Lecture 9

Waldemar Hebisch

December 14, 2021

1 Constrained optimization

Recall general optimization problem: given a set S and a function $f : S \to \mathbb{R}$ find $x_0 \in S$ such that

$$f(x_0) = \max_{x \in S} f(x).$$

Usually we have

$$S = \{x : \forall_{i \in E} g_i(x) = 0, \forall_{i \in I} g_i(x) \leq 0\}$$

where E is called set of equality constraints, I is called set of inequality constraints. When all equality constraints are linear, and f and all g_i are convex the problem is called convex optimization problem.

In convex optimization problem feasible set is convex and under reasonable regularity conditions is closed. Our definition of convex optimization problem requires constraints of special form, so when S is convex but constraints are not convex, then formally this is not a convex optimization problem. On the other hand, when S is convex and closed then

$$g(x) = \inf_{y \in S} \|x - y\|^2$$

is convex function such that $S = \{x : g(x) \leq 0\}$. So when feasible set is convex and closed and goal function is convex we can rewrite our problem as convex optimization problem. However, for g above gradient ∇g vanishes at boundary of S which may cause trouble. Also this g can be difficult to compute. Usually we want to take advantage of specific form of f and g_i, and sometimes we transform problem to simpler form.

Remark: It is important that in constrained problems optimal point may be at the boundary of set S. If we know that optimum is attained in the interior (for example when $f(x)$ goes to infinity when x approaches boundary of S), then we can mostly disregard constraints and use methods from previous lectures. But we need to be more careful when optimal point is at the boundary. In particular derivative criterion $\nabla f(x_\infty) = 0$ is no longer valid.

There are various forms of convex optimization problems of varying complexity

- linear problems: constraints and goal function are affine
- quadratic problems: constraints are affine, goal function is convex quadratic
- quadratically constrained quadratic problems: constraints and goal function are convex quadratic
- second order cone problems
- semi-definite problems
- cone problems

Why so many different forms? Each has specific properties that makes it simpler than more general form. If we can recognize problem as some of simpler forms, then we can take advantage of this form and use more efficient solution method. In particular, up to semi-definite problems such forms have relatively simple self-concordant barrier function.

Second order cone problems have affine goal function and constraints of form

$$\|Ax + b\| \leq \langle c, x \rangle + d.$$

If needed adding extra variables we can transform convex quadratic constraints into second order cone constraints. Also, by adding extra variable t and constraint $f - t \leq 0$ we can transform problem into problem with linear goal function. So in this sense second order cone problems are more general than quadratically constrained quadratic problems.

In linear programing problems special role is played by $C = \mathbb{R}^n_+$. C is a convex cone and linear constraints may be written as

$$Ax + b \in C.$$

In cone problems we replace C by more arbitrary closed convex cone possibly in different space. And each constraint may use its own cone. In semi-definite problems we use cone of positive definite matrices. Note that

$$\|x\| \leq t$$

is equivalent to

$$0 \leq \begin{pmatrix} tI & x \\ x^T & t \end{pmatrix}$$

so we can rewrite second order cone constraints as semi-definite constraints.

Example: Basis pursuit problem: minimize $\|x\|_1$ under constraint $Ax = b$ can be transformed to linear problem by writing $x_i = u_i - w_i$ and minimizing $\sum(u_i + w_i)$.

Example: LASSO problem: minimize $\|Ax - b\|^2$ under constraint $\|x\|_1 \leq t$ can be transformed into quadratic problem (again writing $x_i = u_i - w_i$ and using $\sum(u_i + w_i) \leq t$ and $u_i \geq 0$, $w_i \geq 0$ as constraints).

Example: SVM lead to quadratic problem.

Example: Consider differentiable $f : C \mapsto \mathbb{R}$ where $C \subset \mathbb{R}^n$ (most interesting case is $n = 2$). Surface of the graph of f is

$$\int_C \|(f'(x), 1)\|_2 dx$$

which is convex (quadratic) infinite dimensional problem. Taking f from finite dimensional space (say via expression on finite grid) leads to finite dimensional quadratic problem. We can put constraints on f, say limiting maximal value, requesting specific values at boundary, etc.

Example: Limiting maximal eigenvalue of symmetric matrix A by M leads to constraint

$$A \leq MI$$

Example: Limiting norm of A by t leads to constraint

$$0 \leq \begin{pmatrix} tI & A \\ A^T & tI \end{pmatrix}$$

Non-convex problem lead to various difficulties. We say that vector v is tangent to S at x if and only if

$$\liminf_{t \to 0_+} \frac{d(x + tv, S)}{t} = 0.$$

Set of tangent vectors is a cone (is closed under multiplication by positive numbers). We call it tangent cone and denote it by TS_x.

Remark: This definition is rather general. For regular sets one could use more restrictive definition, replacing lim inf by lim. In fact, for for convex set S we can using lim we get the same set TS_x, but there are differences in less regular cases.

Remark: When $S_1 = \{(t, 0) : t \geq 0\}$, $S_2 = \{(0, t) : t \geq 0\}$, $S = S_1 \cup S_2$, $x = (0, 0)$ then $TS_x = S$. So TS_x may be non-convex.

Example: Let $g(x) = x_1^2 + x_2^2 - 1$ and $S = \{x \in \mathbb{R}^2 : g(x) = 0\}$. Of course this is just unit circle in the plane. Tangent vectors at $x = (1, 0)$ are of form $(0, s)$, but $(1, 0) + (0, s) = (1, s) \notin S$. In this case tangent cone is just a line.

Example: Let $S = \mathbb{R}^2_+$ and $x = (0, 0)$. Then set of tangent vectors at x is just S. Let $f(y) = y_1 + y_2$. Then f attains minimum at x, but $\nabla f(x)$ is nonzero. However, for all $v \in TS_x$ we have $\langle \nabla f(x), v\rangle \geq 0$.

Example: Let $S = \{(0, 1/n) : n > 0 \text{ and } n \in \mathbb{Z}\}$ where $\mathbb{Z}$ denotes integer numbers. Then $TS_0 = \{(0, t) : t > 0 \text{ and } t \in \mathbb{R}\}$.

Example: Let $S = \{(0, 2^{-n}) : n \in \mathbb{Z}\}$. Then TS_0 is the same as previous example. Definition with lim would give $TS_0 = \{0\}$.

We have the following general optimality condition:

Lemma 1.1 *If f is differentiable at x and $x = \operatorname{argmin}_{y \in S} f(y)$, then for all $v \in TS_x$ we have*

$$\langle \nabla f(x), v\rangle \geq 0$$

Proof: Fix $v \in TS_x$. By definition of $\liminf$ there exists sequence of $t_i \geq 0$ going to 0 such that $d(x + t_i v, S) = o(t_i)$. By definition of distance there exists $y_i \in S$ such that $d(x + t_i v, y_i) \leq 2d(x + t_i v, S)$. By definition of derivative

$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + o(d(x, y)).$$

Put $x_i = x + t_i v$. We have $d(x, x_i) = O(t_i)$ and $d(x_i, S) = o(t_i)$, so $d(y_i, x_i) = o(t_i)$ and $d(y_i, x) = O(t_i)$. Consequently

$$\langle \nabla f(x), y_i - x_i \rangle = o(t_i)$$

and

$$\begin{aligned} f(y_i) - f(x) &= \langle \nabla f(x), y_i - x \rangle + o(t_i) \\ &= \langle \nabla f(x), x_i - x \rangle + o(t_i). \end{aligned}$$

Now

$$\begin{aligned} \langle \nabla f(x), v \rangle &= \lim_{i \to \infty} \frac{\langle \nabla f(x), x_i - x \rangle}{t_i} \\ &= \lim_{i \to \infty} \frac{f(y_i) - f(x)}{t_i} \geq 0. \end{aligned}$$

The last inequality follows because $f(y_i) \geq f(x)$ so all terms are nonnegative. □

Remark: Above we could use weaker definition of derivative. We will come back to this when discussing subgradient.

Now we need to describe TS_x. For this we need notion of active constraints: constraint $g_i \leq 0$ is called active at x when $g_i(x) = 0$.

Lemma 1.2 *If g_i are differentiable at x then*

$$TS_x \subset \{v : \forall_{i \in E} \langle \nabla g_i(x), v \rangle = 0, \quad \forall_{i \in J} \langle \nabla g_i(x), v \rangle \leq 0\}$$

where J is set of active inequality constraints at x. If additionally g_i have continuous derivative and matrix formed from $\nabla g_i(x)$, $i \in E \cup J$ has full rank, then inclusion above is equality.

Outline of the proof: The first part follows from optimality lemma: for each $i \in E \cup J$ function $-g_i$ attains minimal value over S at x. The second part follows from inverse function theorem: after change of variables all constraints are linear and result is obvious. □

To write optimality conditions in nicer form we need notion of dual cone. When C is a convex cone, then

$$C^* = \{x : \forall_{y \in C} \langle x, y \rangle \geq 0\}$$

is called dual cone. It is a closed convex cone.

Lemma 1.3 *If C is convex cone, then*

$$(C^*)^* = \bar{C}$$

where $\bar{C}$ is closure of C.

Proof: Clearly $\bar{C} \subset (C^*)^*$. By separating hyperplane lemma if $y \in (C^*)^* - \bar{C}$, then there is v and b such that

$$\langle y, v\rangle < b$$

and

$$\forall_{x \in \bar{C}} \langle x, v\rangle \geq b.$$

Since $\bar{C}$ is a cone this means $b \leq 0$. Also, the inequality above remains valid for $b = 0$. But then $v \in C^*$, so first condition gives contradiction with $y \in (C^*)^*$, so no such y can exist. □

Now, let C be convex cone spanned by $\nabla g_i(x), -\nabla g_i(x)$ for $i \in E$ and $-\nabla g_i(x)$ for $i \in J$. Under assumptions of previous lemma we have

$$TS_x = C^*.$$

Our optimality condition can be written as

$$\nabla f(x) \in (TS_x)^* = (C^*)^* = C.$$

In this way we obtained Karush-Kuhn-Tucker (KKT) optimality conditions:

Lemma 1.4 *If g_i have continuous derivative and matrix formed from $\nabla g_i(x)$, $i \in E \cup J$ has full rank, then there exist $\lambda_i \in \mathbb{R}$, such that for $i \in J$ we have $\lambda_i \geq 0$ and*

$$\nabla f(x) - \sum_{i \in E} \lambda_i \nabla g_i(x) + \sum_{i \in J} \lambda_i \nabla g_i(x) = 0.$$

Example: Entropy $f(x) = -\sum_{i=1}^n x_i \log(x_i)$ with constraints $g(x) = \sum_{i=1}^n x_i = 1$, $x_i \geq 0$. When $x_i = 0$ corresponding term in f is 0, so we expect those constraints to be inactive. Then

$$\nabla g(x) = (1, \ldots, 1),$$

$$\nabla f(x) = -(\log(x_1) + 1), \log(x_2) + 1), \ldots, \log(x_n) + 1)$$

Since ∇g is always nonzero, with g as the only constraint KKT conditions hold so $\log(x_i) + 1 = \lambda$ is independent of i. Consequently $x_1 = x_2 = \cdots = x_n = \frac{1}{n}$ and value is

$$-n\frac{1}{n}\log(\frac{1}{n}) = \log(n).$$

Making $x_i = 0$ active effectively decreases n, so indeed this is optimal value.

Example: minimize $2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2$ subject to $x_1^2 + x_2^2 \le 5$, $3x_1 + x_2 \le 6$. We have

$$\nabla f(x) = (4x_1 + 2x_2 - 10, 2x_1 + 2x_2 - 10).$$

Solving for $\nabla f(x) = 0$ gives $x = (0, 5)$ which does not satisfy the first constraint. With one constraint (first or second) active assumption of the lemma are satisfied, so KKT equations hold. When only second constraint is active we get $x = (\frac{2}{5}, \frac{24}{5})$ which again does not satisfy first constraint (and has negative λ). When only first constraint is active we get two complex solutions, one solution with negative λ and $x = (1, 2)$ with $\lambda = 1$. This time second constraint is satisfied, so this is candidate solution giving $f(x) = -20$. There are also two points where both constraints are active. Checking values of f we see that both give bigger value than -20, so $(1, 2)$ is optimal point.

From example it should be clear that large number of inequality constraints leads to combinatorial explosion of cases. We need better methods than directly solving KKT conditions.

Methods for convex problems:

- conditional gradient
- projected gradient
- penalty methods
- barrier methods

Conditional gradient algorithm (also called Frank-Wolfe algorithm) is given by formulas:

$$d_i = \operatorname{argmin}_{z \in S} \langle \nabla f(x_i), z - x_i \rangle,$$

$$x_{i+1} = x_i + \alpha_i (d_i - x_i).$$

α_i can be determined by line search or prescribed, for example $\alpha_i = \frac{2}{i+2}$ (goes to 0 when i goes to ∞).

- since f is convex, when $d_i = x_i$ we are at optimum
- generates feasible points
- does not need projection
- requires optimization with linear objective to find d_i

Lemma 1.5 *If f is convex and ∇f is Lipschitz continuous with constant M, x_∞ is optimal point, $\alpha_i = 2/(i+2)$, then for $i > 0$*

$$f(x_i) - f(x_\infty) \le \frac{2MD^2}{i+2}$$

where $D = \sup_{x,y \in S} \|x - y\|$.

Proof: By definition of d_i we have

$$\langle \nabla f(x_i), x_\infty - x_i \rangle \geq \langle \nabla f(x_i), d_i - x_i \rangle.$$

By convexity of f

$$f(x_\infty) \geq f(x_i) + \langle \nabla f(x_i), x_\infty - x_i \rangle \geq f(x_i) + \langle \nabla f(x_i), d_i - x_i \rangle.$$

Since ∇f is Lipschitz continuous we have decay estimate

$$f(x_{i+1}) \leq f(x_i) + \alpha_i \langle \nabla f(x_i), d_i - x_i \rangle + \alpha_i^2 \frac{MD^2}{2}.$$

Subtracting estimate for $f(x_\infty)$ we get

$$f(x_{i+1}) - f(x_\infty) \leq (\alpha_i - 1)\langle \nabla f(x_i), d_i - x_i \rangle + \alpha_i^2 \frac{MD^2}{2}.$$

When $i = 0$ we have $\alpha_0 = 1$ so

$$f(x_1) - f(x_\infty) \leq \frac{MD^2}{2} \leq \frac{2MD^2}{1+2}$$

so we get the result for x_1. For larger i we use induction. Subtracting from decay estimate α_i times estimate for $f(x_\infty)$ we get

$$f(x_{i+1}) - \alpha_i f(x_\infty) \leq (1 - \alpha_i) f(x_i) + \alpha_i^2 \frac{MD^2}{2}$$

so

$$f(x_{i+1}) - f(x_\infty) \leq (1 - \alpha_i)(f(x_i) - f(x_\infty)) + \alpha_i^2 \frac{MD^2}{2}.$$

By inductive assumption

$$f(x_i) - f(x_\infty) \leq \frac{2MD^2}{i+2}$$

so

$$\begin{aligned} f(x_{i+1}) - f(x_\infty) &\leq (1 - \alpha_i)\frac{2MD^2}{i+2} + \alpha_i^2 \frac{MD^2}{2} \\ &= (1 - \frac{2}{i+2})\frac{2MD^2}{i+2} + \frac{4}{(i+2)^2}\frac{MD^2}{2} \\ &= \frac{i(i+3)}{(i+2)^2}\frac{2MD^2}{i+3} + \frac{(i+3)}{(i+2)^2}\frac{2MD^2}{i+3} \end{aligned}$$

However

$$\frac{i(i+3)}{(i+2)^2} + \frac{(i+3)}{(i+2)^2} = \frac{i^2 + 3i + i + 3}{(i+2)^2} = \frac{(i+2)^2 - 1}{(i+2)^2} \leq 1$$

so we have

$$f(x_{i+1}) - f(x_\infty) \le \frac{2MD^2}{(i+1)+2}$$

which ends inductive proof. □

Remark: Can not get better rate for strongly convex f.
Remark: Exact line search will do the same or better.

1.1 Projection onto convex set

For the next algorithm we need notion of projection:

$$\mathrm{Proj}_S(x) = \mathrm{argmin}_{z\in S} \|z - x\|$$

Since set of z such that $\|z - x\| \le t$ is compact, minimum above is attained when S is closed, so Proj_S is well defined (a priori definition involves choice). We have

Lemma 1.6 *For convex S*

$$\|\mathrm{Proj}_S(x) - \mathrm{Proj}_S(y)\| \le \|x - y\|.$$

We will prove this later considering proximal methods.
Remark: In particular, lemma implies that projection onto convex set is uniquely defined (no need for choice).
Example: When S is a hyperplane we get usual orthogonal projection.
Example: When S is a halfspace, then $\mathrm{Proj}_S(x) = x$ for $x \in S$, otherwise it is projection onto boundary hyperplane.
Example: When S is euclidean unit ball, then

$$\mathrm{Proj}_S(x) = \begin{cases} x & \text{when } \|x\| \le 1 \\ \frac{x}{\|x\|} & \text{otherwise} \end{cases}$$

Example: When S is l^∞ unit ball, then

$$\mathrm{Proj}_S(x)_i = \begin{cases} x_i & \text{when } |x_i| \le 1 \\ \mathrm{sign}(x_i) & \text{otherwise} \end{cases}$$

where $\mathrm{sign}(x_i)$ is sign of x_i.

1.2 Projected gradient algorithm

Projected gradient algorithm is given by formula:

$$x_{i+1} = \mathrm{Proj}_S(x_i - \alpha_i \nabla f(x_i))$$

where Proj_S denotes projection onto S.

- thanks to projection generates feasible points
- reasonably good convergence properties
- projection may be expensive to compute

Lemma 1.7 *Assume that* $mI \leq \nabla^2 f \leq MI$ *and optimal point* x_∞ *belongs to interior of* S. *Then for projected gradient descent with constant step size* $\alpha = \frac{2}{M+m}$ *we have*

$$\|x_i - x_\infty\| \leq C^i \|x_0 - x_\infty\|$$

where

$$C = \frac{M-m}{M+m}$$

Remark: The same as in unconstrained case.

Proof: Put $y_{i+1} = x_i - \alpha \nabla f(x_i)$. Proof in unconstrained case shows bound for $\|y_i - x_\infty\|$. Then nonexpansiveness of projection gives result.

Without assumption about x_∞ we get slightly weaker result, using $\alpha_i = \frac{1}{M}$ and $C = (1 - \frac{1}{M})$.

We also get convergence for convex (non necessarily strongly convex) f. We will prove all this considering proximal methods. We can apply Nesterov acceleration to get faster convergence.

1.3 Penalty methods

Instead of minimizing f in penalty methods we optimize

$$f(x) + \lambda h(x)$$

where $h(x) = 0$ for $x \in S$ and $h(x) > 0$ for $x \notin S$ for sequence of $\lambda > 0$ going to infinity. h above is called penalty function.

One possible choice for h is

$$h(x) = \sum_{i \in E} g_i^2(x) + \sum_{i \in I} (g_i)_+^2(x)$$

where $(g_i)_+$ is positive part of g_i, that is $(g_i)_+(x) = g_i(x)$ when $g_i(x) > 0$ and $(g_i)_+(x) = 0$ otherwise.

For convex S another possibility is

$$h(x) = \min_{z \in S} \|z - x\|^2 = \|x - \mathrm{Proj}_S(x)\|^2.$$

In general $f(x) + \lambda h(x)$ may fail to have minimal point, but under reasonable conditions for large λ

$$x_\lambda = \mathrm{argmin}_x (f(x) + \lambda h(x))$$

is well defined and converges to feasible point x_∞.

Lemma 1.8 *When f is continuous and*

$$x_\lambda \to x_\infty \in S,$$

then

$$x_\infty = \text{argmin}_{x \in S} f(x)$$

Proof: Fix $\varepsilon > 0$. Since f is continuous and $x_\lambda \to x_\infty$ for large λ we have

$$f(x_\lambda) \geq f(x_\infty) + \varepsilon$$

By definition of x_λ we have

$$\inf_{x \in S} f(x) \geq f(x_\lambda) + \lambda h(x_\lambda) \geq f(x_\lambda)$$

so

$$\inf_{x \in S} f(x) \geq f(x_\lambda) \geq f(x_\infty) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary we have

$$\inf_{x \in S} f(x) \geq f(x_\infty).$$

But $x_\infty \in S$ so we have equality

$$\min_{x \in S} f(x) = f(x_\infty).$$

□

Features of penalty methods:

- f needs to be defined on whole $\mathbb{R}^n$ (or at least open neighbourhood of S)
- can use efficient unconstrained method like Newton method
- in principle $f(x) + \lambda h(x)$ becomes badly conditioned for large λ, not a problem when λ is gradually increased and we use Newton method (or preconditioning)

1.4 Barrier methods

Instead of minimizing f in barrier methods we optimize

$$f(x) + \lambda h(x)$$

where $h(x) \geq 0$ and $h(x)$ goes to infinity when x goes to boundary of S for sequence of $\lambda > 0$ going to 0. h above is called barrier function. In pure barrier methods we only allow inequality constraints.

Popular choice of barrier function h are

$$h(x) = \sum_{i \in I} \frac{1}{g_i^2(x)}$$

and

$$h(x) = -\sum_{i \in I} \log(-g_i(x)).$$

It has many advantages to use self-concordant barrier function, for many important convex sets S such barriers are easy to construct.

Lemma 1.9 *If f is continous, h is finite in interior of S, S is equal to closure of its interior,*

$$x_\lambda = \mathrm{argmin}_{x \in S}(f(x) + h(x)),$$

$$x_\lambda \to x_\infty,$$

then

$$x_\infty = \mathrm{argmin}_{x \in S} f(x).$$

Proof: Fix $\varepsilon > 0$. By definition of inf there exists $y \in S$ such that

$$f(y) \leq \inf_{x \in S} f(x) + \varepsilon.$$

y may be on boundary of S, but since f is continuous and S is closure of its interor there exists z in interior of S such that

$$f(z) \leq \inf_{x \in S} f(x) + 2\varepsilon.$$

Since $h(z)$ is finite for small λ we have

$$f(z) + \lambda h(z) \leq \inf_{x \in S} f(x) + 3\varepsilon$$

so

$$\begin{aligned} f(x_\lambda) \leq f(x_\lambda) + \lambda h(x_\lambda) &= \mathrm{argmin}_{x \in S}(f(x) + h(x)) \\ &\leq f(z) + \lambda h(z) \leq \inf_{x \in S} f(x) + 3\varepsilon. \end{aligned}$$

Since $x_\lambda \to x_\infty$ and f is continous also

$$f(x_\infty) \leq \inf_{x \in S} f(x) + 3\varepsilon.$$

Since $\varepsilon > 0$ were arbitrary we have

$$f(x_\infty) \leq \inf_{x \in S} f(x)$$

so

$$f(x_\infty) = \min_{x \in S} f(x).$$

□

Features of barrier methods:

- need feasible starting point
- goes only trough feasible points
- can use efficient unconstrained method like Newton method
- barrier function is typically badly conditioned in classical sense
- self-concordant barriers have very good convergence properties with Newton method

1.5 Interior point methods

Simplex method solves linear programming problem by moving trough boundary points of feasible set. Another possibility is to find sequence of points in the interior of feasible set that converges to optimal solution. This happens in barrier methods (which are now dominant form of interior point methods). For linear constraints it is usual to use logarithmic barrier. More precisely, we replace inequality constraints $x \geq 0$ by barrier

$$\phi(x) = -\sum_{i=1}^{m} \log(x_i)$$

Equality constraints remain (and must be treated separately). For $\lambda > 0$ we get problem of minimizing

$$f(x) + \lambda\phi(x)$$

This is equivalent to minimizing

$$\frac{1}{\lambda} f(x) + \phi(x)$$

Since f is affine (as we have linear programming problem), function above is self-concordant, so we have good convergence properties for Newton method. We will say more about this later. However, before doing this we will look into duality.

1.6 Further reading

Stephen Boyd, Lieven Vandenberghe, Convex Optimization, chapters 4, 8.1, 10, 11.

David G. Luenberger, Yinyu Ye, Linear and Nonlinear Programming, chapters 11, 12, 13.

Jorge Nocedal, Stephen J. Wright, Numerical Optimization, chapter 12.