SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Uncertain mixed strategy games

av

Leo Kurpatow

2020 - No K40

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Uncertain mixed strategy games

Leo Kurpatow

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2020

Uncertain mixed strategy games

Sasha Kurpatow

August 19, 2020

Abstract

This paper will deal with how to transform an uncertain 2 person game to a robust tractable program and how to solve such a program. The program that we will be using are second order programs (SOCP) and semi-definte programs (SDP). Both of these programs are some common conic programs (CP), which we will introduce through adapting linear programming (LP). There are some theorems and definitions that will frequently occur such as the duality theorem and Karuch-Kuhn-Tuckers conditions, that will be essential when we are creating our aspiring program. We will in the last section discuss how and in which fields robust programming is applied in.

Keywords— linear programming, conic programming, robust programming, game theory, zero sum game, uncertain bimatrix game.

Contents

1 Introduction

There are many reasons why mathematics is such an important subject, and one of the most underlying reason, which one simply cannot ignore is its applications. A mathematical program is a structure in the form of an objective function that orientates in a space, and its data is a restriction to this space. Our goal is to either find an optimal value of our objective function within this space, or to collect solutions that can generate such a value, wheres an optimal value is a maximum or minimum value. However, for any realistic problems uncertainty will be an inevitable factor. The mathematical interpretation of uncertainty is that our data is only so called nominal values, which can vary with a perturbation. For a mathematical program with a lot of constraints a rather small perturbation can affect the feasibility of our solution and make it infeasible, thus resulting that the true answer is very different. Hence, we want to make our program robust such that it takes these perturbations in account. Whiles we want to make our program robust we still want to get a solution, i.e. we want to create a tractable model. There are different ways one can acquire such a model, in this paper we have chosen to create a tractable model by assuming that the uncertainty set can be described in the form of a box or ellipsoid.

There are many ways to acquire a mathematical program, if it is involving agents it is quite common that our data takes form of decisions. A 2-person game or a bimatrix game is a decision problem involving two agents, where each agent wants to optimise their payoff. Hence, each agent can transform their decision problem to a mathematical program and their opponents possible response will be the agents constraints. In the case of an uncertain bimatirx game one or more of the components of the game must be uncertain. We will only treat the scenario where an agent is uncertain of their opponents strategy and when an agent is uncertain of their own payoff.

We will give two separate sections for the duality theorem and Karush-Kuhn-Tucker conditions, because their role in uncertain bimatrix games are particular important. The duality is best described through an example:

Suppose that we try to maximise our profit given that we don't exceed our budget in terms of costs, its dual reformulation would then be that we want to minimize our costs given that our profit must exceed our maximal profit in our previous problem.

We can do a similar analogy in a bimatrix, given that the opponent wants to minimize the players payoff.

In the upcoming section we will focus on linear programming, and then in the conic program section show how we can generalise those methods in order to solve none linear problems, which an uncertain bimatrix game can be transformed to.

2 Linear programming

2.1 Formalising an LP

We begin by observing a general mathematical program (MP), i.e. a maximisation or minimisation problem under some constraints. When we create an MP we will first have to analyse the decision variables, call them x. These variables are bounded in some data set K and if K is not a finite set then that is in itself a constraint, i.e. a restriction to the feasibility. However, if nothing is mentioned, we will assume that $x \in R^n$. We can describe our constraints as a required aggregation that needs to be fulfilled, i.e. a constraint function of the form $g(x)\,\Omega\,b$. Where Ω can be an inequality or equality relation. Thus, our data set and constraint functions will create a feasible space, which we denote as $\mathcal{F}$. The function that we want to optimise in relation to $\mathcal{F}$ is called the objective function or cost function. If we now assume that we have an MP of the form of a minimisation with m constraints then we can give a formal notation of a MP

$$\begin{aligned} \min_{\text{(or max)}} \; & f(x) \\ \text{s.t.} \;\; & g(x)\,\Omega\,b, \end{aligned} \tag{MP}$$

where $x \in R^n$, $b \in R^m$ and $f : R^n \to R^m$. We have denoted "subjected to" as "s.t" and we will use this denotation throughout the paper. Assume that we are given a minimisation problem, then the constraints must have a lower bound in order for the program to have a minimal value. Thus, for such a problem to be meaningful Ω must then either denote; "$>$", "$=$" or "$\geq$". We can make a similar setting for a maximisation problem. We can now give a definition for an LP.

Definition 2.1. *An MP is an LP when its objective function and constraints are linear.*

A useful technique that we will use when we have an MP is the introduction of slack variables, that are variables which allow us to manipulate the Ω sign, but the most frequent usage is to transform an inequality to an equality. A specific MP can be interpret as a program that has certain features, i.e. our structure and data must fulfill certain requirements. We will now introduce a linear program and we will out of convention refer this as the standard form of an LP.

$$\begin{aligned} \min \; & c^T x \\ \text{s.t.} \;\; & Ax \geq b, \end{aligned} \tag{LP}$$

whereas A is an $m \times n$ constraint matrix with real variables and $b \in R^m$. Hence, our standard LP model is a minimisation problem with inequality constraints. A more common convention for a standard LP which we can find on [2, page 15] is that we set the optimal value as ∞ when it is infeasible. Why? It is mainly for algorithmic reasons which ensures us to always find a new candidate, since any other feasible solution will give a smaller value. Notice that if it is feasible and unbounded from below, then the objective value is $-\infty$. We can now give a notation of the optimal value.

Notation 1. *If an LP has a nonempty feasible set $\mathcal{F}$ and is bounded below (i.e. there exists a finite infimum for $\mathcal{F}$), then we can define the optimal value c^* as the infimum* *value amongst the feasible solutions, i.e.*

$$\mathcal{F} = \{x : Ax - b \geq 0\} \neq \emptyset$$
$$c^* = \inf_{\{x:Ax-b\geq 0\}} c^T x.$$

2.2 Certifying solvability for LP

A natural question is to ask, how do we verify that an LP problem is solvable? To get an intuition to this we can aggregate our inequalities with a positive vector λ, into one equation, which we will denote as Cons(λ). Formally, Cons(λ) is the the inner product such that

$$\langle \lambda, Ax \rangle \ := \ \lambda^T Ax \geq \lambda^T b. \tag{Cons(λ)}$$

If we assume that our objective function is strict greater then the vector a, i.e. we assume that a is a lower bound to our problem. Then there should not exists any solution such that our objective function equals a, hence we can verify feasibility by verifying that the following system has no solutions

$$\begin{aligned} -c^T x + a &> 0 \\ (Ax - b)\, \Omega\, 0&, \end{aligned} \tag{$\mathcal{S}$}$$

where Ω is either "$>$" or "$\geq$". If we now rewrite our original LP as a summation of all the affine combinations with the vector $\lambda \geq 0$, i.e. as Cons(λ). Then we get the following system

$$(\sum_{i=1}^{m} \lambda_i a_i)^T x \, \Omega \, \sum_{i=1}^{m} \lambda_i b_i. \tag{Cons(λ)}$$

As mentioned above, our LP has no solutions if $\mathcal{S}$ has no solutions. Since $\lambda \geq 0$ we see that we can make the the following conventions for Cons(λ)

$$\sum_{i=1}^{m} \lambda_i a_i)^T = d^T$$
$$\sum_{i=1}^{m} \lambda_i b_i = e,$$

where d is a vector and e is a scalar. This gives us the refined system

$$d^T x \, \Omega \, e$$

Thus, for our LP to be infeasible then our refined system needs to be unsolvable. We can clearly see that if we set $d = 0$ then we will get a contradiction for all $e > 0$ and if Ω is a strict inequality then we will get a contradiction for $e = 0$ as well.

We now want to propose a theorem which describes the feasibility of an LP, it is similarly to the one in [2, page 18].

Theorem 2.1. *A LP that is a minimisation problem, i.e.*

$$\begin{aligned} &min\ c^T x \\ &s.t\ Ax \, \Omega \, b \end{aligned} \tag{LP'}$$

where Ω is "$\geq$" or "$>$", is infeasible if and only if there exists an affine combination with a vector $\lambda \geq 0$ such that the following system is solvable

$$\begin{cases} (a) \ \sum_{i=1}^{m} \lambda_i > 0 \\ (b) \ \sum_{i=1}^{m} \lambda_i a_i = 0 \\ (c) \ \sum_{i=1}^{m} \lambda_i b_i \ \Omega \ 0. \end{cases}$$

2.3 The Dual theorem

The duality is a technique that systematically generates a lower bound to the "primal" problem. By primal we mean the optimisation problem that was given to us. We are essentially transforming our problem such that it is an optimisation problem in respect to our constraints. A consequence from our transformation is that our dual problem is always convex, which we will show later on. In practices it is easier to implement algorithms to such a program. Because, a convex program with a differentiable objective function must the local optimal point be a global optimal point as well, a simple geometric interpretation will verify this. Hence, the dual holds many positive properties. Since we will from now on be working with two programs simultaneously, we will refer the programs as the primal and the dual, where the primal is the program we are given. Assume that our primal is the following program

$$c^* = \min_{x \in R^n} f(x) \qquad \text{s.t. } g(x) \geq b, \ i = 1, ..., m. \tag{P}$$

In order to formulate the dual program (D), we will have to create a program such that its objective function is a lower bound to the primal programs objective function. We can accomplish this in the following manner:

Introduce $\lambda \in R_+^m$ and let $f_\lambda(x)$ denote the Lagrangian function. If we now minimise $f_\lambda(x)$ with respect to x, then we can create a function $\phi(\lambda)$ such that

$$\phi(\lambda) = \min \ f(x) + \lambda^T (b - g(x)).$$

If we choose a $\lambda \geq 0$ such that $\phi(\lambda)$ is the maximal value, then it will be a lower bound to the primal problem, i.e. the dual to the primal. This can be written as

$$\max \ \{\phi(\lambda) : \lambda \geq 0\}, \tag{D}$$

where $\phi(\lambda)$ is the objective function to the dual. We will now work on the maximisation problem to the LP problem of the form

$$c^* = \min_{x \in R^n} f(x) \qquad \text{s.t. } g(x) \geq b, \ i = 1, ..., m. \tag{P}$$

Now, introduce $\lambda \geq 0$ so that we

$$\begin{aligned} &\min \ c^T x + \lambda^T (b - Ax) \\ &\text{s.t.} \quad x \geq 0, \end{aligned}$$

where the objective function is equal to

$$\lambda^T b + (c^T - \lambda^T A)x.$$

Since $\lambda^T b$ is independent of x will we be working on $(c^T - \lambda^T A)x$. If

$$c^T - \lambda^T A \geq 0,$$

then $x = 0$ will do the job. Otherwise, if some component of

$$c^T - \lambda^T A < 0,$$

say the i-th, then $x_i = \infty$ gives minimum. Since we will maximise the minimal Lagrange function in order to get the dual, then the dual function must be be $\lambda^T b$. Therefor the dual is

$$\max \{b^T \lambda : A^T \lambda - c \geq 0, \lambda \geq 0\}. \tag{D}$$

We can now give a formal statement.

Theorem 2.2. *(Weak duality) The optimal value in the dual is always less then or equal to the optimal value in the primal.*

This is evident from how we choose to define the dual. If we assume that the optimal value for the prime is c^* and the optimal value for the dual is b^*, then we can express the duality gap as:

$$\sigma = c^* - b^*,$$

where σ is the duality gap. We can now present the duality theorem in LP, which we will present in a similar fashion as in [2, page 23]

Theorem 2.3. *(The strong duality theorem) If we have two programs like in (P) and (D) then the optimal values for (P) and (D) are equal when;*

$$\begin{aligned} (i)&\ \textit{The primal is feasible and bounded below} \\ (ii)&\ \textit{The dual is feasible and bounded above} \\ (iii)&\ \textit{The primal is solvable} \\ (iv)&\ \textit{The dual is solvable} \\ (v)&\textit{Both the primal and the dual are feasible.} \end{aligned}$$

The proof can be read in [2], but readily it is: If the (P) is feasible and bounded below (i) implies that c^* exists and weak duality implies that this value must equal to b^*. If (D) is is feasible and bounded above then (P) must be so as well. If both of the program are feasible then that implies that both of them need to be bounded.

We can now summarise the above as a more convenient theorem.

Theorem 2.4. *A pair (x^*, y^*) are feasible solutions to the primal and dual if and only if the complementary slackness holds or the duality gap is zero, where*

$$\lambda_i = 0 \ \textit{or} \ b_i - \sum_{j=1}^{n} a_{ij} x_j = 0 \qquad \text{(complementary slackness)}$$

$$c^T x^* - b^T y^* = 0. \qquad \text{(zero duality gap)}$$

The zero duality gap holds by definition. The complementary slackness is a reformulation of feasibility for both the programs.

2.4 Quadratic programming problem

We will in this section show how the theory presented earlier is applied by studying the quadratic programming (QP) problem. A quadratic function in variables $x = (x_1, ..., x_n)$ is a polynomial in x where the maximum degree of nominals is equal to two. Assume that; Q is an $n \times n$ semi-definite positive matrix, A is a $m \times n$ matrix, b is an n-vector and c is an n-vector. Now we consider the quadratic programming problem (QP)

$$\begin{aligned} \min \ & \frac{1}{2}x^T Qx + c^T x, \\ \text{s.t.} \ & Ax = b \\ & x \geq 0. \end{aligned} \tag{QP}$$

If we begin by observing the geometry of this sort of problem then we can notice that we are minimising a convex quadratic function over a polyhedron described by $\{x : Ax = b, x \geq 0\}$, where the solution are the points where the quadratic function and the polyhedral set (or polyhedron) intersect. From this we get that for us to gain a solution there exists two geometrical positions, which are;

(i) The optimal solution is on the boundary of the polyhedron.
(ii) The optimal solution is inside the polyhedron.

Since our objective function is a quadratic function we know that the level curves are ellipses. Thus, in (i) we have that our optimal solution x^* is the intersection of one of the hyperplanes and the tangent to an ellipsoidal.

For the case in (ii) we have

$$\arg\min_x \frac{1}{2}x^T Qx + c^T x = -Q^+ c,$$

is inside the polyhedron. From this we can see that $-Q^+ c$ is the center of the ellipsoidal level curves. Where Q^+ is the pseudo inverse of Q and in particular $Q^+ = Q^{-1}$ if Q is invertible.

Note that if the polyhedron is empty the the problem is infeasible, i.e. there is no solution; and if the polyhedron is unbounded in the opposite direction of c and $Qc = 0$ then the QP is unbounded. However, if Q is positive definite, then QP cannot be unbounded, for a more detailed explanation see [18, page 127-130].

Let us now find the dual problem of this QP problem.

We can write the Lagrange function as

$$\begin{aligned} L(x, \lambda, \mu) &= \frac{1}{2}x^T Qx + c^T x - \lambda^T x + \mu^T (b - Ax) \\ &= \frac{1}{2}x^T Qx + x^T (c - A^T \mu - \lambda) + b^T \mu. \end{aligned}$$

where $\lambda \geq 0$, $\lambda \in R^n$, $\mu \in R^m$ are Lagrange multipliers. Solving

$$0 = \nabla_x L(x, \lambda, \mu) = Qx + c - A^T \mu - \lambda,$$

yields

$$x = Q^{+}(A^T\mu + \lambda - c), \tag{1}$$
$$\iff Qx = A^T\mu + \lambda - c. \tag{2}$$

We can substitute (1) with $L(x, \lambda, \mu)$, this gives us the dual objective. Hence,

$$\phi(\lambda, \mu) = \begin{cases} B^T\mu - \frac{1}{2}(A^T\mu + \lambda - c)Q^{+}(A^T\mu + \lambda - c) & \text{if } A^T\mu + \lambda - c \in \text{Im}(Q) \\ -\infty & \text{otherwise} \end{cases}$$

This implies that the dual problem is

$$\begin{aligned} &\max\ b^T\mu - \frac{1}{2}(A^T\mu + \lambda - c)Q^{+}(A^T\mu - c) \\ &\text{s.t. } A^T\mu + \lambda - c \in \text{Im}(Q) \\ &\qquad \lambda \geq 0. \end{aligned}$$

From (2), we also write the dual problem as follows

$$\begin{aligned} &\max\ b^T y - \frac{1}{2}x^T Qx \\ &\text{s.t. } Qx = A^T y + \lambda - c \\ &\qquad \lambda \geq 0. \end{aligned} \tag{D}$$

This is uncommon if we are interested in dualising the problem, because there are both primal and dual variables in this formulation. However, this is the case when we reformulate QP as a conic programming problem, which we will do later.

If we were to assume that Q is positive definite, i.e. $Q \in S^n_{++}$, then that implies that $Q^{+} = Q^{-1}$. Hence, our dual problem would then be

$$\begin{cases} \max\limits_{\lambda,\, \mu}\ b^T\mu - \frac{1}{2}(A^T\mu + \lambda - c)Q^{-1}(A^T\mu + \Lambda - c) \\ \text{s.t. } \lambda \geq 0. \end{cases}$$

If we assume that Slater's condition are fulfilled, then strong duality gives us;
(i) The primal problem of QP is infeasible if and only if

$$\exists y:\ y^T A \leq 0 \text{ and } b^T y > 0$$

(ii) The dual problem is infeasible if and only if

$$\exists x \geq 0:\ Ax = 0,\ Qx = 0 \text{ and } c^T x < 0.$$

2.5 KKT conditions and Slater's conditions

The KKT and Slater's conditions are formulated in different ways depending on if the MP is with or without equality constraint. Consider a general mathematical programming in the form

$$\begin{aligned} &\min\ f(x) \\ &\text{s.t. } g(x) \leq 0 \\ &\qquad h(x) = 0, \end{aligned} \tag{MP}$$

where

$$\begin{aligned}
&f:\ R^n \to R\\
&g(x) \leq 0 \iff g_i(x) \leq 0,\ i = 1, ..., m\\
&h(x) = 0, \iff h_j(x) = 0,\ j = 1, ..., l\\
&x \in X \subseteq R^n.
\end{aligned}$$

We assume that X is an open set and for the sake of simplicity we will also assume that f and g_i are continuous and differentiable in the optimum x^*. Let us now consider a system that only has inequality constraints, i.e.

$$\begin{aligned}
\min\ & f(x)\\
\text{s.t. } & g(x) \leq 0\\
& x \in X.
\end{aligned} \tag{$MP_{\leq}$}$$

If we let

$$I = \{i :\ g_i(x^*) < 0\},$$

then we can give Slater's condition for a $(MP_{\leq})$ and (MP), i.e.

Slater's condition for $MP_{\leq}$:
The set X is open, each g_i is pseudo convex for $i \in I$ at a feasible point x^* and each g_i for $i \notin I$ is continuous at x^* and there is an $x \in X$ such that $g_i(x) < 0$ for all $i \in I$ [18, 223].

Slater's condition for MP:
If we now observe for a general (MP) then we have that; the set X is open, each g_i for $i \in I$ is pseudoconvex at x^*, each g_i for $i \notin I$ is continuous at x^*, each h_i for $i = 1, ..., l$ is quasiconvex, quasiconcave and continuosly differentiable at x^*, and $\nabla h_i(x^*)$, for $i = 1, ..., l$ are liniearly independent. Furthermore, there is an $x \in X$ such that $g_i(x) < 0$ for all $i \in I$ and $h_i(x) = 0$ for all $i = 1, ..., l$.

We can now summarise these two to the following

Slater's condition for a convex programming problem:

$$\begin{aligned}
\min\ & f(x)\\
\text{s.t. } & g(x) \leq 0\\
& Ax = b\\
& x \in X.
\end{aligned}$$

The set X is open, f, g are differentiable over X, and A is a $m \times n$ matrix where $m \leq n$ has full row rank.

We say that the problem satisfies the Slater's conditions if

$$\exists x_0 \in X :\ g_i(x_0) < 0,\ i = 1, ..., m,\ Ax_0 = b,$$

which guarantees strong duality [4, page 283].

Theorem 2.5. *Consider the problem (MP). Suppose (MP) (the primal problem) and its dual have equal optimal value, which are attained at x^* and (λ^*, μ^*). Then*

$$\lambda_i^* g_i(x^*) = 0,\ i = 1, ..., m.$$

Proof. Since the duality gap is zero, we have

$$f(x^*) = \phi(\lambda^*, \mu^*)$$

where $\phi(\lambda, \mu)$ is the dual objective function. By definition

$$\phi(\lambda, \mu) = \min_{x \in X} (f(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{l} \mu_j h_j(x))$$

Thus

$$\phi(\lambda^*, \mu^*) = \min_{x \in X} (f(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{j=1}^{l} \mu_j^* h_j(x))$$

By definition of minimum we have

$$\begin{aligned} &\min_{x \in X} (f(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{j=1}^{l} \mu_j^* h_j(x)) \\ &\leq f(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{j=1}^{l} \mu_j^* l_j(x^*). \end{aligned}$$

Now x^* is optimum of (MP) we have $g(x^*) \leq 0$, $h(x^*) = 0$, so altogether we have a chain of inequalities

$$\begin{aligned} f(x^*) \leq \phi(\lambda^*, \mu^*) \leq f(x^*) &+ \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{j=1}^{l} \mu_j^* l_j(x^*) \\ \leq f(x^*). \end{aligned}$$

Hence, they are all equal, i.e.

$$\begin{aligned} f(x^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{j=1}^{l} \mu_j^* h_j(x^*) \\ \iff \sum_{i=1}^{m} \lambda_i^* g_i(x^*) = 0, \text{ since } h_j(x^*) = 0. \end{aligned}$$

Now $g_i(x^*) \leq 0$, $\lambda_i^* \geq 0$ and the sum is eqial to zero implies

$$\lambda_i^* g_i(x^*) = 0,\ i = 1, ..., m.$$

□

Theorem 2.6. *Under the same condition as the ones in the previous theorem, if (λ^*, μ^*) is optimal of the dual problem then*

$$x^* = arg \min_x L(x, \lambda^*, \mu^*)$$

where

$$L(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x). \qquad \text{(Lagragian function)}$$

Proof. We have from the proof of the previous theorem

$$\begin{aligned} &\min_x \{f(x) + \sum_{i=1}^{m} \lambda_i^* g_i(x) + \sum_{j=1}^{l} \mu_j^* h_j(x)\} \\ =&f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{j=1}^{l} \mu_j^* h_j(x^*) \end{aligned}$$

which means x^* is the minimiser of the Lagrange function $L(x, \lambda^*, \mu^*)$ □

If we would now assume differentiability, then $x^* = \arg \min_x L(x, \lambda^*, \mu^*)$ implies

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla q_n(x^*) + \sum_{j=1}^{l} \mu_j^* h_J(x^*) = 0.$$

This leads to the following necessary condition for existence of primal and dual solution. Next theorem is an immediate consequence of the previous result.

Theorem 2.7. *(The KKT necessary condition). Assume that f, g_i $(i = 1, ..., m)$, h_j $(j = 1, ..., l)$ in (MP) are differentiable over the open set X. Let x^* and (λ^*, μ^*) be any primal and dual optimal solution with no duality gap. Then the KKT conditions holds. That is,*

$$\begin{aligned} &(i) g_i(x^*) \leq 0,\ i = 1, ..., m,\ h_j(x) = 0,\ j = 1, ..., l && \text{(primal feasibility)} \\ &(ii) \lambda_i^* \geq 0,\ i = 1, ..., m && \text{(dual feasibility)} \\ &(iii) \lambda_i^* g_i(x^*) = 0,\ i = 1, ..., m && \text{(complementary slackness)} \\ &(iv) \nabla f(x^*) + \sum_{i=1} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{l} \mu_j^* \nabla h_j(x^*) = 0. && \text{(Lagrange stationarity)} \end{aligned}$$

Theorem 2.8. *(KKT sufficient theorem) Consider the optimisation problem (MP). Assume that; X is an open and convex set, f, $\{g_i\}_{i=1}^m$ and $\{h_j\}_{j=1}^l$ are differentiable and convex on X. If (x^*, λ^*, μ^*) is a KKT point, i.e. it satisfies the KKT conditions. Then x^* and (λ^*, μ^*) are primal and dual optimal with zero duality gap.*

The proof is rather technical so we give a proof of a little special case where $h_i(x)$ are affine functions, which is indeed what we need in this test. This affect the optimisation problem so that it becomes categorized as a convex optimisation problem.

Proof. Since (x^*, λ^*, μ^*) is a KKT point we have that x^*, (λ^*, μ^*) are primal and dual feasible. Note that $\{h_j(x)\}_{j=1}^l$ are affine plus convex and

$$\sum_{j=1}^{l} \mu_j h_j(x)$$

is also convex. This together with $\lambda^* \geq 0$ (KKT condition (ii)) gives that $L(x, \lambda^*, \mu^*)$ is convex in X (positive combination of convex functions). So, the Lagrange stationarity at x^* implies x^* is a minimiser of $L(x, \lambda^*, \mu^*)$, which is sufficiently optimal due to convex of $L(x, \lambda^*, \mu^*)$. Hence, by the KKT condition's

$$\begin{aligned}\phi(\lambda^*, \mu^*) &= L(x^*, \lambda^*, \mu^*) = f(x^*) + \sum_{i=1}^{m} \lambda_i g_i(x^*) + \sum_{j=1}^{l} \mu_j^* h_j(x^*) \\ &= f(x^*).\end{aligned}$$

In other word, the duality gap is 0. Thus x^* and (λ^*, μ^*) are primal and dual optimal. □

Theorem 2.9. *Assume that the optimisation problem (MP) is; convex and f, $\{g_i\}_{i=1}^m$, $\{h_j\}_{j=1}^n$ are differentiable and convex on the open convex set. If (MP) satisfies the Slater's condition, then the KKT condition are sufficient and necessary for optimality.*

3 Conic programming

3.1 Formalising a CP

We will in this section show how the idea of solving LP problems can be generalised to solving a fairly large class of nonlinear problems. Basically we transform the polyhedral cone in LP to a convex cone formed by the nonlinear constraints. Note that the nonlinear objective function is not essential, because it can be reformulated by minimising t with an additional constraint $f(x) < t$ where $t \in R$.

Recall that the LP problem has a linear objective function and the constraint $Ax \geq b$, which is $Ax - b \geq 0$. So we have a natural ordering to get this "$\geq$" (component wise). That means $Ax - b$ is in the cone R^n_+ with the natural "$\geq$". Now, we will pass to nonlinear constraints to a well-order cone, which has similar properties as R^n_+. That is, we need to find "$\geq$" in terms of the cone. We can now formalise a standard CP as the following

$$\begin{aligned} &\min c^T x \\ &\text{s.t. } Ax\, \Omega_K\, b \end{aligned} \tag{CP}$$

where Ω_K is "$>$" or "$\geq$" such that it satisfies the basic properties of a standard ordering and $K \subset R^n$ is a convex cone. For it to be a standard ordering (like vector inequality) then Ω_K must satisfy

$$\begin{aligned} (i)&\,\text{reflexivity, } a \geq a \\ (ii)&\,\text{anti-symmetry, if } a \geq b \text{and } b \geq a \text{ then } b = a \\ (iii)&\,\text{transitivity, if } a \geq b,\ b \geq c \text{ then } a \geq c \\ (iv)&\,\text{Homogeneity, if } a \geq b \text{ then } \lambda a \geq \lambda b \\ (v)&\,\text{Additive, if } a \geq\ b,\ c \geq d \text{ then } a + c \geq b + d. \end{aligned}$$

where $a, b, c, d \in R^m$. Usually we only refer that the order is good if its inner product satisfies (i) to (v). This is the case if the set K can be formalised such that

$$K := \{a \in E : a \succeq 0\},$$

where E is the Euclidean space. We can now precise Ω_K as the following

$$Ax\ \Omega_K\ b := \{x : Ax\, \Omega b, x \in K\}.$$

We can now propose a proposition.

Proposition 1. *If K is of good order then K must be a pointed cone, i.e.*

$$\begin{aligned} &(i)\, a', a \in K \implies a + a' \in K && \text{(K is nonempty and closed under addition)} \\ &(ii)\, a \in K,\ \lambda \in R,\ \lambda \geq 0, \implies \lambda a \in K && \text{(K is a conic set)} \\ &(iii)\, a \in K,\ -a \in K \implies a = 0, && \text{(K is pointed)} \end{aligned}$$

Proof. K is clearly nonempty, just take any positive point in E. If $a \geq_K 0$ and $b \geq_K 0$ then $a + b \geq_K 0$, thus (i). The proof for (ii) is analogously to the one in (i). And we can clearly see for (iii) that the only point that satisfies $a \in K$, $-a \in K$ is 0. □

We will in the next section look at some particular cones that frequently occurs in CP.

3.2 Some interesting cones

We will now observe some special cones; the positive orthant, the Lorentz cone and the semi-definit cone. These cones frequently occur and they are proper cones.

Definition 3.1. *(Proper cone) A cone K is a proper cone if it is convex, closed, pointed and has a nonempty interior.*

Definition 3.2. *(Self dual) Let K be a nonempty subset in R^n. The K^* is a dual cone if*

$$K^* = \{y : \langle x, y\rangle \geq o, \text{ for all } x \in K\}$$

Assume K is a nonempty convex cone. If $K^* = K$ then K is called self dual.

3.2.1 The positive orthant

The positive orthant is when K can be formalised as

$$K = R_+^m \subset R^n$$

Proposition 2. *The positive orthant is a proper cone.*

Proof. Since the positive orthant contains all of its line segments implies that it has to be convex and R_+^m is clearly nonempty. The cone is closed because the boundary is included in R_+^m and clearly is 0 the only point which satisfies $-a, a \in R_+^m$. □

Proposition 3. *The positive orthant is self dual.*

Proof. The dual cone to K is such that

$$K_* = \{a : a^T x \geq 0, \forall\, x \in K\}.$$

We can clearly see that if K is the positive orthant then K_* must be so as well. □

3.2.2 Lorentz cone

The Lorentz cone (or sometimes know as the second order cone or the ice cream cone) is the set K that can be formalised as

$$K = L^m = \left\{ x = (x_1, ..., x_m)^T \in R^m : x_m \geq \sqrt{\sum_{i=1}^{m-1} x_i^2} \right\}$$

Lemma 3.1. *The Lorentz cone is a proper cone*

Proof. We are going to separate this proof into four stages. Whereas, in the first stage we prove that the set is convex, in the second stage that the set is closed, the third stage that the interior is none empty and in the fourth stage that the set is pointed.

i) We can see that convexity follows directly from triangle inequality. For any $x, y \in L$ let

$$\tilde{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_{m-1} \end{pmatrix}, \tilde{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_{m-1} \end{pmatrix}.$$

We can now apply the triangle inequality:

$$\begin{aligned} ||\lambda x + (1-\lambda)y||_2 &\leq ||\lambda x||_2 + ||(1-\lambda)y||_2 \\ &= \lambda||\tilde{x}||_2 + (1-\lambda)||\tilde{y}||_2 \\ &\leq \lambda x_m + (1-\lambda)y_m. \end{aligned}$$

ii) L is closed because the boundary is included in L

iii) int(L)$\neq \emptyset$.
We can see that $(0, 0, 1) \in \text{int}(L)$, hence the interior of L cannot be empty.

iv) L is a pointed cone if $L \cap (-L) = \{0\}$, which is the case since all elements in L are greater then or equal to zero.

Altogether gives that L is a proper cone. $\square$

Lemma 3.2. *The Lorentz cone is self dual, $L = L^*$*

Proof. We prove that $L \subseteq L^*$ and $L^* \subseteq L$.

$L \subseteq L^*$:
Hence, if $x_i y \in L$ then Cauchys-Schwarz inequality gives

$$x^T y = x_n y_n + \sum_{i=1}^{n} x_i y_i \geq x_n y_n - \sqrt{\sum_{i=1}^{n-1} x_i^2} \sqrt{\sum_{i=1}^{n-1} y_i^2} \geq 0.$$

The last inequality clearly holds because $x, y \in L$. This inequality tells us that

$$\begin{aligned} &x_n^2 \geq \sum_{i=1}^{n-1} x_i^2, \quad -y_n^2 \geq \sum_{i=1}^{n-1} y_i^2 \\ &x_n \geq 0,\ y_n \geq 0 \end{aligned}$$

Thus the inner product for the vector spaces are the same, i.e. if $x \in L$ then $x \in L^* \iff L \subseteq L^*$.

$L^* \subseteq L$:
Assume $y \in L^*$, i.e. $y^T x \geq 0\ \forall x \in L$. If $y = (y_1, ..., y_{n-1}, y_n - 1) = 0$, then consider $x_1 = ... = x_{n-1} = 0$ and $x_n = 1$. Then

$$y^T x \geq 0 \iff y_n \geq 0 \iff y_n^2 \geq \sum_{i=1}^{n-1} y_i^2 \iff y \in L.$$

Otherwise let $x_n = \sum_{i=1}^{n-1} y_i^2$ and $x_i = -y_i$ for $i = 1, ..., m$. Then

$$y^T x \geq 0 \iff y_n \sqrt{\sum_{i=1}^{n-1} y_i^2} - \sum_{i=1}^{n-1} y_i^2 \geq 0$$

$$\iff y_n^2 \geq \sum_{i=1}^{n-1} y_i^2,\ y_n \geq 0 \iff y \in L.$$

Thus the inner product for the vector spaces are the same, i.e. if $y \in L^*$ then $y \in L \iff L^* \subseteq L$. □

3.2.3 Semi-definite cones

Let

$$S^n := \{A \in R^{n\times n} : A^T = A\}, \qquad \text{(symmetric set)}$$

$$S_+^n = \{A \in S^n : A \succeq 0\} = \{A \in S^n : x^T Ax \geq 0 \forall x \in R^m\} \quad \text{(positive semi-definite)}$$

$$S_{++}^n = \{A \in S_+^n : A \succ 0\} \qquad \text{(positive definite)}$$

Clearly S_+, S_{++} are cones. We will now define an inner product on S^n as

$$\langle A, B\rangle = Tr(AB) = \sum_{i,j} A_{i,j} B_{i,j},$$

which is the so-called Frobenius inner product, where $Tr(\cdot)$ is the the trace of the matrix $\cdot$.

Proposition 4. *S_+^n is a proper cone.*

Proof. (i) S_+^n is closed and convex.

We can rewrite S_+^n as an intersection of H_x, i.e.

$$S_+^n = \underset{x\in R^n}{\cap} \{A \in S^n : x^T Ax \geq 0\} = \underset{x\in R^n}{\cap} H_x$$

where $H_x := \{A \in S^n : x^T Ax \geq 0\}$.

Now H_x is closed and convex. Thus must S_+^n be convex and closed as well.

(ii) S_+^n are pointed i.e. we claim that

$$(S_+^n) \cap (-S_+^n) = \{0\}.$$

If there exists an A such that $A \in (S_+^n) \cap (-S_+^n)$ then the eigenvalues of A must equal zero, resulting that $A = 0$.

(iii) int(S_+^n)=$S_{++}^n \neq \emptyset$

We can prove this by shoving that int(S_+^n) $\subseteq S_{++}^n$ and vice versa and then that there exist an interior point in S_{++}^n. Hence, we prove this by a matrix 2-norm.

int(S_+^n) $\subseteq S_{++}^n$:

Let $A \in$ int(S_+^n), then $\exists\epsilon > 0$ (that is sufficietnly small) such that $||A - X||_2 \leq \epsilon$. This implies that $x \in S_+^n$.

Let $X = A - \epsilon I$, where I is the $n \times n$- identity matrix. We then have

$$||A - X||_2 = ||\epsilon I||_2 = \epsilon \implies X = A - \epsilon I \in S_+^n.$$

Since all the eigenvalues are nonnegative and the eigenvalues of $A - \epsilon I$ are $\lambda - \epsilon$, which is nonnegative, where λ is the eigenvalue of A. From this it follows that

$$\lambda_i \geq \epsilon > 0,$$

where λ_i is the eigenvalue of A, $i = 1, ..., n$ counted with multiplicities

$$\implies A \in S_{++}^n.$$

$S_{++}^n \subseteq \text{int}(S_+^n)$:

Let $A \in S_{++}^n$. Then the minimum eigenvalue of A $\lambda_{\min} > 0$. Consider the following set

$$B := \{M \in S^n : ||M - A||_2 \leq \frac{\lambda_{\min}}{2}\}.$$

We will show that $B \subseteq S_{++}^n$ (i.e. A is interior point of S_+^n). We have:

$$||M - A|| \leq \frac{\lambda_{\min}}{2} \iff |x^T(M - A)x| \leq \frac{\lambda_{\min}}{2}, \forall x : ||x|| = 1$$

then

$$\begin{aligned} &\implies x^T M x \geq x^T A x - \frac{\lambda_{\min}}{2} \geq \frac{\lambda_{\min}}{2} > 0 \\ &\implies M \in S_{++}^n \implies B \subseteq S_{++}^n \end{aligned}$$

□

Proposition 5. *S_+^n is self dual.*

Proof. S_+^n is self dual, which is equivalent to $S_+^n = (S_+^n)^*$. (Under scalar product Tr(AB)).

$S_+^n \subseteq (S_+^n)^*$:

Assume $B \in S_+^n$, then by spectral theorem,

$$B = \sum_{i=1}^{n} \lambda_i v_i v_i^T,\ \lambda_i \geq 0$$

and $v_i \perp v_j$ $(j \neq i)$, $||v_i|| = 1 \forall i$.

Now $A \in S^n_+$, i.e. all $v^T Av > 0$ for all no zero vector v. Using the cyclic property of the trace we have

$$\begin{aligned}
Tr(AB) &= Tr(\sum_{i=1}^{n} \lambda_i A v_i v_i^T) = \sum_{i=1}^{n} Tr(\lambda_i A v_i - v_i^T) \\
&= \sum_{i=1}^{n} \lambda_i Tr(A v_i v_i^T) \\
&= \sum_{i=1}^{n} \lambda_i Tr(v_i^T A v_i) \\
&= \sum_{i=1}^{n} \lambda_i v_i^T A v_i,\ v_i^T A v_i \geq 0 \\
&\geq 0 \\
&\implies A \in (S^n_+)^*.
\end{aligned}$$

$(S^n_+)^* \supseteq S^n_+$:
Assume that

$$B \in S^n : Tr(AB) > 0\ \forall A \in S^n_+.$$

We will now show that $B \geq 0$. Take $A = xx^T\ \forall x \in R^n$ we get $Tr(xx^T B) = x^T Bx \geq 0\ \forall x \in R^n \implies B \in S^n_+$. $\square$

3.3 Some common conic programs

It is clear that an LP problem is a CP since we introduce the CP by an analog of LP where the cone K is R^n_+ and the ordering is the componentwise inequality. Now we show that there are other optimization problems.

3.3.1 Second order programming

We start with a reformulation of the QP discussed in section 2.4. Note that

$$Q \in S^n_+ \iff \exists F \in R^{k \times n} \text{ such that } Q = F^T F$$

(which is of most interest when $k \ll n$. Then $x^T Qx = ||Fx||_2^2$. Now the conic reformulation is

$$\begin{aligned}
\min\ & r + c^T x \\
\text{s.t. } & Ax = b \\
& x \geq 0 \\
& (1, r, F(x)) \in L_r^{k+2}
\end{aligned} \tag{CP_{QP}}$$

where L_r^{k+2} is the rotated cone defined as

$$\begin{aligned}
&L_r^{k+2} = \{x \in R^{k+2} :\ 2x_1x_2 \geq x_3^2 + ... + x_n^2, x_1 \geq 0, x_2 \geq 0\} \\
&x \in L^{k+2} \iff T_{k+2}x \in L_r^{k+2}
\end{aligned}$$

with

$$T_{k+2} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & I_n \end{pmatrix}$$

The dual problem of (CP_{QP}) is

$$\begin{aligned} &\max\ b^T y - u \\ &\text{s.t.}\ -F^T v = A^T y - c + s \\ &\quad s \geq 0 \\ &\quad (u, 1, v) \in L_r^{k+2} \end{aligned}$$

Now we are in the position to justify why we had an uncommon version of the dual to the QP earlier. In an optimal primal dual solution we have $r = \frac{1}{2}Fx^2$. Hence, the complementary slackness for the cone L_r^{k+2} demands $v = Fx$ and $-FF^T v = Qx$ and $u = \frac{1}{2}|v|^2 = \frac{1}{2}x^T Qx$. This is why the dual has both dual and primal variables.

A second order conic problem is a conic problem for which the cone K is a direct product of ice-cream cones:

$$\begin{aligned} K &= L^{m_1} \times ... \times L^{m_k}. \\ &= \left\{ y = \begin{pmatrix} y[1] \\ \dots \\ y[k] \end{pmatrix} : y[i] \in L^{m_i},\ i = 1, ..., k \right\}. \end{aligned}$$

Hence we can substitute K with L^m and we can write the data matrix as:

$$[A_i; b_i] = \begin{bmatrix} D_i & d_i \\ p_i^T & q_i \end{bmatrix},$$

where D_i is of the size of $(m_i - 1) \times \dim x$. We can write this problem as

$$\min_x \{c^T x :\ ||D_i x - d_i||_2 \leq p_i^T x - q_i,\ i = 1, ..., k\}.$$

Second order cone program (SOCP) is a generalisation of linear and quadratic programming that allows affine combinations of variables to be constrained inside a special convex set K. Wheres the set K is a second order cone whiles the constraints and the objective are an affine combination of variables. A way to prove that K_n is a convex set is by expressing it as the intersection of a finite or infinite number of halfspaces, i.e.

$$K_n = \cap\{(x, t), x \in R^n \times R : x^T u \leq t\},$$

which is a cone for any $z \in K_n$ if it holds that $\alpha z \in K_n\ \forall \alpha \geq 0$.

A SOCP in standard form is

$$\begin{aligned} &\min_{x \in R^n} c^T x \\ &\text{s.t.}\ ||A_i x + b_i||_2 \leq c_i^T x + d_i,\ i = 1, ..., m. \end{aligned}$$

An equivalent form is the standard conic form

$$\min_{x \in R^n} c^T x$$
$$\text{s.t. } (a_i x + b_i,\, c_i^T x + d_i) \in K_m,\ i = 1, ..., m.$$

SOCP duality

$$p^* = \min_x c^T x$$
$$\text{s.t. } ||A_i x + b_i||_2 \leq c_i^T x + d_i,\ i = 1, ..., m$$

we have

$$\begin{aligned} p^* &= \min_x \max_{\lambda \geq 0} c^T x + \sum_{i=1}^m \lambda_i(||A_i x + b_i||_2 - c_i^T x - d_i) \\ &= \min_x \max_{||u_i||_2 \leq \lambda_i,\, i=1,..,m} c^T x + \sum_{i=1}^m (u_i^T (A_i x + b_i) - \lambda_i(c_i^T + d_i)) \end{aligned}$$

Applying the maxmin inequality we obtain $p^* \geq d^*$

$$d^* = \max_{||u_i||_2 \leq \lambda_i,\, i=1,..,m} \min_x c^T x + \sum_{i=1}^m (u_i^T (A_i + b_i) - \lambda_i(c_i^T x + d_i)).$$

Solving x we get the dual problem, and applying Slater's condition we get the dual theorem.

3.3.2 Semi-definite program

A semi-definite program (SDP) is a convex optimisation problem where the objective is subjected to constraints, which are positive semi-definit. And, positive semi-definite constraints implies that the symmetric matrices depends affinely on a vector of variables $x \in R^n$.

Given $F_i \in S^m$, $i = 0, ..., n$ is a semi-definite program, we can usually reformulate to the following:

$$p^* = \min_{x \in R^n} c^T x$$
$$\text{s.t. } F(x) \succeq 0,$$

where $F(x) = F_0 + \sum_{i=1}^n x_i F_i$, $F(x) \succeq 0$ iff $\lambda_{\min}(F(x)) \geq 0$.

If p^* and d^* are attained, we say that the Lagrangian formula $\mathcal{L}$ has a saddle point at the primal and dual, which we denote as (x^*, λ^*, v^*).

We will use the letters $\mathcal{A}, \mathcal{B}$ etc. to indicate that it is a linear mapping.

Hence, a semi-definite program is usually written:

$$\min_x \{c^T x : \mathcal{A}x - B \geq_{S_+^m} 0\}.$$

We note that

$$\mathcal{A}x = \sum_{j=1}^{n} x_j A_j, \quad x = (x_1, ..., x_n)^T \in R^n$$

$$\min_x \{c^T x : x_1 A_1 + ... + x_n A_n - B \succeq 0\}.$$

The dual of the semidefinit program is

$$\max_\Lambda \{\langle B, \Lambda \rangle \equiv \mathrm{Tr}(B\Lambda) : \mathrm{Tr}(A_i \Lambda) = c_i \; i = 1, ..., n; \Lambda \succeq 0\}.$$

For us to apply strong duality the program has to be positive definite, which it is when

$$\langle \Lambda, \mathcal{A}x - B \rangle = 0.$$

4 Robust linear programming

4.1 Defining Robust linear programming

We will in a similar manner as described in [4] go through how we can define a robust linear program, RLP. As the name hints, robust programming is a method where we analyse how to make the problem robust against uncertainty, i.e. we bound our data within a set such that the perturbation is within this set. Without taking the perturbation into account, then one can get a direct contradicting solution. We can illustrate this with an example:
A drunk person tries to walk home in the middle of the highway, will he get home safely? Our nominal values would indicate that there exits a narrow path that could allow him to get home safely. However, taking the perturbation in account, then he or she would most likely not be able to maintain to walk on the straight narrow path. Thus, our nominal values indicates that he or she will get home safely whiles when we take perturbation in account then he or she will not get home safely. This is a silly example, but it still emphasises that a rather small perturbations can have a big impact on the solution.

Lets now reformulate our original LP problem such that every variable is a function of an uncertainty $\mathcal{U}$, i.e.

$$\left\{\min_{x} \{c^T x + d : Ax \leq b\}\right\}_{(c,d,A,b)\in\mathcal{U}} \tag{ULP}$$

where $\mathcal{U} \subset R^{(m+1)\times(n+1)}$. We now want to show how the robust counterpart can be computationally tractable if we make mild convexity and compatibility assumptions for the uncertainty set U. As a start we can can formulate our uncertainty set $\mathcal{U}$ as a sum of the nominal values and their respective perturbation. Hence,

$$\mathcal{U} = \left\{(c^T, d, A, b) = (c_0^T, d_0, A_0, b_0) + \sum_{l=1}^{L} \zeta_l (c_l^T, d_l, A_l, b_l) : \zeta \in \mathcal{Z} \subset R^L\right\}.$$

Where $\mathcal{Z}$ is our perturbation set, and index zero indicates that it is our nominal value and ζ is our perturbation vector. Thus, our solution will be robust feasible if it satisfies

$$Ax \leq b \quad \forall (c, d, A, b) \in \mathcal{U}.$$

We now want to give a definition for the robust value, so that we can formalise the robust counterpart. If we denote the robust value as $\hat{c}(x)$ for a feasible solution x, then we can give the following definition.

Definition 4.1. *If $\hat{c}(x)$ is the uncertain value of $c(x)$ for an uncertain linear program then we can ensure robustness by taking the largest value that $c^T x + d$ can take over all the realisations of the data in the uncertainty set, i.e.*

$$\hat{c}(x) = \underset{(c,d,A,b)\in\mathcal{U}}{sup} [c^T x + d].$$

We can formulate the robust counterpart, which is equivalent to the uncertain LP

$$\min_{x,t} \left\{t : c^T x - t \leq -d,\ Ax \leq b,\ \forall (c, d, A, b) \in \mathcal{U}\right\}. \tag{RC}$$

If we now look at two cases; (i) where the objective is certain but the data remains uncertain, and (ii) where the objective is uncertain but the data is certain.

(i) Assuming that we have a certain objective then we can rewrite our robust counterpart:

$$\min_x \{c^T x + d : Ax \leq b, \forall (A,b) \in \mathcal{U}. \tag{RC}$$

We can now replace all of its original constraints with its robust counterparts. We do this by extending the uncertainty set $\mathcal{U}$ to $\hat{\mathcal{U}}$, where $\hat{\mathcal{U}}$ is a direct product of all the robust counterparts $\mathcal{U}_i$. Hence, we project our uncertainty set $\mathcal{U}$ onto our data space. This leaves us with

$$\begin{aligned}
&(Ax)_i \leq b_i \iff a_i^T x \leq b_i,\ \forall [a_i; b_i] \in \mathcal{U}_i \\
&\mathcal{U}_i = \{[a_i; b_i] : [A, b] \in \mathcal{U} \\
&\hat{\mathcal{U}} = \mathcal{U}_1 \times ... \times \mathcal{U}_n.
\end{aligned}$$

Notice that we can extend $\mathcal{U}_i$ to its convex hull, because if x is a robust feasible solution then it will remain feasible if we do this extension. If $[\overline{a}_i; \overline{b}_i] \in \text{conv}(\mathcal{U}_i$:

$$\begin{aligned}
[\overline{a}_i; \overline{b}_i] &= \sum_{j=1}^{J} \lambda_j [a_i^j; b_i^j] \\
\implies \overline{a}_i^T x &= \sum_{j=1}^{J} \lambda_j [a_i^j]^T x \leq \sum_j \lambda_j b_i^j = \overline{b}_i
\end{aligned}$$

where $[a_i^j; b_i^j] \in \mathcal{U}_i$, $\lambda \geq 0$ such that $\sum_j \lambda_j = 1$. By similar argument will x remain robust feasible if we extend to its closure set. Hence, if we have a certain objective function then we can extend $\mathcal{U}$ to its convex and closed robust counterpart. Thus we can rewrite $\mathcal{U}_i$ to the following

$$\begin{aligned}
\mathcal{U}_i &:= \{[a_i; b_i] = [a_i^0; b_i^0] + \sum_{l=1}^{L} \zeta_l [a_i^l; b_i^l] : \zeta \in \mathcal{Z}_i\} \\
\mathcal{U} &= \{[a; b] = [a^0; b^0] + \sum_{l=1}^{L} \zeta_l [a^l; b^l] : \zeta \in \mathcal{Z}\},
\end{aligned}$$

where $\mathcal{Z}$ is a closed convex perturbation set. Thus, all a and b that fulfill the uncertainty requirements must give a tractable representation of the robust counterpart. Lets now observe how we can reformulate our robust counterpart given some popular uncertainty sets $\mathcal{Z}$.

4.2 Tractable under ellipsoid uncertainty

Assuming that our objective function is certain, we can then formulate our uncertainty set so that it is an ellipsoid. However, this is equivalent to a ball, by the same reasoning as before. Thus, we assume that the uncertainty set is a ball with a radius of Ω and that it is centered at the origin, i.e.

$$Z = Z^E = \{\zeta \in R^L : ||\zeta||_2 \leq \Omega\}.$$

Now, for an uncertain LP we have for any $x \in R^n$ the following must hold

$$
\begin{aligned}
&\sum_{l=1}^{L}(\zeta_l a^l)x + [a^0]^T x \leq \sum_{l=1}^{L}(\zeta_l b^l) + b^0 \\
&\sum_{l=1}^{L}(\zeta_l (a^l)x + b^l) \leq b^0 - [a^0]^T x.
\end{aligned}
$$

In the last step we gathered all of our uncertainty vectors on the left side and the certain values on the right side. Notice that $b - A^T x$ is an upper bound, which means that the maximum value of left side will still be less than or equal to the value on the right side. Thus, we substitute our left side with its maximum.

$$
\begin{aligned}
&\max_{||\zeta||_2 \leq \Omega} \{\sum_{l=1}^{L} \zeta_l (a^l)x + b^l)\} \leq b - [a^0]^T x \\
\iff &\Omega \sqrt{\sum_{l=1}^{L}((a^l)x + b^l)^2} \leq b^0 - [a^0]^T x \\
\iff &w \leq b^0 - [a^0]^T x
\end{aligned}
$$

We substituted the greatest value with w. Since our sets consists of closed and convex sets means that we can freely apply strong duality. Hence,

$$
\min_{\lambda \in R^L} \{w^T \lambda : \lambda = (\sum_{l=1}^{L} a^l)x - b, \, \lambda \geq 0\} \leq b^0 - [a^0]^T x
$$

We can now formalise our robust counterpart to our uncertain LP

$$
\begin{aligned}
\min_{x \in R^n, \, \lambda \in R^L} \; & c^T x \\
\text{s.t.} \; & w^T \lambda_i \leq b_i - a_i^T x \\
& \lambda_i = A_i^T x - b_i \\
& \lambda_i \geq 0,
\end{aligned} \tag{RCLP}
$$

wheres the number of decision variables is $n + mL$ and the number of constraints is $(1 + k + 1)m$. This is clearly a $SOCP$ problem, hence if we have a certain objective function and an uncertainty set of an ellipsoid then we can transform our uncertain LP to a $SOCP$ problem.

5 Robust programming applied to 2-person games

5.1 Formalising a bimatrix game

As the name might suggest we call it a bimatrix game, because it consists of two players and we can represent the players payoffs in a matrix. When we formalise such a game we first need to find out what pure strategies each player has access too, by "pure" we emphasise that the strategy is strict. For example under z circumstance player I will make the move q. Thus, we can form two strategy sets of the available strategy for respective player, call them S and T. When a strategy is presented against an opponents strategy this result is a payoff to the respective player. Hence, if we assume that S consists of m strategies and T of n strategies then we can formalise both players payoffs as $n \times m$ matrices. We can now present this in mathematical terms

Player I	Player II
Payoff: $A = \begin{pmatrix} a_{(s_1,t_1)} & ... & a_{(s_1,t_n)} \\ ... & ... & ... \\ a_{(s_n,t_1)} & ... & a_{(s_n,t_m)} \end{pmatrix}$	Payoff: $B = \begin{pmatrix} b_{(s_1,t_1)} & ... & b_{(s_1,t_n)} \\ ... & ... & ... \\ b_{(s_n,t_1)} & ... & b_{(s_n,t_m)} \end{pmatrix}$
Strategy set: S	Strategy set: T
strategy: s_i	Strategy: t_i

We can denote this game, G, as the set of (A, B, S, T). The above is a pure strategy game, but we can generalize it into a mixed strategy game by introducing two probability vectors x and y that determines the likelihood that a player chooses each pure strategy. Notice, that in a pure strategy game x and y must then be e_i and e_j, i.e. where e_i denotes unit vector and the index denotes the position of the unit. Thus, we can formalize the expected payoff in a general mixed strategy game in the following manner

$$\sum_{i=1}^{n} x_i = 1,\ \sum_{j=1}^{m} y_j = 1,\ x_i,\ i = 1, ..., n,\ y_j \geq 0,\ j = 1, ..., m$$
$$w_1 = x^T Ay,\ w_2 = x^T By,$$

where w_1 and w_2 denotes the value of the expected payoff for respective player. We can make our denotation of game G a bit more informative by setting it as (A, B, S, T, X, Y), however out of convention will we be referring a bimatrix game as (A, B). Out of convention we are going to let $\triangle_m$ and $\triangle_n$ denote the m-dimensional and n-dimensional mixed strategies, i.e. the simplices

$$\triangle_m = \{x \in R^m : x_i \geq 0, \forall_i \sum_{i=1}^{m} x_i = 1\}$$
$$\triangle_n = \{y \in R^n : y_i \geq 0, \forall_i \sum_{i=1}^{n} y_i = 1\}.$$

Thus, a strategy in the strategy space can be written as $(x, y) \in \triangle_m \times \triangle_n$.

We now want to write a definition for a Nash equilibrium, which is a mixed strategy pair (x^*, y^*) that no players has any motives to alternate from. Formally, we can write this as the following definition.

Definition 5.1. *A Nash equilibrium is the saddle point (x^*, y^*) for the bimatrix game (A, B) if and only if it satisfies*

$$x^{*T}Ay^* \geq x^T Ay^* \quad \forall x \in \triangle_m,$$
$$x^{*T}By^* \geq x^{*T}By \quad \forall y \in \triangle_n.$$

From this the following theorem holds.

Theorem 5.1. *(Nash's existence theorem) Every bimatrix game (A, B) where each player can choose from a finite amount of pure strategies has at least one Nash equilibrium.*

We can prove this with Brouwer's theorem as they have done in [15] and [17].

Theorem 5.2. *(Brouwer's theorem) Let T be a continuous function that maps a compact convex set $S \subset R^n$ to itself. Then there is a point $x \in S$ such that $T(x) = x$, i.e. there is a fixed point in the set.*

Proof. (Of theorem 5.1) Let x and y be any pair of mixed strategies for the bimatrix game (A,B). We can denote $A_{i,.}$ as the ith row of A and $B_{.,j}$ as the j-th column of B. We can now define the maximum perturbation between our optimal solution and our maximal aggravation and denote them as c_i for player I and d_j for player II. Hence, we make the following definitions

$$\begin{aligned}
c_i &:= \max\{A_{i,.}y - x^T Ay, 0\},\ 1 \leq i \leq m \\
d_j &:= \max\{x^T B_{.,j} - x^T Ay, 0\},\ 1 \leq j \leq n \\
x_i' &= \frac{x_i + c_i}{1 + \sum_{k=1}^m c_k},\ 1 \leq i \leq m \\
y_j' &= \frac{y_j + d_j}{1 + \sum_{l=1}^n d_l},\ 1 \leq j \leq n.
\end{aligned}$$

We can see this as a mapping, i.e. $T(x, y) = (x', y')$. We now claim that T is a continuous map that maps the set of all the mixed strategies to itself, i.e. $\triangle_n \times \triangle_m$ to itself.

$$i) \sum_{i=1}^m x_i' = \frac{1}{1 + \sum_{k=1}^m c_k}(\sum_{i=1}^m x_i + \sum_{i=1}^m c_i) = 1$$

$$\text{Likewise, } \sum_{j=1}^n y_j' = 1.$$

$$(ii)\, x_j' \geq 0\ \forall_j = 1, ..., n \text{ and } y_i' \geq 0\ \forall_i = 1, ..., m$$

because $c_i \geq 0$ and $d_j \geq 0$ by definition.

Thus the strategy space must be the Cartesian product $\triangle_n \times \triangle_m$ which must be convex and compact since $\triangle_n$ and $\triangle_m$ are convex and compact. This means T satisfies the conclusion of Brouwer's fixed point theorem, hence T has a fixed point $(x, y) \in \triangle_n \times \triangle_m$. That is, $T(x, y) = (x, y)$ or equivalently $(x', y') = (x, y)$.

Next we will show that a point satisfies T if and only if it is a Nash equilibrium. We will show this from player ones perspective, since the proof is analogue for the second player, i.e. we can do the same for d_j. So we will have two case

$$\begin{aligned}&(1)\ c_i = 0 \text{ and } d_i = 0\\&(2)\ c_i > 0 \text{ or } d_j > 0.\end{aligned}$$

For (1) we have that $x'_i = x_i$ and that $y'_j = y_j$, hence T is satisfied. We also have that this is a Nash equilibrium, since for a Nash equilibrium to be true for player one we have

$$x^T Ay \geq e_i^T Ay = A_{i,.}y,$$

which is only true when for all i $c_i = 0$. We can in a similar manner show that for it to be Nash equilibrium for the second player then for all j $d_j = 0$, hence (1) implies that it is a Nash equilibrium. Obviously, we then have that (2) implies that it is not an Nash equilibrium. Hence, there exists an index such that $A_{p,.}y > x^T Ay$ and $x_p > 0$. But, since $x^T Ay$ is the weighted average of $\{A_{i,.}y\}_{i=1}^m$ then this would imply

$$x'_p = \frac{x_p}{1 + \sum_{k=1}^m c_k} < x_p \implies x' \neq x,$$

which does not satisfy T. Thus the only point that satisfies Brouwer's theorem is the Nash equilibrium, which proves the theorem. □

We have now showed that every bimatrix game (A, B) has a Nash equilibrium, we will now show a theorem that gives more information about this equilibrium.

5.2 Zero sum games

A zero sum game is when a players outcome is directly inclined with the other players loss, hence referring to our previous denotation of the payoff matrices we would have

$$A + B = 0.$$

As we showed in theorem 5.1, there is at least one equilibrium in such a game. We will now show that we can transform such a problem into a LP problem. Now we claim that in a zero sum game both players opt to go for a strategy that minimizes their respective worst expected payoff. Why? This is best illustrated in a example:

$$\begin{pmatrix} 4 & 6 & 8 & 10 & 15 \\ 9 & 7 & 3 & 0 & -2 \end{pmatrix} \tag{Example 1}$$

We can easily see in this example that the column player could be tempted to maximise her payoff by choosing column five, but then the row player can easily retaliate by choosing row two. Observe, that this follows analogously as the primal and dual of a LP, i.e.

$$\begin{aligned}&\min\ c^T y && \max\ x^T b\\&\text{s.t. } Ay \geq b,\ y \geq 0 && \text{s.t. } x^T A \leq c,\ x \geq 0\end{aligned}$$

where $b = (1_1, ..., 1_m)^T$ and $c = (1_1, .., 1_n)^T$. We will now show that we can refine this formulation by the minmax theorem.

Theorem 5.3. *(Minmax) For any $m \times n$ matrix A, the minmax over all strategies equals the maximimum, i.e.*

$$\underset{x}{max}\ \underset{y}{min}\ x^T Ay = \underset{x}{min}\ \underset{y}{max}\ x^T Ay$$

This equality is the value of the game. If the maximum on the left is attained at y^* and the minimum on the right is attained at x^* then those strategies are optimal and will yield a saddle point from which nobody wants to alternate from. Hence, (x^*, y^*) is a Nash equilibrium that must abide the following relation

$$x^{*T} Ay \leq x^{*T} Ay^* \leq x^T Ay^*, \text{ for } \forall x, y.$$

Proof. Since the zero sum game is a LP problem with convex sets means that we can freely use the dual theorem. This means that there exits an optimal probability vectors such that

$$c^T x^* = y^{*T} b.$$

Since c and b are only ones and zeros this implies that there is a saddle point S such that

$$\sum x_i^* = \sum y_i^* = S.$$

For this to be equal to a normal mixed strategy set, i.e. the sum of probabilities needs to equal one, we can divide both sides by S and and receive the optimal mixed strategies

$$\frac{x^*}{S} = \frac{y^*}{S} = 1.$$

If we now observe for any other strategies x y we have

$$\begin{aligned} &Ax^* \geq b & &y^* A \leq c^T \\ \implies &yAx^* \geq yb & \implies &y^* Ax \leq c^T x \\ \iff &yAx^* \geq 1 & \iff &y^* Ax \leq 1 \end{aligned}$$

The last equivalency holds because every mixed strategy sum is equal to one. Thus we get the following expression

$$\begin{aligned} y^* Ax \leq& 1 \leq yAx^* \\ \implies \quad \frac{y^* Ax}{S} \leq& \frac{1}{S} \leq \frac{yAx^*}{S}. \end{aligned}$$

The last expression tells us that if a player were to use the optimal mixed strategy, i.e. y^*/S or x^*/S, then they can ensure that their opponent can maximum win or lose $1/S$. Hence, we get a saddle point $1/S$ if a player use a maxmin strategy. $\square$

We can now apply the minmax theorem in order to solve our previous example.

The optimal strategy for row player is $(\frac{3}{5}, \frac{2}{5})^T$. Clearly both $y_1^* = (\frac{1}{2}, 0, \frac{1}{2}, 0, 0)^T$ and $y_2^* = (\frac{2}{3}, 0, 0, \frac{1}{3}, 0)^T$ are optimal for the column player. By convexity any point in R^5 between these two vectors is an optimal solution, i.e. $x^* = \lambda x_1^* + (1 - \lambda) x_2^*, \forall \lambda \in [0, 1]$.

The optimal solution (x^*, y^*) to this pair of LP problems obtains the optimal strategy for Y player (column player) $\frac{x^*}{\Phi}$ and the optimal strategy for X player (row player) $\frac{y^*}{\Phi}$, where $\Phi = \sum_{i=1}^{n} x_i^* = \sum_{i=1}^{m} y_i^*$. Note that optimal (mixed strategies) are not unique.

5.3 Bimatrix game as a quadratic programming SDP relaxation approach

In this subsection we will first show that a bimatrix game can be formulated as a quadratic program following [1].

We will now show that the Nash equilibrium can be written as a solution to a quadratic problem, we start by claiming that

$$\begin{aligned}
&(x^*, y^*) \in \triangle_m \times \triangle_n \text{ is a Nash equilibrium for game } (A, B) \text{ iff} \\
&x^* A y^* \geq e_i^T A y^*,\ 1 \leq i \leq m \\
&x^* B y^* \geq (x^*)^T A e_i,\ 1 \leq i \leq n
\end{aligned}$$

Proof. The interpretation of this is that there always exists an optimal pure strategy response to a an optimal mixed strategy. Notice that the optimal payoff to a mixed strategy is a convex combination of pure strategies. Then by Brouwer's theorem, as previously stated, there always exists a pure strategy which is a best response to the other player's strategy. □

We can now prove our previous statement and reformulate it to a QP

$$\begin{aligned}
&\min 0 \\
&\text{s.t. } x^T A y \geq e_i^T A y,\ 1 \leq i \leq m \\
&\qquad x^T B y \geq x^T B e_j\ 1 \leq j \leq n \\
&\qquad x_i \geq 0,\ y_i \geq 0,\ \forall n \leq m,\ 1 \leq j \leq n \\
&\qquad \sum_{i=1}^{m} x_i = 1\ \sum_{j=1}^{n} y_j = 1.
\end{aligned}$$

This can be solved by SDP relaxation, if we let

$$\mathcal{M} := \begin{pmatrix} X & P \\ Z & Y \end{pmatrix} \mathcal{M}' := \begin{bmatrix} \mathcal{M} & \begin{matrix} x \\ y \end{matrix} \\ x^T y^T & 1 \end{bmatrix}$$

with $X \in S^{m \times m}$ (all $m \times m$ symmetric matrices), $Z \in R^{n \times m}$, $Y \in S^{n \times n}$ (all $n \times n$ symmetric matrices), $P = Z^T$, $x \in R^m$, $y \in R^n$. Then the QP can be reformulated to

$$\begin{aligned} \min_{\mathcal{M}' \in S^{(m+n+1)(m+n+1)}} \quad & 0 \\ \text{s.t. } & \mathrm{Tr}(Az) \geq e_i^T Ay \quad \forall 1 \leq i \leq m \\ & \mathrm{Tr}(Bz) \geq x^T Ae_j \quad \forall 1 \leq j \leq n \\ & \sum_{i=1}^{m} x_i = 1, \sum_{j=1}^{n} y_j = 1 \\ & \mathcal{M}' \succeq 0, \, \mathcal{M}'_{m+n+1,m+n+1} = 1 \succeq 0 \square \end{aligned}$$

We can without loss of generality assume that all the entries in A and B are between zero and one, because Nash equlibria are invariant under certain affine transformations in the payoff.

Examples of some bimatrix games:

1. Example for transform any pair of A, B to the pair that has elements bounded by 0 and 1:

$$A = \begin{pmatrix} 1 & -3 \\ -1 & -2 \end{pmatrix} B = \begin{pmatrix} 1 & 3 \\ 2 & -4 \end{pmatrix}$$

First we convert them to have positive elements add $3I$ to A:

$\implies A + 3I = \begin{pmatrix} 4 & 0 \\ 2 & 1 \end{pmatrix}$ then we divide it by 4 (so the maximum is 1) $\frac{1}{4}(A + 3\tau) = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{4} \end{pmatrix}$

Similarly

$$B + 4I = \begin{pmatrix} 5 & 7 \\ 6 & 0 \end{pmatrix} \text{ thus we need to divide it by 7: } \frac{1}{7}(B + 4\tau) = \begin{pmatrix} \frac{5}{7} & 1 \\ \frac{6}{7} & 0 \end{pmatrix}$$

so we can choose $c = \frac{1}{4} \geq 0$, $d = \frac{3}{4}$, $e = \frac{1}{7} \geq 0$, $f = \frac{1}{7}$.

where I is the identity matrix.

2. In the case of the condition (x^*, y^*) is a Nash equilibrium if and only if

$$x^{*T} Ay^* \geq e_i^T Ay^* \, \forall i \text{ and } x^{*T} By^* \geq x^{*T} Be_j \, \forall j$$

For the game given by A, B with our previous example our optimal solution

$$x^* = (\frac{3}{4}, \frac{1}{4})^T, \, y^* = (\frac{2}{3}, \frac{2}{3})^T$$

is a Nash equilibrium because

$$x^{*T)} Ay^* = -\frac{5}{3}, \, Ay^* = (-\frac{5}{3}, -\frac{5}{3})^T \text{ so}$$
$$x^{*T} Ay^* = e_1^T Ay^* = e_2^T Ay^* = -\frac{5}{3} \text{ our expected payoff}$$

and

$$x^{*T}By^* = \frac{5}{4},\ x^{*T}B = (\frac{5}{4}, -\frac{5}{4}) \text{ so}$$
$$x^{*T}By^* = e_i^T Be_1 = x^{*T}Be_2 = \frac{5}{4}$$

3. Non unique solution

$$A = \begin{pmatrix} 6 & 3 \\ 1 & 5 \end{pmatrix}\ B = \begin{pmatrix} 5 & 3 \\ 1 & 6 \end{pmatrix}$$

We then get the optimal strategies

$$x^* = (\frac{5}{7}, \frac{3}{7}),\ y^* = (\frac{3}{7}, \frac{5}{7})$$

(x^*, y^*) is a Nash equilibrium since

$$x^*Ay^* = \frac{27}{7} \text{ and } Ay^* = (\frac{27}{7}, \frac{27}{7})^T \implies x^{*T}Ay^* \geq e_1^T Ay^* = e_2^T Ay^*$$
$$x^{*T}By^* = \frac{27}{7} \text{ and } x^*B = (\frac{27}{7}, \frac{27}{7}) \implies x^{*T}By^* \geq x^{*T}Be_j\ j = 1, 2.$$

Hence we can see that both of the players expected payoff from their respective optimal strategy gives the same value, i.e.

$$EP_x = EP_y = \frac{27}{7}.$$

Note also that if $x^* = (1,0)^T$, $y^* = (0,1)^T$ then (x^*, y^*) is also a Nash equilibrium

$$x^{*T}Ay^* = 5$$
$$x^{*T}By^* = 6$$

and $x^* = (1,0)^T$, $y^* = (0,1)^T$ also give a Nash equilibrium with the expected payoff 6 and 5 respectively. If we denote the three pairs in the order they come as P_1, P_2, P_3. Then we see that P_1 is a mixed strategy and P_2 and P_3 are pure strategies.

This example also shows that (x^*, y^*) obtained from solving:

$$\min_y x^{*T}Ay = \max_x \min_y x^T Ay$$
$$\max_y x^T Ay^* = \max_y \min_x x^T By$$

4.

$$A = \begin{pmatrix} 3 & 2 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 2 \\ 3 & 3 & 2 \\ 3 & 2 & 3 \end{pmatrix}\ B = \begin{pmatrix} 2 & 3 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 1 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

Nash equilibrium's

	X player	EP	Y player	EP
1.	$(0,1,0,0,0)$	$=3$	$(1,0,0)$	$=2$
2.	$(0,1,0,0,0)$	$=3$	$(0,1,0)$	$=2$
3.	$(\frac{1}{2},0,0,\frac{1}{2},0)$	$=3$	$(1,0,0)$	$=2$
4.	$(0,0,0,1,0)$	$=3$	$(1,0,0)$	$=2$
5.	$(\frac{2}{3},0,0,0,\frac{1}{3})$	$=3$	$(1,0,0)$	$=\frac{7}{3}$
6.	$(0,0,0,0,1)$	$=3$	$(1,0,0)$	$=3$

5.4 Robust bimatrix games

In this section we show how a bimatrix game with uncertainty can be solved. The problem is

$$\min_{y\in\triangle_m}\max_{\tilde{A}\in D_A,\tilde{z}\in z^u} y^T\tilde{A}\tilde{z} \qquad ((P_A))$$

$$\min_{z\in\triangle_n,}\max_{\tilde{B}\in D_B,\,\tilde{y}\in Y^u} \tilde{y}^T\tilde{B}z \qquad ((P_B))$$

we deal with the following uncertainty types;
(i) the payoffs are uncertain
(ii) the opponent's strategy is uncertain.

The uncertainty in these two types of the problems will be formulated as box and ellipsoid uncertainty, which will be described soon.

5.4.1 Uncertain payoffs

Now the game problem is the pair of optimisation problems

$$\min_{y\in\triangle_m}\max_{\tilde{A}\in D_A} y^T\tilde{A}z$$

$$\min_{z\in\triangle_n}\max_{\tilde{B}\in D_B} y^T\tilde{B}z$$

Our goal is to determine the uncertainty sets D_A and D_B so that our problem becomes tractable, as described in section 4. To this and we consider two different sets of uncertainty payoffs.

(a) Box Uncertainty:

$$D_A = \{\tilde{A} : \ \tilde{A} = A + M,\ -\underline{M} \leq M \leq \overline{M}\} =: D_A^B$$
$$D_B = \{\tilde{B} : \ \tilde{B} = B + N,\ -\underline{N} \leq N \leq \overline{N}\} =: D_B^B$$

where $\underline{M}$, $\overline{M}$, $\underline{N}$, $\overline{N} \in R^{m\times n}$ have nonnegative components. This uncertainty is called box uncertainty.

(b) Ellipsoid uncertainty

$$D_A^E = D_A^1 \times \cdots \times D_A^n$$
$$D_B^E = D_B^i \times \cdots \times D_B^m$$
$$D_A^i = \{A_j^c + R_j(\triangle p_j) : ||\triangle p_j|| \leq \Gamma_j\},\ j = 1, ..., n,$$

where A_j^c is the j-th column of A, $j = 1, ..., n$ and R_j are $m \times L_j$ matrices, $\triangle p_j = (\triangle p_j^1, ..., \triangle p_j^{L_j})^T \in R^{L_j}$. We consider payoff uncertainty for D_B^W similarly

$$D_B^i = \{B_i^r + (S_i \triangle q_i)^T : ||\triangle q_i|| \leq \Omega_i\},\ i = 1, ..., m,$$

where S_i are $n \times K_i$ matrices, $\triangle q_i = (\triangle q_i^1, ..., \triangle q_i^{K_i}) \in R^{K_i}$ and B_i^r is the i-th row of B, $i = 1, ..., m$.

Note that Γ_j and Ω_i are parameters controlling the robustness and optimality. This uncertainty is called ellipsoidal uncertainty, since if $L_j = \Omega_i = 1$ for all i and j, all $L_j = n$, $j = 1, ..., n$, $K_i = m$, $i = 1, ..., m$, then D_A^i, D_B^i are the standard ellipsoid uncertainty sets in section 4.

Now we study the box uncertainty for player with the payoff $\tilde{A} \in D_A^B$ in detail. The first task is to solve the inner maximisation of $y^T \tilde{A} z$ for $\tilde{A} \in D_A^B$. That is

$$\max_{-\underline{M} \leq M \overline{M}} y^T (A + M) z = y^T (A + \overline{M}) z.$$

So we obtain the following minimisation problem over y:

$$\min_y \ y^T (A + \overline{M}) z$$
$$\text{s.t. } y \geq 0, \sum_{i=1}^m y_i = 1.$$

Let now $\mathbb{1}_m = (1, ..., 1)^T \in R^m$. Then $\sum_{i=1}^m y_i = 1$ is equivalently to $\mathbb{1}_m^T y = 1$. Let further $f(y) = y^T (A + \overline{M}) z$. Since this is an LP problem, the KKT conditions are necessary and sufficient for optimum. Thus, we can use the KKT conditions to solve the problem. The KKT conditions are

$$(i) \nabla [f(y) + (-\lambda)^T y + \alpha (\sum_{i=1}^m y_i - 1) = 0]$$

$$(ii) y \geq 0, \sum_{i=1}^m y_i = 1$$

$$(iii) \lambda \in R_+^m,\ \alpha \in R$$

$$(iv) \lambda_i y_i = 0,\ i = 1, ..., m.$$

Explicitly (i) is $\lambda = (A + \overline{M}) z + \alpha \mathbb{1}_m \geq 0$ since (iii) holds.
Hence the KKT conditions can be stated as

$$y \in R_+^m,$$
$$y \perp (A + \overline{M}) z + \alpha \mathbb{1}_m \in R_+^m,$$
$$\mathbb{1}_m^T y = 1.$$

Similarly, we can solve the problem for the player with payoff $\tilde{B}$, that is the KKT conditions can be rephrase as

$$z \in R_+^n,\ z \perp (B + \overline{N})^T y + \mathbb{1}_n \beta \in R_+^n \text{ and}$$
$$\mathbb{1}_n^T z = 1.$$

Now we can solve the bimatrix game with box uncertainty payoff by solving $x = (y^T, z^T, \alpha, \eta)^T \in R_+^m \times R_+^n \times R \times R$, $Gx \in R_+^{m+n}$, $Hx \in R_+^{m+n}$ and $Gx \perp Hx$, $Cx = d$.

$$G = \begin{pmatrix} I_m & 0 & 0 & 0 \\ 0 & I_n & 0 & 0 \end{pmatrix},\ H = \begin{pmatrix} 0 & A + \overline{M} & \mathbb{1}_n & 0 \\ B^T + \overline{N}^T & 0 & 0 & \mathbb{1}_m \end{pmatrix}$$
$$C = \begin{pmatrix} \mathbb{1}_m^T & 0 & 0 & 0 \\ 0 & \mathbb{1}_n^T & 0 & 0 \end{pmatrix},\ d = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Next we turn to the game with ellipsoid uncertainty payoff. As in the previous situation we solve the inner maximisation problem. By definition of the set D_A^E we have

$$\min_{y \in \triangle_n} \max_{\tilde{A} \in D_A^E} y^T \tilde{A} z. \tag{P_I^E}$$

$$\begin{aligned} \max_{\tilde{A} \in D_A^E} y^T \tilde{A} z &= y^T A z + \max_{||\triangle p_j|| \le \Gamma_j\ 1 \le j \le n} \sum_{j=1}^n z_j y^T R_j (\triangle p_j) \\ &= y^T A z + \sum_{j=1}^n z_j \Gamma_j ||R_j^T y|| \end{aligned}$$

because $y^T R(\triangle p_j) = (R^T y)^T (\triangle p_j) \le ||R^T y|| ||\triangle p_j|| |\triangle p_j|| \le \Gamma_j ||R_j^T y||$ using the Cauchy-Schwarz inequality and matrix computation. Hence, (P_I^E) is reduced to

$$\begin{aligned} &\min_y\ y^T A z + \sum_{j=1}^n z_j \Gamma_j ||R_j^T y||_2 \\ &\text{s.t. } y \ge 0, \mathbb{1}_m^T y = 1. \end{aligned} \tag{$P_{I'}^E$}$$

Similarly we get an optimisation problem for the player with payoff $\tilde{B}$

$$\begin{aligned} &\min_z\ y^T B z + \sum_{i=1}^m y_i \Omega_i ||S_i^T z|| \\ &\text{s.t. } z \ge 0,\ \mathbb{1}_n^T = 1. \end{aligned} \tag{$P_{II'}^E$}$$

Solving $(P_{I'})^E$ and $(P_{II'})^E$ for y and z simultaneously we obtain optimal strategy for the bimatrix game with ellipsoid uncertainty payoffs. These are $SOCP$ problem. These problems can be solved by the KKT conditions. The KKT conditions for $(P_{I'}^E)$

are

$$(i) R^m \ni y \geq o, \mathbb{1}^T y = 1$$

$$(ii) \lambda^y \in R^m_+,\ \mu \in R$$

$$(iii) \lambda^y_i y_i = 0 \quad \forall i = 1, ..., m$$

$$\mu(\mathbb{1}^T_m y - 1) = 0$$

$$(iv) \nabla_y((Az)^T y + \sum_{j=1}^{m} z_j \Gamma_j ||R_j^T y||_2 - \lambda^T y + \mu(II_m^T y - 1) = 0$$

(iv) can explicitly be formulated as

$$(Az)^T + \mu \mathbb{1}^T_m - \lambda^T + \sum_{j=1}^{m} z_j \Gamma_j \frac{(R_j R_j^T y)^T}{||R_j^T y||_2} = 0$$

$$\iff Az + \mu \mathbb{1}_m - \lambda + \sum_{j=1}^{m} \frac{z_j \Gamma_j R_j R_j^T y}{||R_j^T y||_2} = 0.$$

Let

$$u_j = -\frac{z_j \Gamma_j R_j R_j^T y}{||R_j^T y||_2}.$$

Then

$$\lambda = Az + \mu \mathbb{1}_m - \sum_{j=1}^{m} R_j u_j \geq 0 \text{ and } \lambda \perp y.$$

$$\iff R^n_+ \ni y \perp Az + \mu \mathbb{1}_m - \sum_{j=1}^{m} R_j u_j \in R^m_+.$$

Let $\gamma_j = ||R_j^T y||$. Then

$$u_j^T R_j^T y + \Gamma_j z_j \dot{\gamma}_j = -z_j \Gamma_j ||R_j^T y|| + \Gamma_j z_j ||R_j^T y||.$$

Note also that $||u_j||_2 = z_j \Gamma_j$. So we have

$$L^{L_j+1} \ni \begin{pmatrix} u_j \\ \Gamma z_j \end{pmatrix} \perp \begin{pmatrix} R_j^T y \\ \gamma_j \end{pmatrix} \in L^{L_j+1}$$

Analogy the KKT condition for player with $\tilde{B}$ are

$$L^{K_i+1} \ni \begin{pmatrix} t_i \\ \Omega_i y_i \end{pmatrix} \perp \begin{pmatrix} S_i^T z \\ \sigma_i \end{pmatrix} \in L^{K_i+1},\ i = 1, ..., n$$

$$R^n_+ \ni z \perp B^T y + \mathbb{1}_n \eta - \sum_{i=1}^{m} S_i t_i \in R^n_+,\ \mathbb{1}_n^T z = 1.$$

Note that the unknowns are $y, z, u_1, ..., u_n, t_1, ..., t_m, \gamma_1, ..., \gamma_n, \sigma_1, ..., \sigma_m, \xi, \eta$.

5.4.2 Bimatrix game with opponent's uncertainty

Now we discuss the optimisation problem

$$\min_{y\in\triangle_m}\max_{\tilde{z}\in Z} y^T A\tilde{z}$$
$$\min_{z\in\triangle_n}\max_{\tilde{y}\in Y} \tilde{y}^T Az.$$

Here we also consider box and ellipsoid uncertainty where $\overline{u}, \underline{u} \geq 0$. But we have to make sure that $\tilde{z} \geq 0$, $\sum_{i=1}^n \tilde{z}_i = 1$. To this end we need one more constraint on u, that is $\sum_{i=1}^n u_i = 0$. So,

$$Z^B = \{\tilde{z} :\ \tilde{z} = z + u, -\underline{u} \leq u \leq \overline{u}, \sum_{i=1}^n u_i = 0\}.$$

Now define the uncertainty set

$$\mathcal{U} := \{u :\ -\underline{u} \leq u \leq \overline{u},\ \sum_{i=1}^n u_i = 0\}.$$

So

$$\begin{aligned}\max_{\tilde{z}\in Z^B} y^T A\tilde{z} &= \max_{u\in\mathcal{U}} y^T A(z+u)\\ &= y^T Az + \max_{u\in\mathcal{U}} y^T Au.\end{aligned}$$

This implies that we can formulate I^u under box uncertainty to a LP problem

$$\begin{aligned}\max\ & (A^T y)^T u \\ \text{s.t.}\ & -\underline{u} \leq u \leq \overline{u}, \\ & \sum_{i=1}^n u_i = 0 \iff \mathbb{1}_n^T u = 0,\ \mathbb{1}_n^T = (1, ..., 1) \in R^n.\end{aligned} \tag{P_0}$$

Next is to find its dual problem. Note that

$$\begin{cases}\min\ c^T x \\ \text{s.t.}\ Ax = b, x \geq 0\end{cases} \tag{P}$$

$$\begin{cases}\max\ y^T b \\ \text{s.t.}\ A^T y \leq c\end{cases} \tag{D}$$

are a pair of dual problems. We can reformulate (P_o) as

$$\begin{aligned}\max\ & (A^T y)^T u \\ \text{s.t.}\ & u \leq \overline{u} \\ & -u \leq \underline{u} \\ & \mathbb{1}_n^T u \leq 0 \\ & -\mathbb{1}_n^T u \leq -0\end{aligned}$$

We can see that the constraints are equivalent to

$$\begin{pmatrix} I \\ -I \\ \mathbb{1}_n^T \\ \mathbb{1}_n^T \end{pmatrix} u \le \begin{pmatrix} \overline{u} \\ -\underline{u} \\ 0 \\ 0 \end{pmatrix}$$

This is the form on (D). So the dual of it is

$$\min_{r_1,r_2,c_1,c_2} (\overline{u}^T, \underline{u}^T, 0, 0) \begin{pmatrix} r_1 \\ r_2 \\ c_1 \\ c_2 \end{pmatrix} = \overline{u}^T r_1 + \underline{u}^T r_2$$

$$\text{s.t. } (I, -I, \mathbb{1}_n, -\mathbb{1}_n) \begin{pmatrix} r_1 \\ r_2 \\ c_1 \\ c_2 \end{pmatrix} = A^T y$$

$$r_1 \ge 0,\ r_2 \ge 0\, r_1, r_2 \in R^n$$

$$c_1 \ge 0,\ c_2\, c_1, c_2 \in R$$

And this is equivalent to

$$\begin{aligned} \min_{r_1,r_2,c_1,c_2} \ & \overline{u}^T r_1 + \underline{u}^T r_2 \\ \text{s.t. } & r_1 - r_2 + c_1 \mathbb{1}_n - c_2 \mathbb{1}_n = A^T y \\ & r_1 \ge 0, r_2 \ge 0 \\ & c_1, c_2 \ge 0 \end{aligned}$$

Set $\alpha := c_1 - c_2$ we obtain (note that α is sign unconstrained)

$$\begin{aligned} \iff \min \ & \overline{u}^T r_1 + \underline{u}^T r_2 \\ \text{s.t. } & r_1 - r_2 + \alpha \mathbb{1}_n = A^T y \\ & r_1 \ge 0,\ r_2 \ge 0. \end{aligned}$$

Since the game is finite by duality theorem for LP

$$\max_{u \in U} y^T A u = \min_{r_1, r_2, \alpha} \{ \overline{u}^T r_1 + \underline{u}^T r_2 : r_1 - r_2 + \alpha \mathbb{1}_n = A^T y,\ r_1, r_2 \ge 0 \}$$

Substitute this to (I^B)

$$\begin{aligned} \min_{y \in \triangle_m} \max_{\tilde{z} \in Z^B} y^T A \tilde{z} &= \min_{y \in \triangle_m} \max_{u \in U} y^T A (z + u) \\ &= \min_{y \in \triangle_m} (y^T A z + \max_{u \in U} y^T A u) \\ &= \min_{y \in \triangle_m} y^T A z + \overline{u}^T r_1 + \underline{u}^T r_2 \\ \text{s.t. } & r_1 - r_2 + \alpha \mathbb{1}_n = A^T y \\ & r_1,\ r_2 \ge 0. \end{aligned}$$

Thus we can rewrite $I^{B'}$ to the following

$$\begin{aligned} \min_{r_1,r_2,y,u} & y^T Az + \overline{u}^T r_1 + \underline{u}^T r_2 \\ \text{s.t. } & r_1 - r_2 + \alpha 1_n = A^T y \\ & r_1 \geq 0, r_2 \geq 0 \\ & y \geq 0, 1_m^T y = 1. \end{aligned} \tag{$(I^{B'})$}$$

Player II in similar manner we get

$$\begin{aligned} \min_{z,t_1,t_2,p} & y^T Bz + \overline{r}^T t_1 + \underline{U}^T t_2 \\ \text{s.t. } & t_1 - t_2 + \beta \mathbb{1}_m = Bz \\ & t_1 \geq 0, t_2 \geq 0 \\ & z \geq 0, \mathbb{1}_n^T z = 1. \end{aligned} \tag{$(II^{B'})$}$$

So if we solve these two problems we get y and z.

Now we work out the KKT conditions for the linear programming problem from box uncertainty z^B, y^B. For player I introduce Lagrange multiplies $\lambda_3 \in R_+^m$, λ_1, $\lambda_2 \in R_+^n$, $\mu \in R^n$, $\mu_0 \in R$. Then the Lagrange function is

$$f(y, r_1, r_2, \alpha) = z^T A^T y + \overline{u}^T r_1 + \underline{u}^T r_2 + \mu(\mathbb{1}_n \alpha + r_1 - r_2 - A^T y) + \mu_0(\mathbb{1}_m^T y - 1) - \lambda_1^T r_1 - \lambda_2^T r_2 - \lambda_3^T y.$$

Then the KKT conditions are

(i) $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, $\lambda_3 \geq 0$

(ii) $r_1 \geq 0$, $r_2 \geq 0$, $y \geq 0$, $\mathbb{1}_m^T y = 1$, $\mathbb{1}_n \alpha + r_1 - r_2 - A^T y = 0$

(iii) $\underset{(y,r_1,r_2,\alpha)}{\nabla f} = 0$, i.e.

$$(z^T A^T - \mu^T A^T + \mu_0 \mathbb{1}_m^T - \lambda_3^T,\ \overline{u}^T + \mu^T - \lambda_1^T,\ y^T - \mu^T - \lambda_2^T,\ \mu^T \mathbb{1}_n) = 0$$

$$\iff$$

(iii') $\lambda_3 = Az - A\mu + \mathbb{1}_m \mu_0$, $\mu^T \mathbb{1}_n = 0$

$\lambda_1 = \overline{u} + \mu$, $\lambda_2 = \underline{u} - \mu$

(iv) $\mu_i(\mathbb{1}_n \alpha + r_1 - r_2 - A^T y)_i = 0$, $\forall_i = 1, ..., n$

$\mu_0(\mathbb{1}_m^T y - 1) = 0$

$\lambda_{1,i} r_{1,i} = 0$, $\lambda_{2,i} r_{2,i} = 0$, $i = 1, ..., n$, $\lambda_{3,i} y_i = 0$, $i = 1, ..., m$

Simplifying the conditions further we get $Az - A\mu + \mathbb{1}_m\mu_0 \in R_+^m$ (since $\lambda_3 \geq 0$) it is orthogonal to $y \in R_+^m$ (by (iv) and (ii)), which is

$$\begin{aligned} R_+^m \ni y \perp Az - A\mu + \mathbb{1}_m\mu_0 \in R_+^m \\ r_+^n \ni r_1 \perp \lambda_1 \in R_+^n,\, R_+^n \ni r_2 \perp \lambda_2 \in R_+^n \,(\text{by } (iv),\, (i),\, (ii)) \\ \mathbb{1}_n\alpha + r_1 - r_2 = A^T y,\, \mathbb{1}_m^T y = 1,\, (\text{by } (iii')) \\ \overline{u} = \lambda_1 - \mu,\, \underline{u} = \lambda_2 + \mu \,(\text{by } (iii)) \end{aligned}$$

We similarly have

$$\begin{aligned} R_+^n \ni z \perp B^T y - B^T y - B^T\omega + \mathbb{1}_n\eta \in R_+^n \\ R_+^m \ni t_1 \perp v_1 \in R_+^m,\, R_+^m \ni t_2 \perp v_2 \in R_+^m \\ \mathbb{1}_m B + t_1 - t_2 = Bz,\, \mathbb{1}_n^T\omega = 0,\, \overline{v} = v_1 - \omega,\, \underline{v} = v_2 + \omega. \end{aligned}$$

(b) Ellipsoid uncertainty. Now we investigate a solution method for ellipsoid uncertainty in the opponent's strategy. Let z be a nominal value and the center of the ellipsoid. Then the uncertainty is the form

$$z + \sum_{i=1}^{n} \tilde{z}_i = 1 \text{ and } \tilde{z}_i \geq 0,\, i = 1, ...n \text{ or } \mathbb{1}_n^T \tilde{z}_i = 0.$$

which can be guaranteed if

$$\mathbb{1}_n^T(\sum_{k=1}^{K}(\triangle d_k)z^k) = 0 \text{ and } z + (\sum_{k=1}^{K}(\triangle d_k)z^k) = 0.$$

Here we have the uncertainty strategy set

$$Z^E := \{z + \sum_{k=1}^{K}(\triangle d_k)z^k : 1_n^T(\sum_{k=1}^{K}(\triangle d_k)z^k) = 0,\, z + \sum_{k=1}^{K}(\triangle d_k)z^k \geq 0,\, ||\triangle d||_2 \leq 1\}.$$

Recall the player I's task:

$$\min_{y\in\triangle_m,\, \tilde{z}\in Z^E} \max y^T A\tilde{z} \tag{I^E}$$

we first work out $\max\limits_{\tilde{z}\in Z^E} y^T A\tilde{z}$ for a fixed y. Since

$$y^T A\tilde{z} = y^T A(z + \sum_{k=1}^{K}(\triangle d_k)z^k) = y^T Az + y^T A(\sum_{k=1}^{K}(\triangle d_k)z^k)$$

maximising it over Z^E is equal to

$$\begin{aligned} &\max\, y^T A(\sum_{k=1}^{K}(\triangle d_k)z^k \\ &\text{s.t. } \mathbb{1}_n^T(\sum_{k=1}^{K}(\triangle d_k)z^k) = 0,\, z + \sum_{k=1}^{K}(\triangle d_k)z^k \geq 0,\, ||\triangle d||_2 \leq 1. \end{aligned}$$

Assume the existence of the Slater's point then the strong duality holds. We have learned that the dual problem is nicer, so we try to solve this problem. By standard techniques, we get the dual problem as follows:

Change to $\min -y^T A \sum(\triangle d_k) z^k$

Introduce Lagrange multiplier $\lambda_1 \in R^n_+$, $\lambda_2 \in R_+$, $\mu \in R$.

Construct the Lagrangian function and minimize it over $\triangle d$, i.e.

$$\min_{\triangle d} \; - y^T A \sum_{k=1}^{K} (\triangle d_k) z^k + \mu 1_n^T (\sum_{k=1}^{K} (\triangle d_k) z^k) - \lambda_1^T (z + \sum_{k=1}^{k} (\triangle d_k) z^k) + \lambda_2 (||\triangle d||_2 - 1).$$

The optimality conditions gives

$$\begin{aligned} -y^T A z^k + \mu \mathbb{1}_n^T z^k - \lambda_1^T z^k + \frac{\lambda_2 \triangle d_k}{||\triangle d||_2} = 0, \; 1e = 1, ..., K \\ \iff \; -y^T A z^k + \mu \mathbb{1}_n^T z^k - \lambda_1 z^k = -\frac{\lambda_2 \triangle d_2}{||\triangle d||} \end{aligned}$$

plugging this into the lagrangian function gives

$$\begin{aligned} &\min \sum_{k=1}^{K} - \frac{\lambda_2 \triangle d_k}{||\triangle d||_2} \cdot \triangle d_k - \lambda_1^T Z + \lambda_2 ||\triangle d||_2 - \lambda_2 \\ =& -\lambda_2 ||\triangle d||_2 - \lambda_1^T z + \lambda_2 ||\triangle d||_2 - \lambda_2 \\ =& -\lambda_1^T z - \lambda_2 \end{aligned}$$

This shows the dual problem is

$$\begin{aligned} \max_{\mu, \lambda_1, \lambda_2} \; & (-\lambda_1^T z - \lambda_2) \\ \text{s.t. } & \lambda_1 \geq 0, \; \lambda_2 \geq 0 \\ & y^T A z^k = \mu 1_n^T z^k - \lambda_1^T z^k + \frac{\lambda_2 \triangle d_k}{||\triangle d||_2}, \; k = 1, ..., K. \end{aligned}$$

We can eliminate $\triangle d_k$ by reformulating this further, we have

$$\begin{aligned} \min_{\mu \in R, \, \rho \in R^K, \, \lambda_1 \in R_+, \lambda_2 \in R_+} \; & \lambda_1^T z + \lambda_2 \\ \text{s.t. } & y^T A z^k = \mu 1_n^T z^k - \lambda_1^T z^k + \rho_k, \; k = 1, ..., K \\ & ||\rho||_2 \leq \lambda_2 \\ & \lambda_1 \geq 0, \; \lambda_2 > 0. \end{aligned}$$

This is a second order conic optimisation problem. Note that strong duality holds. Hence, we convert the optimisation problem (I^E) to the following $SOCP$ problem over $(y, \mu, \rho, \lambda_1, \lambda_2) \in R^m \times R \times R^K \times R^n \times R$.

$$\begin{aligned} \min \; & y^T A z + z^T \lambda_1 + \lambda_2 \qquad (*) \\ \text{s.t. } & y^T A z^k \leq \mu 1_n^T z^k + \lambda_1^T z^k \\ & \mathbb{1}_m^T y = 1, \; ||\rho||_2 \leq \lambda_2, \\ & y \geq 0, \; \lambda_1 \geq 0, \; \lambda_2 \geq 0. \end{aligned}$$

Let $Z = (z^1, ..., z^K) \in R^{n\times k}$. Then (*) $\iff$

$$\rho = \begin{pmatrix} y^T A z^1 \\ ... \\ y^T A z^k \end{pmatrix} - \begin{pmatrix} \mu 1_1^T z^1 \\ ... \\ \mu 1_n^T z^k \end{pmatrix} + \begin{pmatrix} \lambda_1^T z^1 \\ ... \\ \lambda_1^T z^k \end{pmatrix}$$
$$= (AZ)^T y - z^T 1_n \mu + Z^T \lambda_1.$$

This gives the following SOCP problem over $y, I\cdot\mu,\ I\cdot\rho,\ I\cdot\lambda_1,\ I\cdot\lambda_2$:

$$\begin{aligned} &\min\ y^T A z + z^T \lambda_1^I + \lambda_2^I \\ &\text{s.t. } \rho^I = (Az)^T y - z^T 1_n \mu^I + Z^T \lambda_1^I \\ &\quad 1_m^T y = 1,\ ||\rho^I||_2 \le \lambda_2^I \\ &\quad y \ge 0,\ \lambda_1^I \ge 0,\ \lambda_2^I \ge 0. \end{aligned} \tag{P_I^*}$$

The second player has to solve the following SOCP problem

$$\begin{aligned} &\min\ y^T B z + y^T \lambda_1^{II} + \lambda_2^{II} \\ &\text{s.t. } \rho^{II} = Y^T B z - Y^T 1_m \mu^{II} + Y^T \lambda_1^{II} \\ &\quad 1_n^T z = 1,\ ||B^{II}||_2 \le \lambda_2^{II} \\ &\quad z \ge 0,\ \lambda_1^{II} \ge 0,\ \lambda_2^{II} \ge 0. \end{aligned} \tag{P_{II}}$$

Where the $\rho^{II} = (\rho_1^{II}, ..., \rho_L^{II})^T \in R^L$ from the uncertainty set

$$Y^E := \left\{ y + \sum_{l=1}^{L} (\triangle h_l) y^l : e_m^T (\sum_{l=1}^{L} (\triangle h_l) y^l) = 0, y + \sum_{l=1}^{L} (\triangle h_l) y^l \ge 0,\ ||\triangle h||_2 \le 1 \right\}$$

and $y = (y^1, ..., y^L) \in R^{m\times L}$, and the minimization is over $(z, \mu^{II}, \rho^{II}, \lambda_1^{II}, \lambda_2^{II}) \in R^n \times R \times R^L \times R^m \times R$.

The optimal mixed strategy pair (y^*, z^*) is obtained by solving the SOCP problems (P_I), and (P_{II}) simultaneously.

6 Discussion

As explained in the beginning of this paper can robust programming be applied in most fields, however fields that does tend to get mentioned are within finance and and engineering, examples of problems within these fields are inventory and networking problems. For an inventory problem one can easily imagine a problem such that a manager wants to minimise the inventory cost. For such a problem the manager would have to take in account that; there is a limited space in the inventory, retailers require certain amount of product each month which require that we match that order from our suppliers. The inevitable uncertainty in this problem is the demand that can vary. However, this does not affect the supplier so they will have to create some sort of penalty for how much the manager deviate from the agreed value. Constructing this uncertainty set can actually be very difficult, since it is easy that one chooses an uncertainty set that yields robust solutions that are too conservative, so that the resulting solutions has too low quality towards the objective. Thus, will one have to use probability tools in order to reduce the uncertainty set, in such a way that it does not encamps results that does not interest us.

Acknowledgements

I want to express my gratitude to my mentor Yishao Zouh, who has always been there and supported me in this journey. She has contributed with theories and guidance, but also other fun aspect around this subject. Without her kind direction would this study have been a little success. I would also like to say thanks to Sven Raum who reviewed my work and gave some useful notes.

7 References

References

[1] G. Lou, X. An, J. Xia, Robust optimization with application to game theory, 21 sep 2010

[2] A. Ben-Tal, A. Nemivroski, Lectures on modern convex optimization, 2013, Cambridge

[3] G. Califiore, L. El Ghaoui, Optimizatoin Models, 2014,

[4] A. Ben-Tal., L. El Ghaoui, A. Nemivroski, Robust Optimization, 2009, Princeton Univerisity

[5] V. Jeykumar, G.Y. Li, G.M. Lee, A Robust von Neuman Minimax Theorem for Zero-Sum Games under Bounded Payoff Uncertainty, 2011

[6] G. Strang, Liniear Algebra and its Applications Fourth edition

[7] A. N. Letchford, A. J. Parkes, A Guide to Conic Optimisation and its Applications

[8] T. S. Ferguson, Game Theory

[9] A. Ben-Tal, A. Nemirovski, Efficient Methods in Optimization, Cones, 2019

[10] A.A. Ahmadi, J. Zhang, Semidefinite Programming and Nash Equlibria in Bimatrix Games, 2019

[11] T. Matsumoto, E. Shimemura, The Kuhn Tuckers Conditions and Multisage Games, Waseda University, 1968

[12] W.L. Winston, Operation Research Applications And Algorithms Fourth Edition, Indiana University, 2004

[13] M. Aghassi, D. Bertsimas, Robust Game Theory, 2006

[14] M. Luptácik, Kuhn–Tucker Conditions, University of Economics in Bratislava, 2010

[15] R. Chandrasekaran, Bimatrix Games

[16] A. Ben-Tal, A. Nemirovski, Robust Solutions of Uncertain Liniear Programs

[17] O. Wilson, Two-Person General-Sum Games

[18] M. S. Bazaraa, H.D. Sherali, C.M. Sheitty, Nonliniear programming Theory and Algorithms