^{(\kappa)} \le p_j^{(\kappa)}, \\
& -\zeta_{j0}^T - \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} - \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} + \mathbb{X}_j^{(\kappa)} \le p_j^{(\kappa)}, \\
& \zeta_{i0}^E + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} - \mathbb{E}_i^{(\kappa)} \le q_i^{(\kappa)}, \\
& -\zeta_{i0}^E - \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} - \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} + \mathbb{E}_i^{(\kappa)} \le q_i^{(\kappa)} \\
& (\kappa = 1, \ldots, T;\ j = 1, \ldots, n;\ i = 1, \ldots, m) \\
& \forall X_r^{(\kappa)} \in \mathcal{Q}\Big(\overline{X}_r^{(\kappa)}, (Err^T)_r^{(\kappa)}\Big) \ \ (\kappa = 0, \ldots, T-1;\ r = 1, \ldots, R), \\
& \forall E_s^{(\kappa)} \in \mathcal{Q}\Big(\overline{E}_s^{(\kappa)}, (Err^E)_s^{(\kappa)}\Big) \ \ (\kappa = 0, \ldots, T-1;\ s = 1, \ldots, S).
\end{aligned}
$$

By considering the sets

$$
\begin{aligned}
U_r &:= \Big\{ u_r \in \mathbb{R}^{|C_r|} \ \Big|\ \|u_r\|_\infty \le 1 \Big\},\ r = 1, \ldots, R, \\
V_s &:= \Big\{ v_s \in \mathbb{R}^{|D_s|} \ \Big|\ \|v_s\|_\infty \le 1 \Big\},\ s = 1, \ldots, S
\end{aligned}
$$

another equivalent problem can be obtained:

$$
\begin{aligned}
\text{Minimize } & \sum_{\kappa=1}^{T} \left\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \right\} \\
\text{such that } & \zeta_{j0}^{T} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} + \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \operatorname{diag}\big((Err^T)_r^{(\kappa-1)}\big) \Theta_{jr}^{TT} \big\} \\
& + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} + \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \operatorname{diag}\big((Err^E)_s^{(\kappa-1)}\big) \Theta_{js}^{ET} \big\} \\
& - \overline{\mathbb{X}}_j^{(\kappa)} \le p_j^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ j = 1, \ldots, n),
\end{aligned}
$$

$$
\begin{aligned}
& -\zeta_{j0}^{T} - \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} - \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \operatorname{diag}\big((Err^T)_r^{(\kappa-1)}\big) \Theta_{jr}^{TT} \big\} \\
& - \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} - \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \operatorname{diag}\big((Err^E)_s^{(\kappa-1)}\big) \Theta_{js}^{ET} \big\} \\
& + \overline{\mathbb{X}}_j^{(\kappa)} \le p_j^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ j = 1, \ldots, n),
\end{aligned}
$$

$$
\begin{aligned}
& \zeta_{i0}^{E} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} + \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \operatorname{diag}\big((Err^T)_r^{(\kappa-1)}\big) \Theta_{ir}^{TE} \big\} \\
& + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} + \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \operatorname{diag}\big((Err^E)_s^{(\kappa-1)}\big) \Theta_{is}^{ET} \big\} \\
& - \overline{\mathbb{E}}_i^{(\kappa)} \le q_i^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ i = 1, \ldots, m),
\end{aligned}
$$

$$
\begin{aligned}
& -\zeta_{i0}^{E} - \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} - \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \operatorname{diag}\big((Err^T)_r^{(\kappa-1)}\big) \Theta_{ir}^{TE} \big\} \\
& - \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} - \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \operatorname{diag}\big((Err^E)_s^{(\kappa-1)}\big) \Theta_{is}^{ET} \big\} \\
& + \overline{\mathbb{E}}_i^{(\kappa)} \le q_i^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ i = 1, \ldots, m).
\end{aligned}
$$

With the equations

$$
\begin{aligned}
\max_{u_r \in U_r} \big\{ u_r' \,\overline{\Sigma}_r^{(\kappa)} \,\Theta_{jr}^{TT} \big\} &= \max_{u_r \in U_r} \big\{ -u_r' \,\overline{\Sigma}_r^{(\kappa)} \,\Theta_{jr}^{TT} \big\} = \Big\| \overline{\Sigma}_r^{(\kappa)} \Theta_{jr}^{TT} \Big\|_\infty, \\
\max_{u_r \in U_r} \big\{ u_r' \,\overline{\Sigma}_r^{(\kappa)} \,\Theta_{ir}^{TE} \big\} &= \max_{u_r \in U_r} \big\{ -u_r' \,\overline{\Sigma}_r^{(\kappa)} \,\Theta_{ir}^{TE} \big\} = \Big\| \overline{\Sigma}_r^{(\kappa)} \Theta_{ir}^{TE} \Big\|_\infty, \\
\max_{v_s \in V_s} \big\{ v_s' \,\overline{\Pi}_s^{(\kappa)} \,\Theta_{is}^{ET} \big\} &= \max_{v_s \in V_s} \big\{ -v_s' \,\overline{\Pi}_s^{(\kappa)} \,\Theta_{is}^{ET} \big\} = \Big\| \overline{\Pi}_s^{(\kappa)} \Theta_{is}^{ET} \Big\|_\infty, \\
\max_{v_s \in V_s} \big\{ v_s' \,\overline{\Pi}_s^{(\kappa)} \,\Theta_{is}^{EE} \big\} &= \max_{v_s \in V_s} \big\{ -v_s' \,\overline{\Pi}_s^{(\kappa)} \,\Theta_{is}^{EE} \big\} = \Big\| \overline{\Pi}_s^{(\kappa)} \Theta_{is}^{EE} \Big\|_\infty
\end{aligned}
$$

the regression problem can be represented as follows:

$$
\begin{aligned}
\text{Minimize}\quad & \sum_{\kappa=1}^{T}\left\{\sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)}\right\}\\
\text{such that}\quad & \Bigg|\zeta_{j0}^{T} + \sum_{r=1}^{R}\big[\,\overline{X}_r^{(\kappa-1)}\big]'\Theta_{jr}^{TT}\\
&\quad + \sum_{s=1}^{S}\big[\,\overline{E}_s^{(\kappa-1)}\big]'\Theta_{js}^{ET} \;-\; \overline{\mathbb{X}}_j^{(\kappa)}\Bigg|\\
&\quad + \sum_{r=1}^{R}\big\|\mathrm{diag}\big((Err^{T})_r^{(\kappa-1)}\big)\Theta_{jr}^{ET}\big\|_\infty\\
&\quad + \sum_{s=1}^{S}\big\|\mathrm{diag}\big((Err^{E})_s^{(\kappa-1)}\big)\Theta_{js}^{ET}\big\|_\infty \;\le p_j^{(\kappa)},\\
& \Bigg|\zeta_{i0}^{E} + \sum_{r=1}^{R}\big[\,\overline{X}_r^{(\kappa-1)}\big]'\Theta_{ir}^{TE}\\
&\quad + \sum_{s=1}^{S}\big[\,\overline{E}_s^{(\kappa-1)}\big]'\Theta_{is}^{EE} \;-\; \overline{\mathbb{E}}_i^{(\kappa)}\Bigg|\\
&\quad + \sum_{r=1}^{R}\big\|\mathrm{diag}\big((Err^{T})_r^{(\kappa-1)}\big)\Theta_{ir}^{TE}\big\|_\infty\\
&\quad + \sum_{s=1}^{S}\big\|\mathrm{diag}\big((Err^{E})_s^{(\kappa-1)}\big)\Theta_{is}^{EE}\big\|_\infty \;\le q_i^{(\kappa)}\\
& (\kappa = 1,\dots,T;\; j = 1,\dots,n;\; i = 1,\dots,m).
\end{aligned}
$$

With

$$
\begin{aligned}
\Theta_j^{G} &= \Big[\zeta_{j0}^{T}, \Theta_{j1}^{TT}, \dots, \Theta_{jR}^{TT}, \Theta_{j1}^{ET}, \dots, \Theta_{jS}^{ET}\Big]^{T},\\
\Theta_i^{E} &= \Big[\zeta_{i0}^{E}, \Theta_{i1}^{TE}, \dots, \Theta_{iR}^{TE}, \Theta_{i1}^{EE}, \dots, \Theta_{iS}^{EE}\Big]^{T},\\
c^{(\kappa)} &= \Big[1, \overline{X}_1^{(\kappa)}, \dots, \overline{X}_R^{(\kappa)}, \overline{E}_1^{(\kappa)}, \dots, \overline{E}_S^{(\kappa)}\Big]^{T}
\end{aligned}
$$

we obtain the robust counterpart program for an estimation of the parameters of linear cluster model (LCM) based on polyhedral uncertainty:

$$
\begin{aligned}
\text{Minimize}\quad & \sum_{\kappa=1}^{T}\left\{\sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)}\right\}\\
\text{such that}\quad & \Big|\big[c^{(\kappa-1)}\big]'\Theta_j^{T} - \overline{\mathbb{X}}_j^{(\kappa)}\Big| + \sum_{r=1}^{R}\big\|\mathrm{diag}\big((Err^{T})_r^{(\kappa-1)}\big)\Theta_{jr}^{TT}\big\|_\infty\\
&\qquad + \sum_{s=1}^{S}\big\|\mathrm{diag}\big((Err^{E})_s^{(\kappa-1)}\big)\Theta_{js}^{ET}\big\|_\infty \;\le p_j^{(\kappa)},\\
& \Big|\big[c^{(\kappa-1)}\big]'\Theta_i^{E} - \overline{\mathbb{E}}_i^{(\kappa)}\Big| + \sum_{r=1}^{R}\big\|\mathrm{diag}\big((Err^{T})_r^{(\kappa-1)}\big)\Theta_{ir}^{TE}\big\|_\infty\\
&\qquad + \sum_{s=1}^{S}\big\|\mathrm{diag}\big((Err^{E})_s^{(\kappa-1)}\big)\Theta_{is}^{EE}\big\|_\infty \;\le q_i^{(\kappa)}\\
& (\kappa = 1,\dots,T;\; j = 1,\dots,n;\; i = 1,\dots,m). \qquad \textbf{(RCPP)}
\end{aligned}
$$

For numerical examples and practical applications related to robust regression with polyhedral uncertainty we refer to [26, 42].

6. Robust Regression with Ellipsoidal Uncertainty

Ben-Tal and Nemirovski introduced the concept of robustness for programming problems where data is subject to ellipsoidal uncertainty [3, 5]. In general, an *ellipsoid* in $\mathbb{R}^p$ will be parameterized in terms of its center $c \in \mathbb{R}^p$ and a symmetric non-negative definite *configuration (or shape) matrix* $\Sigma \in \mathbb{R}^{p \times p}$ as

$$\mathcal{E}(c, \Sigma) = \{\Sigma u + c \,|\, \|u\|_2 \leq 1\}.$$

In order to include data uncertainty into our model, we now assume that the states of the clusters of target variables and environmental factors are subject to ellipsoidal uncertainty. That means, our regression analysis will be based on set-valued data

$$\begin{aligned} X_r^{(\kappa)} &\in \mathcal{E}\big(\overline{X}_r^{(\kappa)}, \overline{\Sigma}_r^{(\kappa)}\big) \subset \mathbb{R}^{|C_r|}, \\ E_s^{(\kappa)} &\in \mathcal{E}\big(\overline{E}_s^{(\kappa)}, \overline{\Pi}_s^{(\kappa)}\big) \subset \mathbb{R}^{|D_s|}, \end{aligned}$$

with $\kappa = 0, 1, \ldots, T$. The measurements $\overline{X}_r^{(\kappa)}$ and $\overline{E}_s^{(\kappa)}$ determine the centers of the ellipsoids and the corresponding symmetric shape matrices $\overline{\Sigma}_r^{(\kappa)}$ and $\overline{\Pi}_s^{(\kappa)}$ are given by the variance-covariance matrices of cluster data what also refers to partial correlations and partial variances of cluster elements.

We assume that the constraints are satisfied for all realizations of the states $X_r^{(\kappa)} \in \mathcal{E}\big(\overline{X}_r^{(\kappa)}, \overline{\Sigma}_r^{(\kappa)}\big)$ and $E_s^{(\kappa)} \in \mathcal{E}\big(\overline{E}_s^{(\kappa)}, \overline{\Pi}_s^{(\kappa)}\big)$ and in this way we obtain the following robust regression problem with uncertain ellipsoidal states:

$$\begin{aligned} \text{Minimize} \quad & \sum_{\kappa=1}^{T} \Big\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \Big\} \\ \text{such that} \quad & \Big| \zeta_{j0}^{T} + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} - \overline{\mathbb{X}}_j^{(\kappa)} \Big| \leq p_j^{(\kappa)}, \\ & \Big| \zeta_{i0}^{E} + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} - \overline{\mathbb{E}}_i^{(\kappa)} \Big| \leq q_i^{(\kappa)} \\ & (\kappa = 1, \ldots, T;\ j = 1, \ldots, n;\ i = 1, \ldots, m) \\ & \forall X_r^{(\kappa)} \in \mathcal{E}\big(\overline{X}_r^{(\kappa)}, \overline{\Sigma}_r^{(\kappa)}\big) \ \ (\kappa = 0, \ldots, T-1;\ r = 1, \ldots, R), \\ & \forall E_s^{(\kappa)} \in \mathcal{E}\big(\overline{E}_s^{(\kappa)}, \overline{\Pi}_s^{(\kappa)}\big) \ \ (\kappa = 0, \ldots, T-1;\ s = 1, \ldots, S). \end{aligned}$$

The above problem can be rewritten:

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{\kappa=1}^{T} \Big\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \Big\} \\
\text{such that} \quad & \zeta_{j0}^{T} + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} - \overline{\mathbb{X}}_j^{(\kappa)} \le p_j^{(\kappa)}, \\
& -\zeta_{j0}^{T} - \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} - \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} + \overline{\mathbb{X}}_j^{(\kappa)} \le p_j^{(\kappa)}, \\
& \zeta_{i0}^{E} + \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} + \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} - \overline{\mathbb{E}}_i^{(\kappa)} \le q_i^{(\kappa)}, \\
& -\zeta_{i0}^{E} - \sum_{r=1}^{R} \big[X_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} - \sum_{s=1}^{S} \big[E_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} + \overline{\mathbb{E}}_i^{(\kappa)} \le q_i^{(\kappa)} \\
& (\kappa = 1, \ldots, T;\ j = 1, \ldots, n;\ i = 1, \ldots, m) \\
& \forall X_r^{(\kappa)} \in \mathcal{E}\big(\overline{X}_r^{(\kappa)}, \overline{\Sigma}_r^{(\kappa)}\big) \ \ (\kappa = 0, \ldots, T-1;\ r = 1, \ldots, R), \\
& \forall E_s^{(\kappa)} \in \mathcal{E}\big(\overline{E}_s^{(\kappa)}, \overline{\Pi}_s^{(\kappa)}\big) \ \ (\kappa = 0, \ldots, T-1;\ s = 1, \ldots, S).
\end{aligned}
$$

This problem has an infinite number of constraints as it depends on all possible realizations of ellipsoidal states of targets and environmental factors. Another reformulation of this problem can be obtained when the ellipsoids are represented as follows:

$$
\begin{aligned}
\mathcal{E}\big(\overline{X}_r^{(\kappa)}, \overline{\Sigma}_r^{(\kappa)}\big) &= \Big\{ \overline{X}_r^{(\kappa)} + \overline{\Sigma}_r^{(\kappa)} u_r \ \Big|\ \|u_r\|_2 \le 1 \Big\}, \\
\mathcal{E}\big(\overline{E}_s^{(\kappa)}, \overline{\Pi}_s^{(\kappa)}\big) &= \Big\{ \overline{E}_s^{(\kappa)} + \overline{\Pi}_s^{(\kappa)} v_s \ \Big|\ \|v_s\|_2 \le 1 \Big\}.
\end{aligned}
$$

With

$$
\begin{aligned}
U_r &:= \Big\{ u_r \in \mathbb{R}^{|C_r|} \ \Big|\ \|u_r\|_2 \le 1 \Big\},\ r = 1, \ldots, R, \\
V_s &:= \Big\{ v_s \in \mathbb{R}^{|D_s|} \ \Big|\ \|v_s\|_2 \le 1 \Big\},\ s = 1, \ldots, S
\end{aligned}
$$

we then obtain the equivalent problem

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{\kappa=1}^{T} \left\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \right\} \\
\text{such that} \quad & \zeta_{j0}^{T} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} + \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \, \overline{\Sigma}_r^{(\kappa-1)} \, \Theta_{jr}^{TT} \big\} \\
& + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} + \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \, \overline{\Pi}_s^{(\kappa-1)} \, \Theta_{js}^{ET} \big\} \\
& - \overline{\mathbb{X}}_j^{(\kappa)} \le p_j^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ j = 1, \ldots, n), \\[2ex]
& -\zeta_{j0}^{T} - \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{jr}^{TT} - \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \, \overline{\Sigma}_r^{(\kappa-1)} \, \Theta_{jr}^{TT} \big\} \\
& - \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{js}^{ET} - \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \, \overline{\Pi}_s^{(\kappa-1)} \, \Theta_{js}^{ET} \big\} \\
& + \overline{\mathbb{X}}_j^{(\kappa)} \le p_j^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ j = 1, \ldots, n), \\[2ex]
& \zeta_{i0}^{E} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} + \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \, \overline{\Sigma}_r^{(\kappa-1)} \, \Theta_{ir}^{TE} \big\} \\
& + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} + \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \, \overline{\Pi}_s^{(\kappa-1)} \, \Theta_{is}^{EE} \big\} \\
& - \overline{\mathbb{E}}_i^{(\kappa)} \le q_i^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ i = 1, \ldots, m), \\[2ex]
& -\zeta_{i0}^{E} - \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)}\big]' \Theta_{ir}^{TE} - \sum_{r=1}^{R} \max_{u_r \in U_r} \big\{ u_r' \, \overline{\Sigma}_r^{(\kappa-1)} \, \Theta_{ir}^{TE} \big\} \\
& - \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)}\big]' \Theta_{is}^{EE} - \sum_{s=1}^{S} \max_{v_s \in V_s} \big\{ v_s' \, \overline{\Pi}_s^{(\kappa-1)} \, \Theta_{is}^{EE} \big\} \\
& + \overline{\mathbb{E}}_i^{(\kappa)} \le q_i^{(\kappa)} \quad (\kappa = 1, \ldots, T;\ i = 1, \ldots, m).
\end{aligned}
$$

The equations

$$
\begin{aligned}
\max_{u_r \in U_r} \big\{ u_r' \, \overline{\Sigma}_r^{(\kappa)} \, \Theta_{jr}^{TT} \big\} &= \max_{u_r \in U_r} \big\{ -u_r' \, \overline{\Sigma}_r^{(\kappa)} \, \Theta_{jr}^{TT} \big\} = \Big\| \overline{\Sigma}_r^{(\kappa)} \Theta_{jr}^{TT} \Big\|_2, \\
\max_{u_r \in U_r} \big\{ u_r' \, \overline{\Sigma}_r^{(\kappa)} \, \Theta_{ir}^{TE} \big\} &= \max_{u_r \in U_r} \big\{ -u_r' \, \overline{\Sigma}_r^{(\kappa)} \, \Theta_{ir}^{TE} \big\} = \Big\| \overline{\Sigma}_r^{(\kappa)} \Theta_{ir}^{TE} \Big\|_2, \\
\max_{v_s \in V_s} \big\{ v_s' \, \overline{\Pi}_s^{(\kappa)} \, \Theta_{is}^{ET} \big\} &= \max_{v_s \in V_s} \big\{ -v_s' \, \overline{\Pi}_s^{(\kappa)} \, \Theta_{is}^{ET} \big\} = \Big\| \overline{\Pi}_s^{(\kappa)} \Theta_{is}^{ET} \Big\|_2, \\
\max_{v_s \in V_s} \big\{ v_s' \, \overline{\Pi}_s^{(\kappa)} \, \Theta_{is}^{EE} \big\} &= \max_{v_s \in V_s} \big\{ -v_s' \, \overline{\Pi}_s^{(\kappa)} \, \Theta_{is}^{EE} \big\} = \Big\| \overline{\Pi}_s^{(\kappa)} \Theta_{is}^{EE} \Big\|_2
\end{aligned}
$$

lead to a further description of the regression problem:

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{\kappa=1}^{T} \left\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \right\} \\
\text{such that} \quad & \left| \zeta_{j0}^{T} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)} \big]' \Theta_{jr}^{TT} + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)} \big]' \Theta_{js}^{ET} - \overline{\mathbb{X}}_j^{(\kappa)} \right| \\
& \quad + \sum_{r=1}^{R} \| \overline{\Sigma}_r^{(\kappa-1)} \Theta_{jr}^{ET} \|_2 + \sum_{s=1}^{S} \| \overline{\Pi}_s^{(\kappa-1)} \Theta_{js}^{ET} \|_2 \le p_j^{(\kappa)}, \\
& \left| \zeta_{i0}^{E} + \sum_{r=1}^{R} \big[\overline{X}_r^{(\kappa-1)} \big]' \Theta_{ir}^{TE} + \sum_{s=1}^{S} \big[\overline{E}_s^{(\kappa-1)} \big]' \Theta_{is}^{EE} - \overline{\mathbb{E}}_i^{(\kappa)} \right| \\
& \quad + \sum_{r=1}^{R} \| \overline{\Sigma}_r^{(\kappa-1)} \Theta_{ir}^{TE} \|_2 + \sum_{s=1}^{S} \| \overline{\Pi}_s^{(\kappa-1)} \Theta_{is}^{EE} \|_2 \le q_i^{(\kappa)} \\
& (\kappa = 1, \ldots, T;\ j = 1, \ldots, n;\ i = 1, \ldots, m).
\end{aligned}
$$

Finally, with the vectors

$$
\begin{aligned}
\Theta_j^T &= \left[\zeta_{j0}^T, \Theta_{j1}^{TT}, \ldots, \Theta_{jR}^{TT}, \Theta_{j1}^{ET}, \ldots, \Theta_{jS}^{ET} \right]^T, \\
\Theta_i^E &= \left[\zeta_{i0}^E, \Theta_{i1}^{TE}, \ldots, \Theta_{iR}^{TE}, \Theta_{i1}^{EE}, \ldots, \Theta_{iS}^{EE} \right]^T, \\
c^{(\kappa)} &= \left[1, \overline{X}_1^{(\kappa)}, \ldots, \overline{X}_R^{(\kappa)}, \overline{E}_1^{(\kappa)}, \ldots, \overline{E}_S^{(\kappa)} \right]^T
\end{aligned}
$$

we obtain the robust counterpart program for an estimation of the parameters of linear cluster model (LCM) based on ellipsoidal uncertainty:

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{\kappa=1}^{T} \left\{ \sum_{j=1}^{n} p_j^{(\kappa)} + \sum_{i=1}^{m} q_i^{(\kappa)} \right\} \\
\text{such that} \quad & \left| \big[c^{(\kappa-1)} \big]' \Theta_j^T - \overline{\mathbb{X}}_j^{(\kappa)} \right| + \sum_{r=1}^{R} \| \overline{\Sigma}_r^{(\kappa-1)} \Theta_{jr}^{TT} \|_2 + \sum_{s=1}^{S} \| \overline{\Pi}_s^{(\kappa-1)} \Theta_{js}^{ET} \|_2 \le p_j^{(\kappa)}, \\
& \left| \big[c^{(\kappa-1)} \big]' \Theta_i^E - \overline{\mathbb{E}}_i^{(\kappa)} \right| + \sum_{r=1}^{R} \| \overline{\Sigma}_r^{(\kappa-1)} \Theta_{ir}^{TE} \|_2 + \sum_{s=1}^{S} \| \overline{\Pi}_s^{(\kappa-1)} \Theta_{is}^{EE} \|_2 \le q_i^{(\kappa)} \\
& (\kappa = 1, \ldots, T;\ j = 1, \ldots, n;\ i = 1, \ldots, m). \qquad \textbf{(RCPE)}
\end{aligned}
$$

For further details on robust optimization and the numerical treatment of the corresponding uncertainty-affected programming problems with ellipsoidal uncertainty we refer to [6, 25].

7. Conclusion

In this paper, we offer a new perspective for the identification of the intrinsic parameters of complex target-environment networks under polyhedral and ellipsoidal uncertainty. By this we further extend earlier approaches that are based on interval arithmetics where Chebychev approximation and generalized semi-infinite optimization is applied. In addition, the robust regression approach complements the recently introduced regression analysis for clusters in target-environment networks that is based on ellipsoidal calculus. In future work, combinations of both approaches have to considered so that clusters may take influence on target and environmental clusters as well as single genes and single environmental factors simultaneously. Furthermore, the introduction of least squares type regression problems as well as Chebychev and general p-norm regression models with both polyhedral and ellipsoidal uncertainty can contribute to the rich variety of methods for parameter estimation of gene-environment and eco-finance networks.

Acknowledgement

With this research paper, the authors would like to express their gratefulness to Professor Florian Potra's scientific contributions, by particularly addressing his interest in the areas conic programming and interior point methods, of dynamical systems and computational biology.

References

[1] M.U. Akhmet, J. Gebert, H. Öktem, S.W. Pickl, and G.-W. Weber, *An improved algorithm for analytical modeling and anticipation of gene expression patterns*, Journal of Computational Technologies 10(4) (2005), pp. 3–20.

[2] E. Borenstein, and M.W. Feldman, *Topological signatures of species interactions in metabolic networks*, J. Comput. Biol. 16(2) (2009), pp. 191–200.

[3] A. Ben-Tal, Conic and robust optimization, Lecture notes (2002) Available at http://iew3.technion.ac.il/Home/Users/morbt.phtml.

[4] A. Ben-Tal, and A. Nemirovski, *Robust Convex Optimization*, Mathematics of Operations Research 23(4) (1998), pp. 769–805.

[5] A. Ben-Tal, and A. Nemirovski, *Convex Optimization in Engineering - Modeling, Analysis, Algorithms*, Faculty of Industrial Engineering and Management Technion-City, Haifa 32000, Israel, available at www.isa.ewi.tudelft.nl/ roos/courses/WI4218/tud00r.pdf.

[6] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*, Princeton University Press 2009.

[7] T. Chen, H.L. He, and G.M. Church, *Modeling gene expression with differential equations*, Proc. Pacific Symposium on Biocomputing (1999), pp. 29–40.

[8] T. Ergenç and G.-W. Weber, *Modeling and prediction of gene-expression patterns reconsidered with Runge-Kutta discretization*, in the special issue at the occasion of seventieth birthday of Prof. Dr. Karl Roesner, TU Darmstadt, Journal of Computational Technologies 9(6) (2004), pp. 40–48.

[9] J. Gebert, M. Lätsch, S.W. Pickl, N. Radde, G.-W. Weber, and R. Wünschiers, *Genetic networks and anticipation of gene expression patterns*, Computing Anticipatory Systems: CASYS(92)03 – Sixth International Conference, AIP Conference Proceedings 718 (2004), pp. 474–485.

[10] J. Gebert, M. Lätsch, S.W. Pickl, G.-W. Weber, and R. Wünschiers, *An algorithm to analyze stability of gene-expression pattern*, in M. Anthony, E. Boros, P.L. Hammer, and A. Kogan (guest eds.), special issue *Discrete Mathematics and Data Mining II* of Discrete Appl. Math. 154(7) (2006), pp. 1140–1156.

[11] J. Gebert, H. Öktem, S.W. Pickl, N. Radde, G.-W. Weber, and F.B. Yılmaz, *Inference of gene expression patterns by using a hybrid system formulation – an algorithmic approach to local state transition matrices*, in G.E. Lasker and D.M. Dubois (eds.) Anticipative and predictive models in systems science I, IIAS (International Institute for Advanced Studies) in Windsor, Ontario, 2004, pp. 63–66.

[12] J. Gebert, M. Lätsch, E.M.P. Quek, and G.-W. Weber, *Analyzing and optimizing genetic network structure via path-finding*, Journal of Computational Technologies 9(3) (2004), pp. 3–12.

[13] A. Gökmen, S. Kayalgil, G.-W. Weber, I. Gökmen, M. Ecevit, A. Sürmeli, T. Bali, Y. Ecevit, H. Gökmen, and D.J. DeTombe, *Balaban Valley Project: Improving the Quality of Life in Rural Area in Turkey*, International Scientific Journal of Methods and Models of Complexity 7(1) (2004).

[14] J.R. Harris, W. Nystad, and P. Magnus, *Using genes and environments to define asthma and related phenotypes: applications to multivariate data*, Clinical and Experimental Allergy 28(1) (1998), pp. 43–45.

[15] M.D. Hoon, S. Imoto, K. Kobayashi, N. Ogasawara, and S. Miyano, *Inferring gene regulatory networks from time-ordered gene expression data of Bacillus subtilis using differential equations*, Proc. Pacific Symposium on Biocomputing (2003), pp. 17–28.

[16] A. Işcanoğlu, G.-W. Weber, and P. Taylan, *Predicting default probabilities with generalized additive models for emerging markets*, invited lecture, Graduate Summer School on New Advances in Statistics, METU (2007).

[17] H.D. Jong, *Modeling and simulation of genetic regulatory systems: a literature review*, J. Comput. Biol. 9 (2002), pp. 103–129.

[18] W. Krabs and S. Pickl, *A game-theoretic treatment of a time-discrete emission reduction model*, Int. Game Theory Rev. 6 (1) (2004), pp. 21–34.

[19] E. Kropat, G.-W. Weber, and B. Akteke-Öztürk, *Eco-finance networks under uncertainty*, in J. Herskovits, A. Canelas, H. Cortes, and M. Aroztegui (eds.), *Proceedings of the International Conference on Engineering Optimization* (ISBN 978857650156-5, CD), EngOpt 2008, Rio de Janeiro, Brazil (2008).

[20] E. Kropat, G.-W. Weber, S. Belen, *Dynamical gene-environment networks under ellipsoidal uncertainty - Set-theoretic regression analysis based on ellipsoidal OR*, to appear in M. Peixoto, D. Rand, A. Pinto (eds.): Dynamics, Games and Science, Springer-Verlag 2010.

[21] E. Kropat, G.-W. Weber, C.S. Pedamallu, *Regulatory Networks under Ellipsoidal Uncertainty - Data Analysis and Prediction by Optimization Theory and Dynamical Systems*, to appear in the book on "Data Mining" in preparation, D. Holmes (ed.), Springer-Verlag 2010.

[22] E. Kropat, G.-W. Weber, J.-J. Rückmann, *Regression analysis for clusters in gene-environment networks based on ellipsoidal calculus and optimization*, Dynamics of Continuous, Discrete and Impulsive Systems, Series B: Applications & Algorithms 17(5) (2010), pp. 639-657.

[23] Y.F. Li, S. Venkatesh, and D. Li, *Modeling global emissions and residues of pesticided*, Environmental Modeling and Assessment 9 (2004), pp. 237–243.

[24] Q. Liu, J. Yang, Z. Chen, M.Q. Yang, A.H. Sung, and X. Hunag, *Supervised learning-based tagSNP selection for genome-wide disease classifications*, BMC Genomics, 9, 1 (2007).

[25] A. Nemirovski, *Modern Convex Optimization*, Lecture at PASCAL Workshop, Thurnau, Germany, March 16-18 (2005).

[26] A. Özmen, *Robust conic quadratic programming applied to quality improvement - A robustification of CMARS*, Master Thesis, Middle East Technical University, Ankara, Turkey, 2010.

[27] S. Özöğür, *Mathematical modelling of enzymatic reactions, simulation and parameter estimation*, MSc. thesis at Institute of Applied Mathematics, METU, Ankara (2005).

[28] M. Partner, N. Kashtan, and U. Alon, *Environmental variability and modularity of bacterial metabolic network*, BMC Evolutionary Biology 7 (2007), 169doi:10.1186/1471-2148-7-169.

[29] S. Pickl, *Der τ-value als Kontrollparameter - Modellierung und Analyse eines Joint-Implementation Programmes mithilfe der dynamischen kooperativen Spieltheorie und der diskreten Optimierung*, Thesis, Darmstadt University of Technology, Department of Mathematics, 1998.

[30] S. Pickl, *An iterative solution to the nonlinear time-discrete TEM model - the occurence of chaos and a control theoretic algorithmic approach*, AIP Conference Proceedings 627(1) (2002), pp. 196–205.

[31] S. Pickl, E. Kropat, and H. Hahn, *The impact of uncertain emission trading markets on interactive resource planning processes and international emission trading experiments*, Climatic Change (2010), DOI 10.1007/s10584-010-9912-8.

[32] S. Pickl and G.-W. Weber, *Optimization of a time-discrete nonlinear dynamical system from a problem of ecology - an analytical and numerical approach*, Journal of Computational Technologies 6(1) (2001), pp. 43–52.

[33] M. Taştan, *Analysis and prediction of gene expression patterns by dynamical systems, and by a combinatorial algorithm*, MSc Thesis, Institute of Applied Mathematics, METU, Turkey, 2005.

[34] M. Taştan, T. Ergenç, S.W. Pickl, and G.-W. Weber, *Stability analysis of gene expression patterns by dynamical systems and a combinatorial algorithm*, in HIBIT – Proceedings of International Symposium on Health Informatics and Bioinformatics, Turkey '05, Antalya, Turkey, 2005, pp. 67–75.

[35] M. Taştan, S.W. Pickl, and G.-W. Weber, *Mathematical modeling and stability analysis of gene-expression patterns in an extended space and with Runge-Kutta discretization*, in *Proceedings of Operations Research 2005*, Bremen, September 2005, Springer, pp. 443–450.

[36] P. Taylan, G.-W. Weber, and A. Beck, *New approaches to regression by generalized additive models and continuous optimization for modern applications in finance, science and technology*, in the special issue in honour of Prof. Dr. Alexander Rubinov, B. Burachik and X. Yang (guest eds.) of Optimization 56, 5-6 (2007), pp. 675698.

[37] P. Taylan, G.-W. Weber, and F. Yerlikaya, *A new approach to multivariate adaptive regression spline by using Tikhonov regularization and continuous optimization*, to appear in TOP (the Operational Research journal of SEIO (Spanish Statistics and Operations Research Society)), Selected Papers at the Occasion of 20th EURO Mini Conference Continuous Optimization and Knowledge-Based Technologies, Neringa, Lithuania, May 20-23, 2008.

[38] P. Taylan, G.-W. Weber, L. Liu, and F. Yerlikaya, *On foundations of parameter estimation for generalized partial linear models with B-splines and continuous optimization*, Computers and Mathematics with Applications (CAMWA) 60(1) (2010), pp. 134-143.

[39] Ö. Uğur, S.W. Pickl, G.-W. Weber, and R. Wünschiers, Operational research meets biology: An algorithmic approach to analyze genetic networks and biological energy production, Preprint no. 50, Institute of Applied Mathematics, METU, 2006, submitted for the special issue of *Optimization* at the occasion of the 5th Ballarat Workshop on Global and Non-Smooth Optimization: Theory, Methods, and Applications (2006).

[40] Ö. Uğur, and G.-W. Weber, *Optimization and dynamics of gene-environment networks with intervals*, in the Special Issue at the occasion of the 5th Ballarat Workshop on Global and Non-Smooth Optimization: Theory, Methods and Applications, November 28-30, 2006, of J. Ind. Manag. Optim. 3(2) (2007), pp. 357–379.

[41] G.-W. Weber, S.Z. Alparslan-Gök, and N. Dikmen, Environmental and life sciences: gene-environment networks - optimization, games and control - a survey on recent achievements, Preprint at Institute of Applied Mathematics, invited paper, in the Special Issue of *Journal of Organisational Transformation and Social Change* 5(3) (2008), pp. 197–233.

[42] G.-W. Weber, I. Batmaz, and A. zmen, *Robust conic quadratic programming - A robustification of CMARS*, submitted to Optimization.

[43] G.-W. Weber, S.Z. Alparslan-Gök, and B. Söyler, *A new mathematical approach in environmental and life sciences: gene-environment networks and their dynamics*, Environmental Modeling & Assessment

14(2) (2009), pp. 267-288.

[44] G.-W. Weber, R. Branzei, and S.Z. Alparslan Gök, *On cooperative ellipsoidal games*, 24th Mini EURO Conference - On Continuous Optimization and Information-Based Technologies in the Financial Sector, MEC EurOPT 2010, Selected Papers, ISI Proceedings, Izmir, Turkey, June 23-26, 2010, pp. 369-372.

[45] G.-W. Weber, R. Branzei, and S.Z. Alparslan Gök, *On the ellipsoidal core for cooperative games under ellipsoidal uncertainty*, in the proceedings of 2nd International Conference on Engineering Optimization, Lisbon, Portugal, September 6-9 (2010) (on a CD-Rom).

[46] G.-W. Weber, E. Kropat, B. Akteke-Öztürk, and Z.-K. Görgülü, *A survey on OR and mathematical methods applied on gene-environment networks*, in the Special Issue on *Innovative Approaches for Decision Analysis in Energy, Health, and Life Sciences* of Central European Journal of Operations Research (CEJOR) at the occasion of EURO XXII 2007 (Prague, Czech Republic, July 8-11, 2007), Vol. 17, No. 3 (2009), pp. 315–341.

[47] G.-W. Weber, E. Kropat, A. Tezel, and S. Belen, *Optimization applied on regulatory and eco-finance networks - survey and new developments*, in M. Fukushima et al. (guest eds.) Pac. J. Optim. 6(2) (2010), pp.319–340.

[48] G.-W. Weber, S. Özögür-Akyüz, and E. Kropat, *A review on data mining and continuous optimization applications in computational biology and medicine*, Embryo Today, Birth Defects Research (Part C) 87 (2009), pp. 165–181.

[49] G.-W. Weber, P. Taylan, S.-Z. Alparslan-Gök, S. Özöğür, and B. Akteke-Öztürk, *Optimization of gene-environment networks in the presence of errors and uncertainty with Chebychev approximation*, TOP, the Operational Research journal of SEIO (Spanish Statistics and Operations Research Society) 16(2) (2008), pp. 284-318.

[50] G.-W. Weber, P. Taylan, K. Yıldırak, and Z.K. Görgülü, *Financial regression and organization*, in the Special Issue on Optimization in Finance, of Dynamics of Continuous, Discrete and Impulsive Systems - Series B (DCDIS-B) 17, 1b (2010), pp. 149-174.

[51] G.-W. Weber, P. Taylan, Z.-K. Görgülü, H. Abd. Rahman, and A. Bahar, *Parameter estimation in stochastic differential equations*, to appear as a Chapter in the Book *Dynamics, Games and Science* (in Honour of Mauricio Peixoto and David Rand), Springer-Verlag 2010; M. Peixoto, D. Rand and A. Pinto (eds.).

[52] G.-W. Weber, and A. Tezel, *On generalized semi-infinite optimization of genetic networks*, TOP 15(1) (2007), pp. 65–77.

[53] G.-W. Weber, A. Tezel, P. Taylan, A. Soyler, and M. Çetin, *Mathematical contributions to dynamics and optimization of gene-environment networks*, in the Special Issue: In Celebration of Prof. Dr. Dr. Hubertus Th. Jongen's 60th Birthday, D. Pallaschke, O. Stein (guest eds.), of *Optimization* 57(2) (2008), pp. 353-377.

[54] G.-W. Weber, Ö. Uğur, P. Taylan, and A. Tezel, *On optimization, dynamics and uncertainty: a tutorial for gene-environment networks*, in the Special Issue *Networks in Computational Biology* of *Discrete Appl. Math.*, 157(10) (2009), pp. 2494–2513.

[55] F.B. Yılmaz, *A mathematical modeling and approximation of gene expression patterns by linear and quadratic regulatory relations and analysis of gene networks*, MSc Thesis, Institute of Applied Mathematics, METU, Ankara, Turkey, 2004.

[56] F.B. Yılmaz, H. Öktem, and G.-W. Weber, *Mathematical modeling and approximation of gene expression patterns and gene networks*, in F. Fleuren, D. den Hertog, and P. Kort (eds.) Operations Research Proceedings (2005), pp. 280–287.