Robust Optimization

Vishal Kashyap (Mixed Mode)

***Abstract*— This paper presents a study on a polynomial-time interior-point algorithm for a class of nonlinear saddle-point problems that involve semi-definiteness constraints on matrix variables. These problems originate from robust optimization formulations of convex quadratic programming problems with uncertain input parameters.**

I. INTRODUCTION

Optimizing an objective function in real life applications with uncertain parameters is greatly affected even by small perturbations in these parameters. Be it the co-variance matrix in Mean-Variance optimization or noisy sensor data, estimation of these parameters can be difficult. In these situations one works with the best estimation of these parameters. Robust Optimization comes to the rescue and generates results considering the uncertainties. In this paper we consider the following convex quadratic optimization problem :

$$\min_x c^T x + \frac{1}{2} x^T Q x$$
$$s.t. Ax \geq b$$

where $A \in R^{m \times n}, b \in R^m, c \in R^n$ and $Q \in S_n^+$ are the input data and $x \in R^n$ are the decision variables. We will study the solution to this problem with box uncertainty in c and Q, i.e $\{Q : Q_L \leq Q \leq Q_U\}$ and $\{c : c_L \leq c \leq c_U\}$, where the inequalities are component wise. Note that here Q is symmetric and therefore will have $\frac{n(n+1)}{2}$ inequalities. These box uncertainties are a natural choice for many applications and provide flexibility and modelling as the component wise inequality can be independently set. [3] shows that such a problem can be formulated as a saddle point problem. The rest of the paper is organized as follows. In Section II saddle point formulation is discussed followed by a central-path (along the saddle points with proximity defined) algorithm. Section III gives an algorithm to solve the said problem and gives to its polynomial time complexity. Finally in section IV we apply this algorithm to Markowitz model for portfolio optimization on SP top500 stocks and present an analysis and the results.

II. PROBLEM FORMULATION

A. *Problem Formulation*

In robust optimization approach, a conservative view point is chosen and one tries to find the best worse-case performance.We model the uncertainty in the objective function using the following uncertainty set :

$$y := y(c_L, c_U, Q_L, Q_U)$$
$$:= \{(c, Q) \in R^n \times S_n : c : c_L \leq c \leq c_U,$$
$$Q : Q_L \leq Q \leq Q_U, Q \geq 0\}$$

Now we can define

$$\phi(x, y) = c^T x + \frac{1}{2} x^T Q x$$

Note that ϕ is a quadratic function of x and a linear function of c and Q. Given an instance of decision variable x, the worst case realization of ϕ(the objective function can be given by) :

$$f(x) := \max_{(c,Q) \in y} \phi(x, c, Q)$$

To get a robust solution we need to minimize the worst-case performance, this is clearly given by minimizing f(x):

$$\min_x f(x) = \min_x \{\max_{(c,Q) \in y} \phi(x, c, Q)\} \tag{1}$$

Alternatively we can first do the best allocation of decision variables based on an instance of c and Q (QP) and then take the worst case performance on y.Denote:

$$g(c, Q) := min_x \phi(x, c, Q)$$
$$\max_{(c,Q) \in y} g(c, Q) = \max_{(c,Q) \in y} \{\min_x \phi(x, c, Q)\} \tag{2}$$

We will follow these steps being listed out to reach an optimal point for this robust problem:

1) Show that it is a convex-concave saddle point problem.
2) Introduce self concordant barrier functions to handle inequalities.
3) Introduce ϕ_t, a function combining the convex-concave SPP and barrier functions. (interior point method)
4) Show that ϕ_t has a unique saddle point which can be reached using newton decrements, i.e it is reached when newton step is 0.
5) Being close to the saddle point of ϕ_t gives bound on the duality gap of original problem. (bound dependent on t)
6) One can follow the central path (increase t) to reach a saddle point of original problem.

We will now show these one by one. Note that all the above steps have been covered in great detail [3], we review these as a preview to the application problem.

B. *Convex-Concave SPP*

Problems 1 and 2 are duals of each other. The problem 1 is the primal whereas 2 is the dual. Therefore the weak duality $g(c, Q) \leq f(x)$ holds and can be easily checked using the epigraph trick.

The function $\phi(x, y)$ is convex-concave, i.e for any fixed value of $(c, Q) \in y$ it is convex (Q is PSD) and for any fixed value of x it is concave (in fact, linear). However these functions maybe non-smooth and we will outline a saddle-point approach to solve them using self-concordant barrier functions for the inequalities.

It can be shown that optimal values of 1 and 2 are equal, i.e strong duality holds and also the function ϕ has a saddle point. Thus the function $\phi(c, Q, x)$ acts as a saddle function and we want to find $\overline{x}$ and $(\overline{c}, \overline{Q})$ such that the following saddle point property holds:

$$\phi(\overline{x}, c, Q) \leq \phi(\overline{x}, \overline{c}, \overline{Q}) \leq \phi(x, \overline{c}, \overline{Q}), \forall x, (c, Q) \in y \quad (3)$$

This saddle-point formulation will help us in solving the robust optimization problem and will also give a proximity measure that will help us monitor the progress of the algorithm. Let X^0 and Y^0 denote the interior of the the sets X and Y respectively. Define the following barrier functions:

C. Introducing Barrier Functions

$$F(x) = -\sum_{i=1}^{m} log[Ax - b]_i \forall x \in X^0 \quad (4)$$

$$\begin{aligned} G(x) = &-\sum_{j=1}^{n} log(c_j^U - c_j) - \sum_{j=1}^{n} log(c_j - c_j^L) \\ &- \sum_{1 \leq i \leq j \leq n} log(Q_{ij}^U - Q_{ij}) - \sum_{1 \leq i \leq j \leq n} log(Q_{ij} - Q_{ij}^L) \\ &- logdet(Q), \forall (c, Q) \in y^0 \end{aligned}$$

It is verifiable that $F(x)$ and $G(y)$ are self-concordant functions with paramters m and $n^2 + 4n$ respectively. Self-concordance will allow us to do a better analysis of convergence to the saddle point as we will be using newton's decrement. Consider the following saddle-barrier function for $t \geq 0$:

$$\phi_t(x, y) := t\phi(x, y) + F(x) - G(y) \quad (5)$$

Just as the case of ϕ we can prove for ϕ_t that it is a convex-concave function with a saddle point which is unique. We will show that this saddle point can be reached using newton decrements.

Proposition : For $t > 0$ and $(x_t, y_t) \in (X_0, Y_0)$, the Newton decrements $\eta(\phi_t, \hat{x}, \hat{y}) = 0$ if and only if $(\hat{x}, \hat{y})$ are the saddle points of ϕ_t

Proof: One can introduce dual variables $\lambda_t \in R^m, \delta_t^U, \delta_t^L \in R^n$ and $V_t^L, V_t^U, Z_t \in S^n$ for solving 5 with the conditions:

$$(Ax_t - b) \cdot_t = \mu_t e, \quad (6)$$

$$(c_t - c_L) \cdot d_t^L = \mu_t e, \quad (7)$$

$$(c_U - c_t) \cdot d_t^U = \mu_t e, \quad (8)$$

$$(Q_t - Q_L) \cdot V_t^L = \mu_t E, \quad (9)$$

$$(Q_U - Q_t) \cdot V_t^U = \mu_t E, \quad (10)$$

$$Q_t Z_t = \mu_t I \quad (11)$$

We also define for fixed $\hat{x}, \hat{y}$ the following functions:

$$\phi_{\hat{y}}^t = t\phi(x, \hat{y}) + F(x) \forall x \in X^0 \quad (12)$$

$$\phi_{\hat{x}}^t = t\phi(\hat{x}, y) - G(y) \forall y \in Y^0 \quad (13)$$

For the newton decrement $\eta(\phi_t, \hat{x}, \hat{y}$ to be zero, $\Delta\phi_{\hat{x}}^t = 0$ and $\Delta\phi_{\hat{y}}^t = 0$. Since $\phi_{\hat{y}}^t$ is a strictly convex function of x, and $-\phi_{\hat{x}}^t$ is a strictly convex function of y, it follows that $\hat{x}$ minimizes $\phi_{\hat{y}}^t$ and $\hat{y}$ maximizes $\phi_{\hat{x}}^t$. Therefore $(\hat{x}, \hat{y})$ must be a saddle point.On the other hand if $(\hat{x}, \hat{y})$, solves (6)-(11) we must have

$$\Delta\phi_{\hat{x}}^t = 0 and \Delta\phi_{\hat{y}}^t = 0$$

Which in turn implies the newton decreament is 0.This completes the proof.

So it has been established that taking newton decrements will lead to the central path, next it will be proved that being on the central path bounds the duality gap of the original problem (ϕ).

To Prove: If $(\bar{x}, \bar{y})$ satisfies $\eta(\phi_t, \bar{x}, \bar{y} \leq \beta$, with $beta \leq \frac{1}{2}$, the following inequality holds:

$$\nu(\bar{x}, \bar{y}) = f(\bar{x}) - g(\bar{y}) \leq (1 + \frac{6\beta}{\sqrt{n^2 + 4n + m}})(n^2 + 4n + m)\mu_t$$

Proof: Let $\phi_{\bar{y}} := \phi(x, \bar{y})$, where $\bar{y}$ is fixed. It has been shown in [3] Lemma 2.8 that there exists a unique minimizer of $\phi_{\bar{y}}^t$, let this be $\hat{x}^t$. Then using first order Taylor series expansion we can write

$$\begin{aligned} \phi(x, \bar{y}) &\geq \phi_{\bar{y}}(\hat{x}^t) + (x - \hat{x}^t)\Delta\phi_{\bar{y}}(\hat{x}^t) \\ &= \phi_{\bar{y}}(\hat{x}^t) + \mu_t (x - \hat{x}^t)^T \Delta F(\hat{x}^t) \end{aligned}$$

, the equality follows using conditions that must hold for $\bar{x}^t$ to be a minimizer. Since, F is a self-concordant barrier function with parameter m it must hold that,

$$(x - \hat{x}^t)^T \Delta F(\hat{x}^t) \leq m$$

We have,

$$\phi(x, \bar{y}) \geq \phi_{\bar{y}}(\hat{x}^t) + \mu_t m \forall x \in X^0$$

therefore ,

$$g(\bar{y}) = min_x \phi(x, \bar{y}) \geq \phi_{\bar{y}}(\hat{x}^t) + \mu_t m \quad (14)$$

Using self-concordance property of F we can write

$$|\nabla F(x)^T h| \leq \sqrt{m}\sqrt{h^T \nabla^2 F(x) h}$$

Thus

$$\begin{aligned} F(\hat{x}^t) - F(\bar{x}) &\leq (\bar{x} - \hat{x}^t)^T \nabla F(\hat{x}^t) \\ &\leq \sqrt{m}\sqrt{(\bar{x} - \hat{x}^t)^T \nabla^2 F(\hat{x}^t)(\bar{x} - \hat{x}^t)} \end{aligned}$$

In our case the newton decrement can be simplified to

$$\eta(\phi_t, x, y) := \sqrt{\eta^2(\phi_y^t, x) + \eta^2(-\phi_x^t, y)}$$

From the assumption $\eta(\phi^t, \bar{y}, \bar{x}) \leq \beta$, we have ,

$$\eta(\phi_{\bar{y}, \bar{x}}^t) \leq \beta_x and \eta(-\phi_{\bar{x}, \bar{y}}^t) \leq \beta_y$$

Using the fact that $\phi^t_{\bar{y}}$ is self-concordant and inequalities (2.16) and (2.17) from reference [2], we obtain the following

$$\phi^t_{\bar{y}}(\bar{x}) - \phi^t_{\bar{y}}(\hat{x}^t) \leq \beta_x$$

$$\sqrt{(\bar{x} - \hat{x}^t)^T \nabla^2 \phi^t_{\bar{y}}(\hat{x}^t)(\bar{x} - \hat{x}^t)} \leq \frac{\beta_x}{1 - \beta_x} \leq 2\beta_x$$

Combining the above two inequalities we get ,

$$\phi_{\bar{y}}(\bar{x}) - \phi_{\bar{y}}(\hat{x}^t) =$$

$$\mu_t(\phi^t_{\bar{y}}(\bar{x}) - \phi^t_{\bar{y}}(\hat{x}^t)) + \mu_t(F(\hat{x} - F(\bar{x}))$$

$$\leq \beta_x \mu_t + \sqrt{m}\sqrt{(\bar{x} - \hat{x}^t)^T \nabla^2 F(\hat{x}^t)(\bar{x} - \hat{x}^t)}\mu_t$$

$$\leq \beta_x \mu_t + \sqrt{m}\sqrt{(\bar{x} - \hat{x}^t)^T \nabla^2 \phi^t_{\bar{y}}(\hat{x}^t)(\bar{x} - \hat{x}^t)}\mu_t$$

$$\leq (1 + 2\sqrt{m})\beta_x \mu_t$$

Combining this inequality with 14 gives,

$$\phi(\bar{x}, \bar{y}) - g(\bar{y}) \leq m\mu_t + (1 + 2\sqrt{m})\beta_x \mu_t \tag{15}$$

Fixing $\bar{x}$ instead of $\bar{y}$ a symmetric argument yields,

$$f(\bar{x}) - \phi(\bar{x}, \bar{y}) \leq (1 + \frac{3\beta_y}{\sqrt{n^2 + 4n}}(n^2 + 4n)\mu_t \tag{16}$$

Combining 15 and 16, we get the desired result.

Thus,a measure of proximity to central path has been given and its relation with duality gap has been shown to be bounded.Hence, we can traverse along this path using newton decrements to get to a saddle point. The algorithm has been given in [3], but we state it here for the sake of completeness.

Algorithm:

- Choose α and β that satisfy the following equations:

$$\gamma := 1.3[(1 + \alpha)\beta + \alpha\sqrt{n^2 + 4n + m}] < 1,$$

$$\frac{\gamma^2(1 + \gamma)}{1 - \gamma} \leq 1.3\beta$$

  Ensure proximity to central path by selecting t_0, x_0, y_0 such that $\eta(\phi_{t_0}(x_0, y_0)) \leq \beta$ set k = 0
- Check the inequality

$$t_k \geq \frac{1}{\epsilon}(1 + \frac{6\beta}{n^2 + 4n + m})(n^2 + 4n + m)$$

  If satisfied stop, else continue
- $t_(k + 1) = (1 + \alpha)t_k$. Take newton step :

$$(x_{k+1}, yk + 1) =$$

$$(x_k, y_k) - [\nabla^2 \phi(t_{k+1})(x_k, y_k)]^{-1} \nabla \phi_(t_{k+1})(x_k, y_k)$$

  k = k + 1. Return to step 2.

This algorithm is proven to have polynomial complexity in [3].To get to ϵ accuracy of duality gap $O(\sqrt{n^2 + 4n + m}log(\frac{1}{\epsilon})$ iterations are needed. In this study methods to reduce runtime were used and variants of the original form were studied for very large scale problems.

D. Simulation

The above discussed convex-concave problem arises in famous Markovitz Portfolio optimization, the first term in the objective ($c^T x$) is the mean maximization, here $c = -profit$ and x is the weight vector for different options or stocks. The second term ($x^T Q x$) in the risk part with Q being the covariance of stocks. So in totality we want to allocate weights to stocks so as to maximize profit and minimize risk by allocating weights to diverse stocks. In this study, box constraints are put on c and Q to test the algorithm introduced in previous sections in the face of market unpredictability. Instead of using the sample covariance, a better estimate of Q was obtained as discussed Ledoit-Wolf in [2]. A look-back period of 256 was chosen for the estimation of c and Q, while re-balancing was done every 64 days.

The baseline chosen for comparison of results was simple Mean-Variance optimization, i.e the same formulation as above without box constraints. The CVXOPT package available for python was used to solve this. To ensure proper comparison, total booksize was set to 100 units (i.e long + short = 100). Gradient and Hessian were calculated as above. The algorithm provided was improved upon by the following :

- The algorithm in [3] used a single step in the inner loop (to get to the central path with ϕ^t).In this study single step convergence was compared with multiple steps.
- Storing the hessian has major memory needs as the number of stocks increase. Implementing the above will then be unrealistic. A work-around was tested which first selects uncorrelated and high performing stocks using Mean-Variance Optimization (MVO) and then the Robust algorithm was run on these
- Lastly, a lot of run-time was saved by estimating the newton step using least square regression.

The following results were obtained by taking a single step in the inner loop. 100 stocks were randomly chosen from SP top500 and 3 year data was used for the simulations.

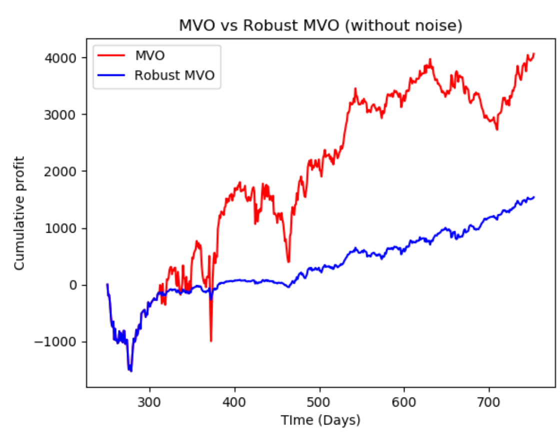

Fig. 1: Comparing performance of Robust algorithm with vanilla MVO without adding noise to stock prices.

Notice in the above figure that MVO gives a better profit compared Robust MVO but the latter is more conservative in does not have large draw-downs. Next, random noise of 2% of the current price was added to the stock prices and the performance evaluated.

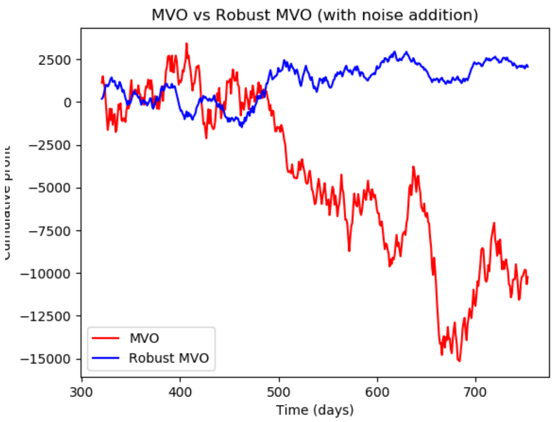

Fig. 2: Comparing performance of Robust algorithm with vanilla MVO with noise added to stock prices.

It can be seen that with uncertainties added Robust MVO maintains a slight profit whereas MVO breaks down. We obtained a bound on duality gap as one traverses the central path, this gap was based only on the dimension of the problem and the choice of initial parameters. To keep the central path close to saddle points of ϕ_t multiple steps were taken instead of just one as suggested in [3]. It was observed that taking multiple steps lets us take the path closer to saddle points.

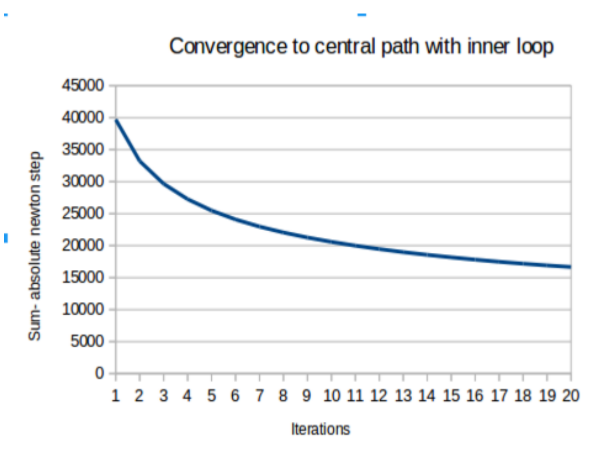

Fig. 3: Convergence closer to central path using inner loop

III. CONCLUSIONS

Robust programming is very promising for modelling uncertainties,and the current approach yielded good results, but it still lacks when it comes to these two major drawbacks : high memory need and very high dependence of results on the initial parameters. We tried to get rid of the second problem by taking multiple newton decrements before updating the value of t (variable of interior point method) but this was not a huge success. Future study goals can be to get better initialization constraints.

REFERENCES

[1] SIAM REVIEW Vol. 53, No. 3, pp. 464501 Theory and Applications of Robust Optimization Dimitris ,Bertsimas ,David B. Brown ,Constantine Caramanis

[2] Ledoit O., Wolf M.. Honey, I Shrunk the Sample Covariance Matrix:Problems in Mean-Variance Optimization, Journal of Portfolio Management , 2004, vol. 30 (pg. 110-19)

[3] Halldrsson, B.V. and R.H. Ttnc. (2003). An Interior-Point Method for a Class of Saddle-Point Prob- lems. Journal of Optimization Theory and Applications 116(3), 559590.

[4] Annals of Operations Research 132, 157187,2004.Robust Asset Allocation,R.H. TTNC,M. KOENIG