Lagrangian Dual.

Find x, subject to

$f_i(x) \leq 0, i = 1, \ldots m.$

Remember calculus (constrained optimization.)

Lagrangian: $L(x,\lambda) = \sum_{i=1}^{m} \lambda_i f_i(x)$

$\lambda_i \geq 0$ - Lagrangian multiplier for inequality i.

For feasible solution x, $L(x,\lambda)$ is

(A) non-negative in expectation

(B) positive for any λ.

(C) non-positive for any valid λ.

If λ, where $L(x,\lambda)$ is positive for all x

(A) there is no feasible x.

(B) there is no x,λ with $L(x,\lambda) < 0$.

Lagrangian:constrained optimization.

$$\begin{array}{ll} \min & f(x) \\ \text{subject to } f_i(x) \leq 0, & i = 1,...,m \end{array}$$

Lagrangian function:

$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$

If (primal) x value v

For all $\lambda \geq 0$ with $L(x,\lambda) \leq v$

Maximizing λ only positive when $f_i(x) = 0$.

If there is λ with $L(x,\lambda) \geq \alpha$ for all x

Optimum value of program is at least α. $OPT \geq \min_x L(x,\lambda)$

Why? A feasible solution has $f_i(x) \leq 0$, so $L(x,\lambda) \leq f(x)$.

Saddle point.

Lagrangian function:

$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$

Primal problem.

x-player "best defense":

$\min_x \max_\lambda L(x,\lambda)$.

x, that minimizes $L(x,\lambda)$ over all $\lambda \geq 0$.

Dual problem:

λ-player "best defense":

$\max_\lambda \min_x L(x,\lambda)$.

λ, that maximizes $L(x,\lambda)$ over all x.

Saddle point: (x,y) with both conditions:

For x-player: $\nabla_\lambda L'(x,\lambda) \leq 0 \implies f_i'(x) \leq 0$.

For λ player: $L'(x,\lambda) = \nabla_x f(x) + \sum_{i=1}^{m} \lambda_i \nabla_x f_i(x) = 0$.

At saddle point. Is $\lambda_i \geq 0$ only if $f_i(x) = 0$? Yes.

Linear Program.

$\min cx, Ax \geq b$

$$\begin{array}{ll} \min & c \cdot x \\ \text{subject to } b_i - a_i \cdot x \leq 0, & i = 1,...,m \end{array}$$

Lagrangian (Dual):

$L(\lambda,x) = cx + \sum_i \lambda_i (b_i - a_i x)$.

or

$L(\lambda,x) = -(\sum_j x_j (a_j \lambda - c_j)) + b\lambda$.

Best λ?

$\max b \cdot \lambda$ where $a_j \lambda = c_j$.

$\max b\lambda, \lambda^T A = c, \lambda \geq 0$

Saddle point: complementary slackness.

Newton's method for root finding.

Find a root of $f(x)$.

At x.

$0 = f(x) + f'(x)(t - x). \implies t = x - \frac{f(x)}{f'(x)}$

$x_{n+1} = x_n - \frac{f(x)}{f'(x)}$.

Convergence Analysis.

Choose α where $f(\alpha) = 0$.

$f(\alpha) = f(x) + f'(x_n)(\alpha - x) + R$

$R = \frac{1}{2!} f''(\psi)(\alpha - x)^2$

For some $\psi \in [\alpha, x]$. (Assume $\alpha < x$.)

Lagrange form of Taylor's series.

$0 = f(\alpha) = f(x) + f'(x)(\alpha - x) + \frac{1}{2} f''(\psi)(\alpha - x)^2$

Rearrange:

$\frac{f(x)}{f'(x)} + (\alpha - x) = \frac{-f''(\psi)}{2f'(x)}(\alpha - x)^2$

Let $x' = x - \frac{f(x)}{f'(x)}$

$\alpha - x' = \frac{-f''(\psi)}{2f'(x)}(\alpha - x)^2$

$|\alpha - x'| = |\frac{f''(\psi)}{2f'(x)}|(\alpha - x)^2$.

If $|\alpha - x| < \varepsilon \implies |\alpha - x| \leq |\frac{f''(\psi)}{2f'(x)}|\varepsilon^2$

Lagrange form for Taylor's Theorem: skipping.

$F(t) = f(t) + f'(t)(x-t) + \frac{f''(t)}{2!}(x-t)^2 \cdots \frac{f^{(k)}}{k!}(x-t)^k$ for $t \in [a,x]$.

Note $F(x) - F(a) = f(x) - F(a) = R(x)$. Remainder in Taylor's.

The mean value theorem: There is $\psi \in [a,x]$, where

$\frac{F'(\psi)}{G'(\psi)} = \frac{F(x)-F(a)}{G(x)-G(a)}$

General version of $G(x) = x$: ψ has slope equal to average slope.

$F'(t)$? Use product rule on kth term:
$\frac{f^{(k+1)}(t)}{k!}(x-t)^k - \frac{f^{(i)}(t)}{(k-1)!}(x-t)^{k-1}$.

Successive terms telescope: E.g.,

$f'(t) + (f''(t)(x-t) - f'(t)(x-t)^0) = f''(t)(x-t)$.
So, $F'(t) = \frac{f^{(k+1)}(t)}{k!}(x-t)^k$.

$R(x) = F(x) - F(a) = \frac{f^{(k+1)}(\psi)}{k!}(x-\psi)^k \frac{G(x)-G(a)}{G'(\psi)}$

Set $G(t) = (x-t)^{k+1}$, $G'(\psi) = -(k+1)(x-t)^k$, $G(a) = (x-a)^{k+1}$, $G(x) = 0$.

$R(x) = F(x) - F(a) = \frac{f^{(k+1)}(\psi)}{(k+1)!}(x-a)^{k+1}$

Multivariate version.

Vector x, functions: $f_1(x), f_2(x), \ldots f_k(x)$.

$x_{n+1} = x_n - J^{-1}(n)F(x_n)$, where $F(x) = (f_1(x), \ldots, f_n(x))$.

where

$$J = \begin{bmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \end{bmatrix}$$

Interior point on the central path.

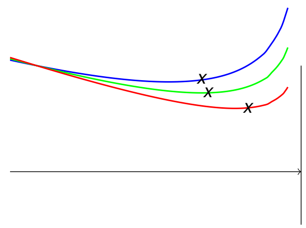
Find x, that minimizes $f_0(x)$ subject to

$f_i(x) \le 0, i = 1, \ldots m$.

Central path:

$\min t f_0(x) - \sum_{i=1}^{m} \ln(-f_i(x))$

The minimizer, $x(t)$, form the **central path.**



The sequence of x's are "central path".

Lagrangian Dual and Central Path.

$\min t f_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition?

Derivative: $t\nabla f_0(x) - \sum_{i=1} \frac{\nabla f_i(x)}{f_i(x)} = 0$ $\nabla f_0(x) - \sum_{i=1} \frac{1}{tf_i(x)} \nabla f_i(x) = 0$

Or, $\nabla f_0(x) = \sum_{i=1} \frac{\nabla f_i(x)}{tf_i(x)}$ (Opposing force fields.)

Recall, Lagrangian: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

Fix λ, optimize for x^* give valid lower bound on solution.
Optimality Condition.
Derivative: $\nabla f_0(x) + \sum_{i=1} \lambda_i \nabla f_i(x) = 0.$

Take $\lambda_i^{(t)} = -\frac{1}{tf_i(x)}$. $x(t) = x^*(\lambda^{(t)})$! Same optimal point!

Value? Found λ where:
$\min_x L(\lambda, x) = f_0(x) + \sum_{i=1} \lambda_i f_i(x) = f_0(x) - \frac{m}{t} \le \min_x \max_\lambda L(\lambda, x)$.

Central point $x(t)$ within $\frac{m}{t}$ of optimal primal!!!!
$L(\lambda, x(t)) \ge f_0(x) - \frac{m}{t} \implies \min_x L(\lambda, x) + \frac{m}{t} \ge f_0(x)$
$\implies OPT + \frac{m}{t} \ge f_0(x)$

Central path.

$\min_x f_0(x), f_i(x) \le 0$.

$\min_x t f_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

$f_0(x(t)) \le OPT + \frac{m}{t}$

Algorithm: take $t \to \infty$.

Finding $x(t)$?

Assume you have $x(t)$, change $t = \mu t$, for $\mu > 1$.
Find $x(\mu t)$.
Idea: newton's method.
Should show new optimal point not too different from old.

Next.

Central Path evolution.

Old point $x = x(t)$ versus $x^+ = x(\mu t)$? Minimizing:
$\mu t f_0(x) - \sum_{i=1}^{n} \ln(-f_i(t))$.

Difference in new objective from old optimal point to new:
$\mu t f_0(x) - \sum_{i=1}^{m} \ln(-f_i(x)) - \mu t f_0(x^+) + \sum_{i=1}^{m} \ln(-f_i(x^+))$

Simplify: $\mu t f_0(x) - \mu t f_0(x^+) + \sum_i \ln(-\frac{f_i(x^+)}{f_i(x)})$

Let $\lambda_i = -\frac{1}{tf_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.
$f_0(x) - L(\lambda, x') \le \frac{m}{t}$ since $\sum_i \lambda_i f_i(x) = -\frac{m}{t}$.

$\mu t f_0(x) - \mu t f_0(x^+) + \sum_i \ln(-\mu t \lambda_i f_i(x^+)) - m \ln \mu$
$\ln(-x) = ln(1 - (1+x)) \le -(1+x)$

$$\begin{aligned}
&\le \mu t f_0(x) - \mu t f_0(x^+) - \sum_i (1 + \lambda_i \mu t f_i(x^+)) - m \ln \mu \\
&= \mu t f_0(x) - \mu t f_0(x^+) - \mu t \sum_i \lambda_i f_i(x^+) - m - m \ln \mu \\
&= \mu t (f_0(x) - (f_0(x^+) + \sum_i \lambda_i f_i(x^+)) - m - m \ln \mu \\
&= \mu t (f_0(x) - L(\lambda, x^+)) - m - m \ln \mu \\
&\le \mu t (\tfrac{m}{t}) - m - m \ln \mu \\
&= m(\mu - 1 - \ln \mu)
\end{aligned}$$

...Central Path Evolution

Old point $x = x(t)$ versus $x^+ = x(\mu t)$?
Minimizing: $F(x) = \mu t f_0(x) + \sum_{i=0}^{n} \ln(-f_i(t))$.

We proved: $F(x) - F(x^+) \leq m(\mu - 1 - \ln \mu)$.

Choose $\mu = 1 + \delta$.
$\ln(1+x) \approx x - x^2/2$. $\implies \ln(1+\delta) = \delta - \delta^2/2$

$F(x) - F(x^+) = m(1 + \delta - 1 - (\delta - \delta^2/2)) = m(\delta^2/2)$.

Choose $\delta = \frac{1}{\sqrt{m}}$ or $\mu = (1+\delta)$.
$F(x) - F(x^+) = (\delta^2/2)) = \frac{1}{2}$.

Modifying t by factor of $(1 + \frac{1}{\sqrt{m}})$, optimal still close!

t can be arbitrarily large!

The value of the objective function can be gigantic
and change by an enormous amount.
but current point is very close to optimal.

Intuition here?

An attempt at intuition.

$\min \sum_i x_i, x \geq 0$.

optimum is 0.

Central path: $F_t(x) = t\sum_i x_i + \sum_i \ln x_i$

Optimum: $x_i = \frac{1}{t}$.

$t \to \mu t$

New optimum: $x_i^+ = \frac{1}{\mu t}$.

Notice: the change in x is quite small.
Roughly $(\mu - 1)\frac{1}{t} = \frac{1}{\sqrt{m}t}$ where t is large.

Intuitively: new point is very close to old point.

Slightly more generally.

Only one vertex on polytope.

n inequalities, n unkonwns: $\min cx, Ax \geq b$.

Is solution bounded or unbounded?

Alg: Linear equation solve for intersection of n inequalities, check if there is some direction of improvement.

Evolution of central path.

Optimal $x(t)$: $\nabla f_0(x) + \sum_{i=1} \frac{1}{tf_i(x)} \nabla f_i(x) = 0$
$tc = -\sum_i \frac{a_i}{a_i x - b_i}$

$s_i = a_i x - b_i$. "Distance' to constraint.
Recall previous example: $x \geq 0$, the x_i are slack variables.
$s = Ax - b$.

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $Ax(\mu t) - b = s(t)/\mu$ works.

Since only n inequalities, can just solve to get next point.

Answer is easy too.

More generally.

General $Ax \geq b, \min cx$.

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ works.

Overdetermined if more than n inequalities, so maybe not possible.

So, need to find solution to: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

Showed solution is at least close in value to old solution on $F(x)$.

One thing to note:
if you know the optimal vertex (tight constraints).
then you are done.

Idea: close enough to tight constraints. Done.

Close enough to a vertex, can jump to vertex.

Cramer's rule, gives estimate of how close the closest two vertices can be.

Interior Point Method.

Find central point.

Recall: $F(x) = tf_0(x) - \sum_i \log(-f_i(x))$.

Find point: $G(x) = \nabla F(x) = 0$.

Newton: find all zeros of vector valued $G(x)$!

$g_1(x) = \frac{t\partial f_0(x)}{\partial x_1} - \sum_i \frac{\partial f_i(x)}{\partial x_1} \frac{1}{f_i(x)}$

Newton:
$\implies |(x_{n+1} - \alpha)| \leq |\frac{g'(\psi)}{2g''(x_n)}|(x_n - \alpha)^2$.

Recall, distance for x to x^+ is pretty small.
On the order of $1/t$.

What about the ratio? $|\frac{g''(\psi)}{2g'(x_n)}|$

Behavior of log barrier.

What about the ratio? $|\frac{g''(\psi)}{2g'(x_n)}|$

What if $f(x) = \log x$ and recall $g(x) = f'(x)$?

$(\log x)' = 1/x$, $(\log x)'' = -1/x^2$, $(\log x)''' = 2/x^3$.

$|(\log x)'''| = |\frac{x}{2}(\log x)''|$.

Thus, this ratio is around $1/x$.

Thus, $|\frac{g''(\psi)}{2g'(x_n)}|(x - x^+) \leq 1/2$.

Quadratic convergence: ratio is small.

Newton's Method.

$f(x) = \log(a_i x - b_i)$

$s_i = a_i x - b_i$

$f'(x) = \frac{a_i}{s_i} \quad f''(x) = \frac{1}{s_i^2} a_i a_i^T \quad f'''(a) = a_i^{\otimes 3} \frac{1}{s_i^3}$

$f(x+u) = f(x) + f'(x) \cdot u + u^{\otimes 2} \cdot f''(x)$

The minimizer? $-\frac{1}{2}(f''(x))^{-1} f'(x)$.

Self-Concordance: $|f'''(x)| \leq 2f''(x)^3/2$.

Newton Method: $f(x) - f(x*) \leq 1/2$, it converges quadratically.

Another Type of IPM strategy.

$\min xc, Ax \geq b$.

$F(x) = tcx - \sum_i \log(a_i x - b_i)$.

$\nabla F(x) = tc - \sum_i \frac{a_i}{s_i}$

Introduce dual variables: λ_i.
Approximate Complementary slackness.
$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$
$s = b - Ax$

Gives another possibility:
Explicitly maintain primal-dual soluion: (x, s)