Journal of Advances and Scholarly Researches in Allied Education
Vol. 12, Issue No. 2, January-2017, ISSN 2230-7540

Analysis on Sub-Gradient and Semi-Definite Optimization

Shashi Sharma*

Mathematics Department, DAV College, Muzaffarnagar-251001

Abstract* – *Semidefinite Programming (SDP) is one of most fascinating parts of mathematical programming with regards to most recent twenty years. Semi definite Programming can be utilized to display numerous practical problems in different fields, for example, curved compelled optimization, combinatorial optimization, and control theory. The sub gradient method is a straightforward algorithm for limiting a no differentiable curved function. The method looks especially like the conventional inclination method for differentiable functions, however with a few eminent exemptions. The subgradient method is promptly reached out to handle problems with requirements. The most broadly known executions of SDP solvers are inside point methods. They give highaccuracy solutions in polynomial time. Sub gradient methods can be much slower than inside point methods (or Newton's method in the unconstrained case). Specifically, they are first-order methods; their exhibition depends particularly on the problem scaling and molding. In this Article, we studied about the Sub-Gradient and Semi-Definite Optimization in detail.

---------------------------------X---------------------------------

I. SUBGRADIENT METHODS

Subgradient methods are iterative methods for taking care of convex minimization problems. Initially created by Naum Z. Shor and others in the 1970s, subgradient methods are united when connected even to a non-differentiable target function. At the point when the target function is differentiable, sub-inclination methods for unconstrained problems utilize a similar hunt course as the method of steepest plunge. Subgradient methods are slower than Newton's method when connected to limit twice consistently differentiable convex functions. In any case, Newton's method neglects to meet on problems that have non-differentiable crimps. As of late, some inside point methods have been proposed for convex minimization problems, yet subgradient projection methods and related bundle methods of drop stay aggressive. For convex minimization problems with expansive number of measurements, subgradient-projection methods are appropriate, in light of the fact that they require little stockpiling. Subgradient projection methods are frequently connected to expansive scale problems with decay techniques. Such decay methods regularly permit a straightforward circulated method for an issue. The subgradient method was initially created by Shor in the Soviet Union in the 1970s. The essential reference on subgradient methods is his book. Another book on the theme is Akgul. Bertsekas is a decent reference on the subgradient method, joined with primal or double decay.

1.1 The Sub Gradient Method

Assume $f: Rn \rightarrow R$ is convex. To minimize f, the subgradient method uses the iteration

$$x^{(k+1)} = x^{(x)} - \alpha_{kg^{(k)}}$$

Here $x^{(k)}$ is the k^{th} iterate, $g^{(k)}$ is any subgradient of f at $x^{(k)}$, and $\alpha_k > 0$ is the k^{th} step size. In this way, at every cycle of the subgradient method, we make a stride toward a negative subgradient. Review that a subgradient of f at x is any vector g that fulfills the imbalance $f(y) \geq f(x) + gT\,(y - x)$ for all y. When f is differentiable, the only possible choice for $g^{(k)}$ is ∇f(x (k)), what's more, the subgradient method at that point diminishes to the gradient method (with the exception of, as we'll see beneath, for the decision of step estimate).

Since the subgradient method isn't a plunge method, usually to monitor the best point discovered up until this point, i.e., the one with littlest function esteem. At each step, we set

$$f_{best}^{(k)} = min\left\{f_{best}^{(k-1)}, \ldots, f\left(x^{(k)}\right)\right\},$$

and set $i_{best}^{(k)} = k\ if\ f\left(x^{(k)}\right) = f_{best}^{(k)}, i.e.\ if\ x^{(k)}$ is the best point discovered up until now. (In a descent method there is no compelling reason to do this —

the present point is dependably the best one up until this point.) Then we have

$$f_{best}^{(k)} = min\{f(x^{(1)}), \ldots, f(x^{(k)})\},$$

i.e., $f_{best}^{(k)}$ is the best objective value found in k iterations? Since $f_{best}^{(k)}$ is diminishing, it has an utmost (which can be −∞). (For comfort of later documentations, we mark the underlying point with 1 rather than 0.)

1.2 Step Size Rules

Several different types of step size rules are used.

- Constant step size. $\alpha_k = h$ is a constant, independent of k.
- Constant step length. $a_k = h/\|g^{(k)}\|$. This means that $\|x^{(k+1)} - x^{(k)}\|2 = h$
- Square summable but not summable. The step sizes fulfill

$$\sum_{k=1}^{\infty} a_k^2 < \infty, \qquad \sum_{k=1}^{\infty} a_k = \infty$$

One typical example is $\alpha_k = a/(b + k)$, where $a > 0$ and $b \geq 0$.

- Non-summable diminishing. The step sizes satisfy

$$\lim_{k\to\infty} a_k = 0, \qquad \sum_{k=1}^{\infty} a_k = \infty$$

Step sizes that fulfill this condition are called decreasing step estimate rules. A commonplace case is $\alpha k = a/\sqrt{k}, where\ a > 0.$

II. CONVERGENCE RESULTS

There are numerous outcomes on merging of the subgradient method. For steady step size and consistent step length, the subgradient algorithm is ensured to unite to inside some scope of the ideal esteem, i.e., we have

$$\lim_{k\to\infty} f_{best}^{(k)} - f^* < \in,$$

where f* denotes the optimal value of the problem, i.e., $f^* = \inf_x f(x)$. (This implies that the subgradient method finds an ²-suboptimal point within a finite number of steps.) The number ² is a function of the step estimate parameter h, and reductions with it.

For the reducing step estimate rule (and accordingly additionally the square summable however not summable step measure rule), the algorithm is ensured to merge to the optimal esteem, i.e., we have

$$limk \to \infty\ f(x(k)) = f*.$$

At the point when the function f is differentiable, we can say more in regards to the joining. For this situation, the subgradient method with consistent step estimate yields union to the optimal esteem, gave the parameter h is sufficiently small.

III. PROJECTED SUBGRADIENT METHOD

One augmentation of the sub gradient method is the anticipated sub gradient method, which takes care of the obliged convex optimization issue

Minimize f(x)

Subject to x ∈ C,

Where C is a convex set, the projected subgradient method is given by

$$x^{(k+1)} = P\left(x^{(k)} - a_{kg}{}^{(k)}\right),$$

Where P is (Euclidean) projection on C, and g (k) is any subgradient of f at $x^{(k)}$. The step measure rules portrayed before can be utilized here, with comparative joining comes about.

The merging evidences for the subgradient method are promptly reached out to deal with the anticipated sub gradient method.

Let $z^{(k+1)} = x^{(k)} - \alpha k g^{(k)}$, i.e., a standard subgradient update, before the projection back onto C. As in the subgradient method, we have

$$\| z^{(k+1)} - x^* \|_2^2 = \| x^{(k)} - a_{kg}{}^{(k)} - x^* \|_2^2$$

$$= \| x^{(k)} - x^* \|_2^2 - 2a_{kg}{}^{(k)T}\left(x^{(k)} - x^*\right) + a_k^2 \| g^{(k)} \|_2^2$$

$$\leq \| x^{(k)} - x^* \|_2^2 - 2a_k\left(f\left(x^{(k)}\right) - f^*\right) + a_k^2 \| g^{(k)} \|_2^2$$

Now we observe that

$$\| x^{(k+1)} - x^* \|_2 = \| P\left(z^{(k+1)}\right) - x^* \|_2 \leq \| z^{k+1} - x^* \|_2,$$

i.e., when we anticipate a point onto C, we draw nearer to each point in C, and specifically, any optimal point. Consolidating this with the imbalance above we get

$$\| x^{(k+1)} - x^* \|_2^2 \leq \| x^{(k)} - x^* \|_2^2 - 2_{\alpha k}\left(f\left(x^{(k)}\right) - f^*\right) + \alpha_k^2 \| g^{(k)} \|_{2'}^2$$

Furthermore, the evidence continues precisely as in the common subgradient method.

IV. SEMI DEFINITE OPTIMIZATION

Semi definite optimization is concerned with picking a symmetric matrix to enhance a linear function subject to linear requirements and a further critical imperative that the matrix be sure Semi definite. It in this manner emerges from the notable linear programming issue by supplanting the vector of variables with a symmetric matrix and supplanting the non-pessimism limitations with a positive Semi definite imperative. (An elective method to compose such an issue is as far as a vector of variables, with a linear target function and a limitation that some symmetric matrix that depends affinely on the variables be sure Semi definite.) This speculation by and by acquires a few critical properties from its vector partner: it is convex, has a rich duality hypothesis (in spite of the fact that not as solid as linear programming's), and concedes hypothetically efficient solution techniques in light of emphasizing inside focuses to either take after the focal way or reduction a potential function. Here we will research this class of problems and overview the ongoing outcomes and methods got.While linear programming (LP) as a subject became quick amid the '50s and '60s, because of the accessibility of the extremely efficient simplex method of G.B. Dantzig, Semi definite optimization (otherwise called Semi definite programming or SDP, the term we will utilize) was slower to pull in as much consideration. Halfway this was on account of, since the doable locale is never again polyhedral, the simple method was not material, albeit related methods do exist. When hypothetically efficient (and also basically valuable) algorithms ended up accessible in the late '80s and '90s, examine in the zone detonated. The ongoing Handbook of Semi definite Programming records 877 references, while the online list of sources on Semi definite programming gathered by Wolkowicz records 722, all since 1990 The advancement of efficient algorithms was just a single trigger of this dangerous development: another key inspiration was the intensity of SDP to demonstrate problems emerging in an extensive variety of regions. Bellman and Fan appear to have been the first to detail a Semi definite programming issue, in 1963. Rather than considering a linear programming issue in vector shape and supplanting the vector variable with a matrix variable, they began with a scalar LP plan and supplanted every scalar variable with a matrix. The subsequent issue (albeit equivalent to the general definition) was fairly lumbering, yet they determined a double issue and set up a few key duality hypotheses, demonstrating that extra normality is required in the SDP case to demonstrate solid duality. Nonetheless, the significance of imperatives requiring that a specific matrix be sure (semi)definite had been perceived substantially before in charge hypothesis: Lyapunov's portrayal of the steadiness of the solution of a linear differential condition in 1890 included simply such a limitation (called a linear matrix imbalance, or LMI). In the mid '70s, Donath and Hoffman and afterward Cullum, Donath, and Wolfe demonstrated that some hard diagram parceling problems could be assaulted by considering a related Eigen value optimization issue – as we will see, these are firmly associated with SDP. At that point in 1979, Lov'asz defined a SDP issue that gave a bound on the Shannon limit of a chart and accordingly found the limit of the pentagon, explaining a longopen guess. Around then, the most efficient method known for SDP problems was the ellipsoid method, and Grotschel, Lovasz, and Schrijver examined in detail its application to combinatorial optimization problems by utilizing it to inexact the solution of both LP and SDP relaxations. Lov'asz and Schrijver later indicated how SDP problems can give more tightly relaxations of (0, 1)- programming problems than can LP.

Fletcher resuscitated enthusiasm for SDP among nonlinear developers in the '80s, and this prompted a progression of papers by Overton and Overton and Womersley; and the references in that. The key commitments of Nesterov and Nemirovski and Alizadeh demonstrated that the new age of inside point methods spearheaded by Karmarkar for LP could be stretched out to SDP. Specifically, Nesterov and Nemirovski built up a general system for tackling nonlinear convex optimization problems in a hypothetically efficient manner utilizing inside point methods, by building up the ground-breaking hypothesis of self-concordant boundary functions. These works prompted the immense late enthusiasm for Semi definite programming, which was additionally expanded by the consequence of Goemans and Williamson which demonstrated that a SDP unwinding could give a provably decent estimation to the maximum cut issue in combinatorial optimization.

V. PROBLEMS

The SDP problem in primal standard form is

$$\begin{array}{rc} & C.X \\ (P) & min_x A_i.X = A_i, i = 1, \ldots m \\ & X \succeq 0, \end{array}$$

where all $A_i \in SIR^{n\times n}$, $b \in IR^m$, $C \in SIR^{n\times n}$ are given, and $X \in SIR^{n\times n}$ is the variable. We also consider SDP problems in dual standard form:

$$(D) \qquad max_{y,s} \qquad \sum_{i=1}^{m} \frac{b^T y}{y_i A_i} + \frac{S = C}{S \succeq 0}$$

Where $y \in IR^m$ and $S \in SIR^{n\times n}$ are the variables, this can also be written as

$$\max_{y} b^T y, \sum_{i=1}^{m} y_i A_i \preceq C$$

Or

$$\max_{y} b^T y, \; C - \sum_{i=1}^{m} y_i A_i \succeq 0$$

in any case, we will see the advantage of having the "slack matrix" S accessible when we examine algorithms. We ought to entirely express "inf" and "sup" rather than "min" and "max" above on the grounds that the problems may be unbounded, as well as in light of the fact that regardless of whether the optimal values are limited they won't not be achieved. We stick to "min" and "max" both to feature the way that we are keen on optimal solutions, not simply values, and on the grounds that we will frequently force conditions that guarantee that the optimal values are in certainty accomplished where limited. The last type of the issue in double standard shape demonstrates that we are endeavoring to enhance a linear function of a few variables, subject to the imperative that a symmetric matrix that depends affinely on the variables is limited to be certain Semi definite. (From this time forward, as is regular in scientific programming, we utilize "linear" to signify "relative" as a rule: in any case, linear administrators will dependably be linear, not relative.)

The restrictions of such methods are being lessened, and they have effectively tackled problems with lattices of request 10,000 and that's only the tip of the iceberg. One constraint is that these more efficient methods typically take care of the double issue, and if a primal close optimal solution is required (as in the maximum cut issue utilizing the strategy of Goemans and Williamson to produce a cut), they may not be as fitting. The point stays energizing and dynamic, and critical advancements can be normal throughout the following quite a long while.

VI. CONCLUSION

The underlying foundations of semidefinite programming can be followed back to both control theory and combinatorial optimization, just as the more classical research on optimization of matrix eigenvalues. We are blessed that numerous fantastic works managing the advancement and applications of SDP are accessible. A choice to smooth approximations is to utilize cutting plane techniques. At an abnormal state, a cutting plane strategy exploits the way that the sub-slope at each point characterizes a hyperplane which is a lower bound on the objective function (this is known as a cutting plane). In this way, by keeping various sub-angles around (from past emphasess), it is conceivable to build an inexorably precise lower headed gauge for the objective function. Obviously, keeping each sub-slope around is costly. Subgradient methods do have a few favorable circumstances over interior-point and Newton methods. They can be quickly connected to a far more extensive assortment of problems than interior-point or Newton methods. The memory prerequisite of subgradient methods can be a lot littler than an interior-point or Newton method, which implies it tends to be utilized for amazingly expansive problems for which interior-point or Newton methods can't be utilized. Morever, by joining the subgradient method with base or double deterioration techniques, it is once in a while conceivable to build up a basic appropriated algorithm for a problem. Regardless, subgradient methods are well worth thinking about.

REFERENCES

1. M. Akg¨ul.(Akg¨ul,2014.)Topics in Relaxation and Ellipsoidal Methods, volume 97 of Research Notes in Mathematics. Pitman,

2. P. Camerini, L. Fratta, and F. Maffioli. (Fratta & Maffioli 2015.) On improving relaxation methods by modifying gradient techniques. Math. Programming Study, pp. 26–34,

3. B. Polyak. (Polyak, 2007) Introduction to Optimization. Optimization Software, Inc.,

4. Y. Nesterov. (Nesteroy 2009) Primal-dual sub gradient methods for convex problems. Mathematical Programming, 120(1): pp. 221–259,

5. N. Shor. (Shor. 2008) Nondifferentiable Optimization and Polynomial Problems. Nonconvex Optimization and its Applications. Kluwer,

6. Christoph Helmberg, Franz Rendl, Robert J.(Helmerg & Robert, 2014) Vanderbei, and Henry Wolkowicz. An interior-point method for semidefinite programming. Technical report, Program in Statistics and Operations Research, Princeton University,

7. Jiashi Feng, Huan Xu, and Shuicheng Yan. (Huan xu & Shuicheng. 2013)Online robust pca via stochastic optimization. In NIPS,

8. Ming Gu. (Gu 2007) Primal-dual interior point methods for semidefinite programming in finite precision. SIAM J. Optimization, 10(2),

9. Michel X. Goemans and David P. Williamson. (Goemans & Williamson.2014) .879 approximationn algorithms for max cut and max 2sat. In STOC,

10. Yuri Nesterov. (Nesterov, 2014)Smoothing technique and its application in semidefinite optimization. CORE Discussion Paper,

---

Corresponding Author

Shashi Sharma*

Mathematics Department, DAV College, Muzaffarnagar-251001