Algorithms

Newton's Method - quadratic convergence (Unconstrained)

given a starting point $x \in \mathrm{dom}\, f$, tolerance $\epsilon > 0$

repeat

1) Compute Newton step and decrement.
$$\Delta x_{nt} = -\nabla^2 f(x)^{-1} \nabla f(x) \qquad \lambda^2 = \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)$$
2) Stopping criterion : quit if $\lambda^2/2 \le \epsilon$
3) Line search : Choose step size t by backtracking line search
4) Update $x = x + t\Delta x_{nt}$

Remember : if f is linearly separable (e.g $-\sum \log x$) then the Hessian is equivalent to the diagonal of the function : $-\nabla^2 f(x) = -\mathrm{diag}(f_{i-n})$.

Newton's Method for Equality-Constrained Minimization (Unconstrained)

given starting point $x \in \mathrm{dom}\, f$ with $Ax = b$, tolerance $\epsilon > 0$,

repeat

1) Compute Newton step and decrement Δx_{nt}, $\lambda(x)$
2) Stopping criterion. quit if $\lambda^2/2 \le \epsilon$
3) Line search. Choose t by backtracking
4) Update $x = x + t\Delta x_{nt}$

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

$$\lambda(x) = (\Delta x_{nt}^T \nabla^2 f(x) \Delta x_{nt})^{1/2}$$

Infeasible Start Newton's Method (Equality-constrained algorithm)

given starting point $x \in \mathrm{dom}\, f$, v, tolerance $\epsilon > 0$, $\alpha \in (0, .5)$, $\beta \in (0,1)$

repeat

1) Compute primal and dual Newton steps Δx_{pd} and Δv_{pd}

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{pd} \\ \Delta v_{pd} \end{bmatrix} = -\begin{bmatrix} \nabla f(x) + A^T v \\ Ax - b \end{bmatrix}$$

r_{dual} ; r_{primal}

2) Backtracking line search on $\|r\|_2$:
   $t = 1$
   while $\|r(x + t\Delta x_{pd}, v + t\Delta v_{pd})\|_2 > (1 - \alpha t)\|r(x, v)\|_2$
   $t = \beta \cdot t$
3) Update. $x = x + t\Delta x_{pd}$, $v = v + t\Delta v_{pd}$

until $Ax = b$ and $\underbrace{\|r(x,v)\|_2 \le \epsilon}$

Notice we use the residual term, not the value of f.

The option to switch to a feasible Newton's Method is always available once the iteration start is feasible again

Solving KKT System by Block Elimination

given KKT system with $H > 0$,

1) Form $H^{-1}A^T$ and $H^{-1}g$ ⟶ Cholesky factorization of H.
2) Form Schur compliment $S = -AH^{-1}A^T$
3) Determine w by solving $Sw = AH^{-1}g - h$ ⟶ Cholesky factorization of $-S$
4) Determine v by solving $Hv = -A^T w - g$

Gradient Descent Method - linear convergence (Unconstrained)

— Algorithms

Given a starting point $x \in \text{dom} f$:

repeat

1) $\Delta x = -\nabla f(x)$
2) line search. Choose step size t via exact or backtracking line search ↳ below.
3) update. $x = x + t\Delta x$

until stopping criterion is satisfied

$\|\nabla f(x)\|_2 \leq \eta$. Usually checked after step 1.

If using backtracking, method works better with large α. $[0.2 - 0.5]$ and mid β $[\approx 0.5]$

If the condition number is > 1000, the gradient method is pretty useless. i.e. if $H \approx I$, then the descent should work well

Backtracking Line Search Method

given a descent direction Δx for f at $x \in \text{dom} f$, $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$, $t = 1$

while $f(x + t\Delta x) > f(x) + \alpha t \nabla f(x)^T \Delta x$,

$t = \beta t$

|| i.e exit condition: $f(t) \leq f(x) - \alpha t \|\nabla f(x)\|_2^2$

Exact Line Search

t is chosen to minimize f along the ray $\{x + t\Delta x \mid t \geq 0\}$

$$t = \underset{s \geq 0}{\text{argmin}} \; \cancel{f(x)} \; f(x + s\Delta x)$$

Exact line search is used when the cost of the minimization problem with one variable is low compared to computing the search direction itself.

Steepest Descent Method - linear convergence (Unconstrained)

Given a starting point $x \in \text{dom} f$:

repeat

1) Compute steepest descent direction $\Delta x_{sd} = \|\nabla f(x)\|_* \Delta x_{nsd}$; $\Delta x_{nsd} = \underbrace{\text{argmin} \{\nabla f(x)^T v \mid \|v\| \leq 1\}}_{\text{geometric problem}}$
2) Line search for t.
3) update. $x = x + t\Delta x_{sd}$

until stopping criterion is satisfied.

Notes: if Euclidean norm is used, then $\Delta x_{sd} = -\nabla f(x)$, thus coinciding with the gradient descent method.

if quadratic norm: $\|z\|_P = (z^T P z)^{1/2} = \|P^{1/2} z\|_2$

$$\Rightarrow \Delta x_{nsd} = -(\nabla f(x)^T P^{-1} \nabla f(x))^{-1/2} P^{-1} \nabla f(x) \; ; \text{dual norm} = \|z\|_* = \|P^{-1/2} z\|_2$$

$$\Rightarrow \Delta x_{sd} = -P^{-1} \nabla f(x)$$

if l_1 norm: $\Delta x_{nsd} = \text{argmin} \{\nabla f(x)^T v \mid \|v\|_1 \leq 1\}$

$$\therefore \Delta x_{sd} = \Delta x_{nsd} \|\nabla f(x)\|_\infty = \frac{-\partial f(x)}{\partial x_i} e_i$$

Sometimes called a coordinate descent algorithm, since we minimize one coordinate at a time. Can trivialize line search process.

Barrier Method (inequality constraints)

Algorithms

given strictly feasible x, $t = t^{(0)} > 0$, $\underbrace{\mu > 1}_{\text{usually around 10 or 20.}}$, tolerance $\epsilon > 0$

repeat

1) Centering Step
   compute $x^*(t)$ by minimizing $t f_0 + \phi$, subject to $Ax = b$, starting at x. } initial centering step, usually Newton's method
2) Update $x = x(t)$
3) Stopping Criterion: quit if $m/t < \epsilon$
4) Increase t : $t = \mu t$

} outer iteration. Dual feasible points are only retrievable here.

Note: Infeasible Start Newton's Method can be used in Centering Step.

Primal-Dual Interior-Point Method

given x that satisfies $f_1(x) < 0, \ldots, f_m(x) \leq 0$, $\lambda > 0$, $\mu > 1$, $\underbrace{\epsilon_{feas} > 0,\ \epsilon > 0}_{\text{Great convergence, so choose small.}}$

repeat

1) Determine t. Set $t = \mu m / \hat{\eta}$
2) Compute primal-dual search direction Δy_{pd}
3) Line search & update
   Determine step length $s > 0$ and set $y = y + s \Delta y_{pd}$.

until $\|r_{pri}\|_2 \leq \epsilon_{feas}$, $\|r_{dual}\|_2 \leq \epsilon_{feas}$, and $\hat{\eta} \leq \epsilon$.

9 Unconstrained Minimization

General Problem: $\min \frac{1}{2}x^T P x + q^T x + r$

$\therefore$ optimality condition: $Px^* + q = 0$

when $P \succ 0$, there is a unique solution, $x^* = -P^{-1}q$. If not, then any solution of $Px^* = -q$ is optimal. No solution $\rightarrow$ unbounded below. This methodology is the basis for Newton's method.

Special Case of General:

$$\min \|Ax-b\|_2^2 = x^T(A^TA)x - 2(A^Tb)^Tx + b^Tb$$

optimality condition: $A^TAx^* = A^Tb$ } Normal Equations of least-squares problem

Unconstrained geometric programming:

$$\min f(x) = \log\left(\sum_{i=1}^m \exp(a_i^T x + b_i)\right)$$

optimality condition: $\frac{1}{\sum \exp(a_i^T x + b_i)} \sum_{i=1}^m \exp(a_i^T x + b_i) a_i = 0$ } No analytical solution so we need an iterative algorithm.

All points in $\mathbb{R}^n$ are valid, so any point can be chosen for $x^{(0)}$

Analytic Center

Consider $\min f(x) = -\underbrace{\sum_{i=1}^m \log(b_i - a_i^T x)}_{\text{cannot be negative}}$; $\mathrm{dom}\, f = \{x \mid a_i^T x < b_i,\ i = 1\ldots m\}$

Thus, the objective function here is the logarithmic barrier for $a_i^T x \le b_i$. The solution x^* is called the analytic center of these inequalities. Obviously, $x^{(0)}$ must satisfy the strict inequality $a_i^T x < b_i$

This can be translated into matrix form:

$$\min f(x) = \log \det F(x)^{-1}, \quad F(x) = F_0 + x_1F_1 + \ldots x_mF_m \text{ i.e is affine.}$$

$\therefore \mathrm{dom}\, f = \{x \mid F(x) \succ 0\}$

All conditions above apply.

Condition # of sublevel sets.

We can show that $\nabla^2 f(x)$ is bounded above and below by $mI \le \nabla^2 f(x) \le MI$, $m > 0$, $M > 0$, $M > m$.

The ratio $k = \frac{M}{m}$ is an upper bound on the ratio of the largest eigenvalue to the smallest. This can also be interpreted geometrically:

$$W(C,q) = \sup q^T z - \inf q^T z. \quad W_{min} = \inf_{\|q\|_2=1} (W(C,q)) \quad W_{max} = \sup_{\|q\|_2=1} W(C,q)$$

$\therefore$ Condition number $=$

$$\mathrm{cond}(C) = \frac{W^2_{max}}{W^2_{min}}$$

} if C is small, then set is spherical, and therefore best approximated by a quadratic function.

Descent Methods

Take the form of $x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)}$.

And must satisfy the equation $\nabla f(x^{(k)})^T \Delta x^{(k)} < 0$ by convexity. i.e. it must make an acute angle with the negative gradient. This is our <u>descent direction</u>.

Newton's Method

The second-order Taylor approximation of f ($\hat{f}$) is

$$\hat{f}(x+v) = f(x) + \nabla f(x)^T v + \frac{1}{2} v^T \nabla^2 f(x) v$$

which is a convex quadratic function of v. It's minimized when

$$v = -\nabla^2 f(x)^{-1} \nabla f(x) \quad \Big\} \text{ This is } \Delta x_{nt} \text{. the Newton step.}$$

Interpreted this way, if $f(x)$ is quadratic or near-quadratic, then Newton's Method should work very well.

Newton's Method is independent of linear/affine changes in coordinates.

Newton decrement:

$$\lambda(x) = \left(\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)\right)^{1/2} = \left(\Delta x_{nt}^T \nabla^2 f(x) \Delta x_{nt}\right)^{1/2}$$

It's used in the backtracking line search for N.M:

$$\nabla f(x)^T \Delta x_{nt} = -\lambda(x)^2$$

Equality Constrained Minimization

§10.1.2 | §4.2.4

Any equality constrained minimization problem can be reduced to an equivalent unconstrained problem by eliminating the equality constraints, at which point you can use the methods in Ch. 9 to solve. Another approach is to solve the dual problem using an unconstrained minimization method, and then recover the solution of the equality constrained problem from the dual solution. ← §10.1.3.

If you can directly handle equality constraints, you can retain important properties, like sparsity, within the function.

Eliminating equality constraints

Find $F \in \mathbb{R}^{n \times (n-p)}$ and vector $\hat{x} \in \mathbb{R}^n$ to parametrize

$$\{x \mid Ax = b\} = \{Fz + \hat{x} \mid z \in \mathbb{R}^{n-p}\}$$

See §C.5 for ways to do this

$\hat{x}$ is any solution to $Ax = b$, and F is null space of A. This makes

$$\min \tilde{f}(z) = f(Fz + \hat{x})$$

From its solution z^*, we can find the solution to $x^* = Fz^* + \hat{x}$. OR, we could find an optimal dual variable $\nu^* = -(AA^T)^{-1} A \nabla f(x^*)$.

We choose elimination matrix F through $N(A)$. However, if $\exists\, T \in \mathbb{R}^{(n-p)\times(n-p)}$, then $\tilde{F} = FT$ is also a suitable elimination matrix.

Using F : $\min f(Fz + \hat{x})$

Using $\tilde{F}$: $\min f(\tilde{F}\tilde{z} + \hat{x}) = f(F(T\tilde{z}) + \hat{x})$

Think of as changing variables in the reduced problem.

Solve via the dual

The dual function for $\min f(x)$ s.t $Ax = b$:

$$g(\nu) = -b^T\nu + \inf_x \left(f(x) + \nu^T Ax\right)$$

$$= \cancel{-b^T\nu - \sup_x\left(f(x) + \nu^T Ax\right)} \quad b^T\nu - \underbrace{\sup_x\left(\left(-A^T\nu\right)^T x - f(x)\right)}$$

$$= -b^T\nu - \underbrace{f^*\left(-A^T\nu\right)}_{\text{by def}}$$

∴ dual problem is

$$\max \quad -b^T\nu - f^*(-A^T\nu)$$

If dual is twice differentiable, then methods for unconstrained minimization can be used to maximize g.

Once ν^* is found, x^* is reconstructed. See example 10.2.

↳ § 5.5.5.

Newton Method with Equality Constraints

Define through second order approximation:

$$\min \ \hat{f}(x+v) = f(x) + \nabla f^T(x) v + \left(\tfrac{1}{2}\right) v^T \nabla^2 f(x) v$$
$$\text{s.t } A(x+v) = b$$

$\therefore \Delta x_{nt}$ is solution to this problem, assuming KKT Matrix is nonsingular. Using the KKT construction & the equality-constrained problem:

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ \omega \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}, \quad \omega \text{ is optimal dual for quadratic problem.}$$

If f is nearly quadratic $x^* \approx x + \Delta x_{nt}$, and $v^* \approx \omega$. Reinterpreted:

$$A x^* = b \qquad \nabla f(x^*) + A^T v = 0. \xrightarrow{\text{substitution}}$$

$$A(x + \Delta x_{nt}) = b, \qquad \nabla f(x + \Delta x_{nt}) + A^T \omega \approx \nabla f(x) + \nabla^2 f(x) \Delta x_{nt} + A^T \omega = 0.$$

using $Ax = b$

$$A \Delta x_{nt} = 0, \qquad \left.\nabla^2 f(x) \Delta x_{nt} + A^T \omega = -\nabla f(x)\right\} \text{Newton Step.}$$

Newton Decrement

$$\left.\lambda(x) = (\Delta x_{nt}^T \nabla^2 f(x) \Delta x_{nt})^{1/2}\right\} \text{Same as unconstrained.}$$

The $v \in \mathbb{R}^n$ is a <u>feasible direction</u> if $Av = 0$; shown by $A\Delta x_{nt} = 0$.

Infeasible Start Newton Method

Start with optimality conditions

$$\left. A x^* = b \qquad \nabla f^*(x) + A^T v = 0 \right\} \text{ x may not be feasible, but } x \in \operatorname{dom} f.$$

sub $x + \Delta x$ for x^* and ω for v^* in first order approximation

$$\nabla f(x + \Delta x) \approx \nabla f(x) + \nabla^2 f(x) \Delta x$$

and now the optimality conditions are

$$A(x + \Delta x) = b \qquad \nabla f(x) + \nabla^2 f(x) \Delta x + A^T \omega = 0$$

$$\therefore \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \omega \end{bmatrix} = -\begin{bmatrix} \nabla f(x) \\ Ax - b \end{bmatrix}$$

Main difference; when residual vanishes, thus reduces to feasible conditions.

Infeasible Start NM with Primal-Dual setup

Primal-Dual: update both primal variable x and dual variable v in order to approx. satisfy optimality conditions. $r(x^*, v^*) = 0$:

$$r(x, v) = \left(r_{dual}(x, v),\ r_{pri}(x, v)\right)$$

$$\underbrace{r_{dual}(x, v) = \nabla f(x) + A^T v}_{\text{dual residual}} \qquad \underbrace{r_{pri}(x, v) = Ax - b}_{\text{primal residual}}$$

§A.4.1

First-order Taylor approximation:

$$r(y+z) \approx \hat{r}(y+z) = r(y) + Dr(y)z \quad ; \quad Dr(y) \in \mathbb{R}^{(n+p)\times(n+p)} \text{ is } \partial r \text{ at } y.$$

Thus, the primal-dual step Δy_{pd} is the step z for which the taylor approx. $\hat{r}(y+z)$ vanishes:

$$Dr(y)\underbrace{\Delta y_{pd}} = -r(y)$$

$$\Delta y_{pd} = (\Delta x_{pd},\ \Delta v_{pd})$$

Evaluating this:

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{pd} \\ \Delta v_{pd} \end{bmatrix} = -\begin{bmatrix} r_{dual} \\ r_{pri} \end{bmatrix} = -\begin{bmatrix} \nabla f(x) + A^T v \\ Ax - b \end{bmatrix}$$

Therefore, we can interpolate that

$$\Delta x_{nt} = \Delta x_{pd}, \qquad w = v + \Delta v_{pd}$$

Notice: this reinterpretation implies that the current value of the dual variable is not needed to compute the primal step, or the dual variable update.

Newton direction at an infeasible point does not imply a descent direction. However, the primal-dual interpretation shows that the norm of the residual does decrease.

$$\frac{\partial}{\partial t} \| r(y + t\Delta y_{pd})\|_2 \Big|_{t=0} = -\|r(y)\|_2$$

Thus $\|r\|_2$ is a good measure of progress in, say, a line search.

Solving an Infeasible Newton's Method involves solving the full KKT system which is presented. There are several ways to do this:

1. Solve the full KKT system. Reasonable when A and H are sparse/small. Use LDL^T factorization.
2. Solve KKT system via elimination; i.e. eliminating the variable v. IP

$$\begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} v \\ w \end{bmatrix} = -\begin{bmatrix} g \\ h \end{bmatrix} \Rightarrow \left.\begin{aligned} Hv + A^T w &= -g \\ Av &= -h \end{aligned}\right\} v = -H^{-1}(g + A^T w)$$

Substitute $\Rightarrow$ $w = \underbrace{(AH^{-1}A^T)}_{\text{Schur compliment of } H}{}^{-1}(h - AH^{-1}g)$. Now we can compute v and w.

11 Interior Point Methods

Now we get inequality constraints :

$$\min \; f_0(x)$$
$$\text{st} \quad f_i(x) \le 0 \;, \; i = 1 \dots m$$
$$Ax = b$$

This problem is assumed to be strictly feasible, meaning that Slater's constraint qualification holds, so there exist dual optimal $\lambda^* \in \mathbb{R}^m$, $\nu \in \mathbb{R}^p$, which satisfy KKT :

$$\left.\begin{aligned} &Ax^* = b \;,\; f_i(x^*) \le 0 \\ &\lambda^* \ge 0 \\ &\nabla f_0(x^*) + \sum_i^m \lambda_i^* \nabla f_i(x^*) + A^T \nu^* = 0 \\ &\lambda_i^* f_i(x^*) = 0 \end{aligned}\right\} \text{Just a quick review}$$

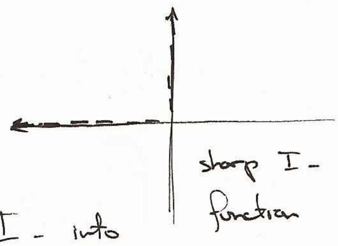
Interior point methods solve the above problems by applying Newton's method to a sequence of equality-constrained problems. Applicable problem forms : LP, QP, QCQP, and GP. Also entropy maximization.

SOCP and SDP are not ready in this way, but can be formed into solvability by Interior Point Methods.

Logarithmic Barrier Function & Central Path

Goal: make inequality constraints into equality constraints, so they can be solved with Newton's Method. Thus, move the inequalities into the objective :

$$\min \; f_0(x) + \sum_{i=1}^m I_-(f_i(x)) \;, \quad I_-(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$$
$$\text{s.t} \quad Ax = b$$



sharp I_- function

This moves everything up, but now the problem is not differentiable. Fix this by turning I_- into a logarithmic barrier instead:

$$\left.\min \; f_0(x) + \sum_{i=1}^m -\left(\frac{1}{t}\right) \log(-f_i(x))\right\} \quad \phi(x) = -\sum_{i=1}^m \log(-f_i(x)) \text{ is logarithmic barrier.}$$

for inequality-constrained problem.

Intuition shows that as $t \to \infty$, the objective function gets closer to the analytical solution of the problem (described with I_-). However, since a large t is hard to minimize given it's Hessian, an iterative approach is used to both grow t and minimize the objective. For future reference :

$$\nabla \phi(x) = \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla f_i(x) \;, \qquad \nabla^2 \phi(x) = \sum_{i=1}^m \frac{1}{f_i(x)^2} \nabla f_i(x) \nabla f_i(x)^T + \sum_{i=1}^m \frac{1}{-f_i(x)} \nabla^2 f_i(x)$$

§A.4.2   §A.4.4.

The log barrier implementation of our objective function can be rewritten

$$\min \; t f_0(x) + \phi(x)$$
$$\text{s.t} \quad Ax = b$$

Central Path

The central path associated with our problem is defined as the set of points $x^*(t)$, $t>0$. They fulfill these necessary & sufficient conditions: $x^*(t)$ is strictly feasible, i.e.

$$Ax^*(t) = b, \quad f_i(x^*(t)) < 0, \quad i = 1 \dots m$$

and there exists $\hat{v} \in \mathbb{R}^p$ s.t.

$$0 = t\nabla f_0(x^*(t)) + \nabla\phi(x^*(t)) + A^T v = t\nabla f_0(x^*(t)) + \sum_{i=1}^{m} \frac{1}{-f_i(x^*(t))} \nabla f_i(x^*(t)) + A^T v$$

From this we can find an important property: every central point yields a dual feasible point, and thus a lower bound on p^*. Defining

$$\lambda_i^*(t) = \frac{1}{-tf_i(x^*(t))}, \quad i = 1 \dots m, \qquad v^*(t) = \frac{\hat{v}}{t} \quad \Big\}\ \text{Both are dual feasible}$$

These are derived from minimizing the Lagrangian derived from our optimality conditions. We can further show that

$$g(\lambda^*(t), v^*(t)) = f_0(x^*(t)) - \frac{m}{t}, \quad m = \#\text{ of inequality constraints.}$$

Thus, the gap between the primal and the dual is just m/t.

Newton Step for Modified KKT equations

The barrier method is given by the linear equations:

$$\begin{bmatrix} t\nabla^2 f_0(x) + \nabla^2\phi(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ v_{nt} \end{bmatrix} = -\begin{bmatrix} t\nabla f_0(x) + \nabla\phi(x) \\ 0 \end{bmatrix}$$

The Newton steps for solving the above is analogous to deriving the Newton steps for the centering problem:

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + A^T v = 0$$

$$-\lambda_i f_i(x) = 1/t, \quad i = 1 \dots m$$

$$Ax = b$$

First, eliminate λ_i using $\lambda_i = -1/(t f_i(x))$ and simplify. Then, replace the non-linear term in the first equation with its Taylor expansion, thus yielding the linear equations

$$Hv + A^T\tilde{v} = -g, \quad Av = 0 \quad \Big\}\ \text{Drastically simplified.}$$

Further simplification shows that

$$v = \Delta x_{nt} \quad \text{and} \quad \tilde{v} = \left(\frac{1}{t}\right) v_{nt}$$

If we did <u>not</u> eliminate λ in this way, we get the <u>primal-dual search directions</u>.

Primal-dual Interior Point Methods

Differences between barrier method: 1) At each iteration, both primal & dual variables are updated.

2) Search direction is obtained from Newton's method, applied to modified KKT.

3) The primal & dual iterates are <u>not</u> necessarily feasible.

Start with the modified KKT conditions expressed as $r_t(x, \lambda, v) = 0$:

$$r_t(x,\lambda,v) = \begin{bmatrix} \nabla f_0(x) + Df(x)^T \lambda + A^T v \\ -\mathrm{diag}(\lambda) f(x) - (1/t)\mathbf{1} \\ Ax - b \end{bmatrix}$$

dual residual, centrality residual, primal residual

If x, λ, v satisfy $r(x,\lambda,v) = 0$, then they are all optimal with duality gap m/t.

Here, $f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix}$, $Df(x) = \begin{bmatrix} \nabla f_1(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{bmatrix}$

Now to perform the Newton step solving r_t <u>without</u> elimination of λ. Denote the current point & Newton step:

$$y = (x, \lambda, v) \qquad \Delta y = (\Delta x, \Delta\lambda, \Delta v)$$

Newton step : $r_t(y + \Delta y) \approx r_t(y) + Dr_t(y)\Delta y = 0$

$$\therefore \begin{bmatrix} \nabla^2 f_0(x) + \sum_{i=1}^m \lambda_i \nabla^2 f_i(x) & Df(x)^T & A^T \\ -\mathrm{diag}(\lambda) Df(x) & -\mathrm{diag}(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{pd} \\ \Delta\lambda_{pd} \\ \Delta v_{pd} \end{bmatrix} = - \begin{bmatrix} r_{dual} \\ r_{cent} \\ r_{pri} \end{bmatrix}$$

Solution of this gives primal-dual search direction.

Dealing with SOCP & SDP

The generalized logarithmic barrier is a good place to start:

$\psi : \mathbb{R}^q \to \mathbb{R}$ for proper cone $K \subseteq \mathbb{R}^q$

1. $\mathrm{dom}\ \psi = \mathrm{int}\ K$, $\nabla^2\psi(y) \prec 0$ for $y \in \mathrm{int}\ K$
2. $\psi(sy) = \psi(y) + \theta \log s$, $s > 0$. θ is the degree of the logarithm

Note: $K = \mathbb{R}^n_+$:
$\psi(y) = \sum \log y_i$
$\nabla\psi = \begin{pmatrix} \frac{1}{y_1} \\ \vdots \\ \frac{1}{y_n} \end{pmatrix}$

Thus applying this definition to SOCP & SDP:

- $K = S^n_+$, $\psi(y) = \log\det y$, degree $\theta = n$, $\nabla\psi(y) = Y^{-1}$.
- $K = S.O.C(n+1)$, $\psi(y) = \log\left(y_{n+1}^2 - \sum_{i=1}^n y_i^2\right)$, degree $\theta = 2$

comes from barrier function & Second Order Cone.

We show that for any y in the cone K, the gradient of $\psi(y)$ is in the dual cone:

$\nabla\psi(y) \succeq_{K^*} 0$, $y^T\nabla\psi(y) = \theta$ for any $y \succ_K 0$

$K^* = \{x \mid y^T x \geq 0,\ \forall y \in K\} \longrightarrow$ definition of dual cone.

proven by contradiction, and derivative of $\psi(sy)$, $s > 0$.

Given this proof, and our definitions for $\psi(y)$ from the problem definitions above, we get generalized barrier functions: $\phi(x) = -\sum_{i=1}^m \psi_i(-f_i(x))$

Now we can attempt to solve $\min\ t f_0(x) + \phi(x)$

s.t $Ax = b$

$x^*(t)$: central path for f_0

If x^* is optimal, then $\exists \omega$:

$$t \nabla f_0(x) + \sum_{i=1}^{m} \nabla f_i(x)^T \nabla \psi_i(-f_i(x)) + A^T \omega = 0$$

and interpolating dual variables:

$$\lambda^* = \frac{1}{t} \nabla \psi_i(-f_i(x)) \quad , \quad \upsilon^* = \frac{\omega}{t}$$

$\therefore \mathcal{L}(x, \lambda^*, \upsilon^*) = t f_0(x) + \sum_i \lambda_i^* f_i(x) + \upsilon^*(Ax-b)$, which means $x^*(t) = \arg\min \mathcal{L}$

and thus our dual problem is

$$g(\lambda^*, \upsilon^*) = \mathcal{L}(x^*(t), \lambda^*, \upsilon^*).$$

Further extrapolating:

$$p^* \geq g(\lambda^*, \upsilon^*) = \mathcal{L}(x^*, \lambda^*, \upsilon^*) = f_0(x^*) + \sum \lambda_i^* f(x^*) + \underbrace{\cancel{\upsilon^{*T}(Ax^*-b)}}_{\text{always } =0}$$

$$= f_0(x^*) - \frac{1}{t} \sum_{i=1}^{m} \theta_i \Big\}$$ This is nearly the same form as a regular inequality problem.

Now, with no inequality approaches to go after, we can solve it using Newton's Method.

16.2) Generator Problem

I think I got this one... to a point

P_{it} : power of generator i at time T

$\sum_i p_{it} = d_t$

$\underline{p}_{it} \le p_{it} \le \bar{p}_{it} \Rightarrow |p_{it+1} - p_{it}| \le \underline{p}_i$ } Have a limit on raising the power in a single time span. (aka ramping constraint)

$\therefore \min \sum_{i,t} g_i(p_{it}) + \Phi(\bar{p})$

s.t $\sum_i p_{it} = d_t$

$$\left.\begin{array}{l} \underline{p}_{it} \le p_{it} \le \bar{p}_{it} \\ |p_{it+1} - p_{it}| \le \underline{p}_i \end{array}\right\}$$

a. Show there exists price Q_t s.t. optimal solvers

$\min \sum \left(g_i(p_{it}) - Q_t p_{it}\right) + \Phi(\bar{p})$

s.t. two constraints above hold.

This comes from the extension of the solution of the $\mathcal{L}(x,\mu,v)$, giving $g(\theta,t) =$

It is the dual of the problem. Related to economics: In a complete market, the price Q must be competitive.