# Interior point on the central path.

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

# Interior point on the central path.

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

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

# Interior point on the central path.

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

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

Central path:

# Interior point on the central path.

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

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

Central path:

$\min tf_0(x) - \sum_{i=1}^{m} \ln(-f_i(x))$

# Interior point on the central path.

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

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

Central path:

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

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

# Interior point on the central path.

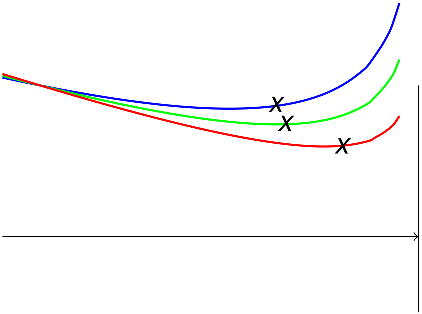
Find $x$, that minimizes $f_0(x)$ subject to

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

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.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition?

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

# Lagrangian Dual and Central Path.

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

Optimality condition? Take Derivative.

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

# Lagrangian Dual and Central Path.

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

Optimality condition? Take Derivative.

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

Or, $\nabla f_0(x) = \sum_{i=1} \frac{\nabla f_i(x)}{t f_i(x)}$

# Lagrangian Dual and Central Path.

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

Optimality condition? Take Derivative.

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

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

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

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, optimize for $x^*$ give valid lower bound on solution.

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, optimize for $x^*$ give valid lower bound on solution.

Optimality Condition.

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, optimize for $x^*$ give valid lower bound on solution.

Optimality Condition.

Derivative:

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, optimize for $x^*$ give valid lower bound on solution.

Optimality Condition.

Derivative: $\nabla f_0(x) + \sum_{i=1} \lambda_i \nabla f_i(x) = 0$.

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, 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 $\lambda_i^{(t)} = -\frac{1}{tf_i(x)}$.

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, 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 $\lambda_i^{(t)} = -\frac{1}{tf_i(x)}$. $\quad x(t) = x^*(\lambda^{(t)})$!

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take 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 $\lambda$, 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!

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, 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 $\lambda$ where:

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, 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 $\lambda$ where:

$\min_x L(\lambda, x) = f_0(x) + \sum_{i=1} \lambda_i f_i(x) = f_0(x) - \frac{m}{t}$

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, 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 $\lambda$ where:

$\min_x L(\lambda, x) = f_0(x) + \sum_{i=1} \lambda_i f_i(x) = f_0(x) - \frac{m}{t}$

Central point $x(t)$ within $\frac{m}{t}$ of optimal primal!!!!

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, 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 $\lambda$ where:

$\min_x L(\lambda, x) = f_0(x) + \sum_{i=1} \lambda_i f_i(x) = f_0(x) - \frac{m}{t}$

Central point $x(t)$ within $\frac{m}{t}$ of optimal primal!!!!

$L(\lambda, x(t)) \geq f_0(x) - \frac{m}{t}$

# Lagrangian Dual and Central Path.

$\min tf_0(x) - \sum_{i=1} \ln(-f_i(x))$

Optimality condition? Take Derivative.

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

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 $\lambda$, 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 $\lambda$ where:

$\min_x L(\lambda, x) = f_0(x) + \sum_{i=1} \lambda_i f_i(x) = f_0(x) - \frac{m}{t}$

Central point $x(t)$ within $\frac{m}{t}$ of optimal primal!!!!

$L(\lambda, x(t)) \geq f_0(x) - \frac{m}{t} \implies \min_x L(\lambda, x) + \frac{m}{t} \geq f_0(x)$

# Lagrangian Dual and Central Path.

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

Optimality condition? Take Derivative.

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

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

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

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

Value? Found $\lambda$ where:

$\min_x L(\lambda, x) = f_0(x) + \sum_{i=1} \lambda_i f_i(x) = f_0(x) - \frac{m}{t}$

Central point $x(t)$ within $\frac{m}{t}$ of optimal primal!!!!

$L(\lambda, x(t)) \geq f_0(x) - \frac{m}{t} \implies \min_x L(\lambda, x) + \frac{m}{t} \geq f_0(x)$

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

# Central path.

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

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

# Central path.

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

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

Optimal: $x(t)$ is feasible.

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

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

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

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

Algorithm: take $t \to \infty$.

# Central path.

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

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

Optimal: $x(t)$ is feasible.

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

Algorithm: take $t \rightarrow \infty$.

Finding $x(t)$?

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

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

Algorithm: take $t \to \infty$.

Finding $x(t)$?

Assume you have $x(t)$, change $t = \mu t$, for $\mu > 1$.

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

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

Algorithm: take $t \to \infty$.

Finding $x(t)$?

Assume you have $x(t)$, change $t = \mu t$, for $\mu > 1$.
Find $x(\mu t)$.

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

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

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.

# Central path.

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

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

Optimal: $x(t)$ is feasible.

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

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.

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

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

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.

# Central path.

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

$\min_x tf_0(x) - \sum_{i>0} \ln(-f_i(x))$

Optimal: $x(t)$ is feasible.

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

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 $x^+ = x(\mu t)$?

# Central Path evolution.

Old point $x = x(t)$ versus $x^+ = x(\mu t)$?

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

# 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:

# 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:

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

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

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

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

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

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

$\ln(-x) = ln(1 - (1 + x)) \leq -(1 + x)$

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

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

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

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

$\ln(-x) = ln(1-(1+x)) \leq -(1+x)$

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

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

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

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

$\ln(-x) = ln(1-(1+x)) \leq -(1+x)$

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

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

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

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

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

$\ln(-x) = ln(1-(1+x)) \leq -(1+x)$

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

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

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

$$= \mu t (f_0(x) - L(\lambda, x^+)) - m - m \ln \mu$$

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

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

$\ln(-x) = ln(1 - (1 + x)) \leq -(1 + x)$

$$\leq \mu t f_0(x) - \mu t f_0(x^+) - \sum_i (1 + \lambda_i \mu t f_i(x^+)) - m \ln \mu$$
$$= \mu t f_0(x) - \mu t f_0(x^+) - \mu t \sum_i \lambda_i f_i(x^+) - m - m \ln \mu$$
$$= \mu t (f_0(x) - (f_0(x^+) + \sum_i \lambda_i f_i(x^+)) - m - m \ln \mu$$
$$= \mu t (f_0(x) - L(\lambda, x^+)) - m - m \ln \mu$$
$$\leq \mu t (\frac{m}{t}) - m - m \ln \mu$$

# 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 $\lambda_i = -\frac{1}{t f_i(x)}$. Remember: $L(\lambda, x) = f_0(x) + \sum_i \lambda_i f_i(x)$.

$L(\lambda, x') \leq \frac{m}{t}$ since $\min_x L(\lambda, x) \leq \frac{m}{t}$

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

$\ln(-x) = ln(1-(1+x)) \leq -(1+x)$

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

# ...Central Path Evolution

Old point $x = x(t)$ versus $x^+ = x(\mu t)$?

# ...Central Path Evolution

Old point $x = x(t)$ versus $x^+ = x(\mu t)$?

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

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

# ...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 $\mu = 1 + \delta$.

# ...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 $\mu = 1 + \delta$.

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

# ...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 $\mu = 1 + \delta$.

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

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

# ...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 $\mu = 1 + \delta$.

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

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

# ...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 $\mu = (1+\delta)$.

# ...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 $\mu = (1+\delta)$.
$F(x) - F(x^+) = (\delta^2/2)) = \frac{1}{2}$.

# ...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!

# ...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!

# ...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!

# ...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

# ...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.

# ...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.

# ...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: $F(x) - F(x^+) \le m(\mu - 1 - \ln \mu)$.

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.

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

# An attempt at intuition.

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

optimum is 0.

# An attempt at intuition.

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

optimum is 0.

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

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

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

$t \rightarrow \mu t$

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

$t \rightarrow \mu t$

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

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

$t \rightarrow \mu t$

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

Notice: the change in $x$ is quite small.

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

$t \rightarrow \mu t$

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.

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

$t \rightarrow \mu t$

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.

# Slightly more generally.

Only one vertex on polytope.

$n$ inequalities, $n$ unkonwns:

# Slightly more generally.

Only one vertex on polytope.

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

# Slightly more generally.

Only one vertex on polytope.

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

Is solution bounded or unbounded?

# 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.

# 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.

# 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 $x(t)$:

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

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

$tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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

$tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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

$tc = -\sum_i \frac{a_i}{a_i x - b_i}$

$s_i = a_i x - b_i$.

# 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.

# 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.

# 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.
  $s = Ax - b$.

# 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 $b - Ax(t) = s(t)$.

# 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 $b - Ax(t) = s(t)$.

# 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.

# 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.

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

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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

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

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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

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

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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 $F(x)$.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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:

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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).

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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.

# More generally.

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

Given solution to $x(t)$ with $b - Ax(t) = s(t)$.
Then $b - Ax(\mu t) = s(t)/\mu$ is optimal: $\mu tc = -\sum_i \frac{a_i}{a_i x - b_i}$

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.

# Interior Point Method.

Find central point.

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

# Interior Point Method.

Find central point.

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

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

# 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

# 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 $G(x)$!

# 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 $G(x)$!

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

# 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:

# 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:

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

# 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.

# 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 $1/t$.

# 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 $1/t$.

# 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? $|\frac{g''(\psi)}{2g'(x_n)}|$

# Newton's Method.

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

# Newton's Method.

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

$$s_i = b_i - a_i x$$

# Newton's Method.

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

$$s_i = b_i - a_i x$$

$$f'(x) = \frac{a_i}{s_i}$$

# Newton's Method.

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

$$s_i = b_i - a_i x$$

$$f'(x) = \frac{a_i}{s_i} \qquad f''(x) = \frac{1}{s_i^2} a_i a_i^T$$

# Newton's Method.

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

$$s_i = b_i - a_i x$$

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

# Newton's Method.

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

$$s_i = b_i - a_i x$$

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

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

# Newton's Method.

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

$$s_i = b_i - a_i x$$

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

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

The minimizer?

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f'(x)$ is linear, goes to $f'(x) = 0$.

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f'(x)$ is linear, goes to $f'(x) = 0$.

Scaled by slope, $f''(x)$, of $f'(x)$.

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f'(x)$ is linear, goes to $f'(x) = 0$.

Scaled by slope, $f''(x)$, of $f'(x)$.

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.
If $f'(x)$ is linear, goes to $f'(x) = 0$.
Scaled by slope, $f''(x)$, of $f'(x)$.

Another Newton Method Analysis:

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.
 If $f'(x)$ is linear, goes to $f'(x) = 0$.
 Scaled by slope, $f''(x)$, of $f'(x)$.

Another Newton Method Analysis:
 potential function: $\|f'(x)\| / \|f''(x)\|$.

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.
If $f'(x)$ is linear, goes to $f'(x) = 0$.
Scaled by slope, $f''(x)$, of $f'(x)$.

Another Newton Method Analysis:
potential function: $\|f'(x)\| / \|f''(x)\|$.

Idea: if $f''(x)$ does not change, then $f'(x) = 0$.

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.
If $f'(x)$ is linear, goes to $f'(x) = 0$.
Scaled by slope, $f''(x)$, of $f'(x)$.

Another Newton Method Analysis:
potential function: $\|f'(x)\| / \|f''(x)\|$.

Idea: if $f''(x)$ does not change, then $f'(x) = 0$.
$f'''(x) \leq f'(x)^{3/2} \rightarrow f''(x)$ does not change much.

# Newton's Method.

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

$s_i = b_i - a_i x$

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

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

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

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

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f'(x)$ is linear, goes to $f'(x) = 0$.

Scaled by slope, $f''(x)$, of $f'(x)$.

Another Newton Method Analysis:

potential function: $\|f'(x)\| / \|f''(x)\|$.

Idea: if $f''(x)$ does not change, then $f'(x) = 0$.

$f'''(x) \leq f'(x)^{3/2} \rightarrow f''(x)$ does not change much.

potential $\|f'(x')\| / \|f''(x')\|$ decreases.

# Behavior of log barrier.

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

# Behavior of log barrier.

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

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

# Behavior of log barrier.

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

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

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

# Behavior of log barrier.

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

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

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

$|(\log x)'''| = 2|(\log x)''|^{3/2}$.

# Behavior of log barrier.

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

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

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

$|(\log x)'''| = 2|(\log x)''|^{3/2}$.

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

# Behavior of log barrier.

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

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

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

$|(\log x)'''| = 2|(\log x)''|^{3/2}$.

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

Newton analysis we did: $|\frac{g''(\psi)}{2g'(x_n)}|(x - x^+) < 1$.

# Behavior of log barrier.

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

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

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

$|(\log x)'''| = 2|(\log x)''|^{3/2}$.

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

Newton analysis we did: $|\frac{g''(\psi)}{2g'(x_n)}|(x - x^+) < 1$.

Quadratic convergence: ratio is small.

# Another Type of IPM strategy.

$\min xc, Ax \geq b$.

# Another Type of IPM strategy.

$\min xc, Ax \geq b$.

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

# Another Type of IPM strategy.

$\min xc, Ax \geq b$.

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

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

# 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: $\lambda_i$.

# 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: $\lambda_i$.

Approximate Complementary slackness.

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$

$s = b - Ax$

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$

$s = b - Ax$

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$

$s = b - Ax$

Predictor-Corrector:

(1) decrease $F(x)$

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$

$s = b - Ax$

Predictor-Corrector:

(1) decrease $F(x)$

(2) Fix complementary slackness.

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$

$s = b - Ax$

Predictor-Corrector:

(1) decrease $F(x)$

(2) Fix complementary slackness.

Gives another possibility:

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

Predictor-Corrector:
(1) decrease $F(x)$
(2) Fix complementary slackness.

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

# 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: $\lambda_i$.

Approximate Complementary slackness.

$\lambda_i s_i = \frac{1}{t}$ verse $\lambda_i s_i = 0$

$s = b - Ax$

Predictor-Corrector:

(1) decrease $F(x)$

(2) Fix complementary slackness.

Gives another possibility:

Explicitly maintain primal-dual soluion: $(x, s)$

# Gradient descent and Newton.

minimize $f(x)$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$$x = x - \alpha f'(x).$$

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y) dx > 0$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y) dx > 0$.
Also: $f'(y) > f'(x) - M(y - x) > 0$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$?

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum?

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f''(x) \geq m$, then $f'(x) = 0$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f''(x) \geq m$, then $f'(x) = 0$. $x$: $(x' - x) \leq f'(x)/m$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f''(x) \geq m$, then $f'(x) = 0$. $x$: $(x' - x) \leq f'(x)/m$.

Estimate of how far is $f'(x)/f''(x)$.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f''(x) \geq m$, then $f'(x) = 0$. $x$: $(x' - x) \leq f'(x)/m$.

Estimate of how far is $f'(x)/f''(x)$. Analysis: the estimate decreases.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y - x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f''(x) \geq m$, then $f'(x) = 0$. $x$: $(x' - x) \leq f'(x)/m$.

Estimate of how far is $f'(x)/f''(x)$. Analysis: the estimate decreases.

# Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y - x) > 0$.
When: $(y - x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f''(x) \geq m$, then $f'(x) = 0$. $x$: $(x' - x) \leq f'(x)/m$.

Estimate of how far is $f'(x)/f''(x)$. Analysis: the estimate decreases.

as long as $f''(x)$ does not change too much.

## Gradient descent and Newton.

minimize $f(x)$.

Gradient descent.

$x = x - \alpha f'(x)$.

Decreases function until gradient changes sign.

If $f''(y) \leq M$ for $y = x - \alpha f'(x)$.

Improve when: $f'(y) > f'(x) - \int_x^y f''(y)dx > 0$.
Also: $f'(y) > f'(x) - M(y-x) > 0$.
When: $(y-x) \leq f'(x)/M$.
set $\alpha = 1/M$.

For linear function: $f'(x)$? Optimum? Is infinitely far.

Newton: $x = x - \frac{f'(x)}{f''(x)}$.

If $f''(x) \geq m$, then $f'(x) = 0$. $x$: $(x' - x) \leq f'(x)/m$.

Estimate of how far is $f'(x)/f''(x)$. Analysis: the estimate decreases.

as long as $f''(x)$ does not change too much.