1 Network optimization

Objective function is $\sum w_\phi U(f_\phi)$, e.g., $U(f) = -\log(f)$.

Wired network:

$$\sum_{\{\phi : l \in P(\phi)\}} f_\phi \leq K_l$$

Define R to be routing matrix, so $R_{l,\sigma} = 1$ if flow ϕ crosses link l. Then $\sum_{\{\phi : l \in P(\phi)\}} f_\phi \leq K_l$ is the same as $\sum_\phi R_{l,\phi} f_\phi \leq K_l$, or we say $Rf \leq K$.

optimization problem

$$\begin{aligned} & \min \sum w_\phi U(f_\phi) \\ \text{s.t. } Rf - K & \leq 0 \\ f & \geq 0 \end{aligned}$$

.

2 Lagrange fomulation

Note that the objective is nonlinear, but constraints are linear define

$$L(f,\mu) = \sum w_\phi U(f_\phi) + \sum \mu_l \left(\sum_\phi R_{l,\phi} f_\phi - K_l \right)$$

where $\mu \geq 0$.

Thm: at optimal point $\nabla L = 0$. This means that $\nabla \sum w_\phi U(f_\phi) = -\sum \mu_l \nabla \left(\sum_\phi R_{l,\phi} f_\phi - K_l \right)$ for some $\mu_l \geq 0$.

Which means that the gradient of the objective function is the same as a weighted sum of the constraint gradients

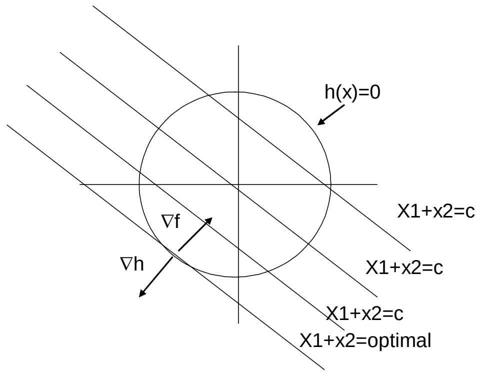

Figure 1:

3 Steepest descent

....

e.g.,

$$\min x_1 + x_2$$
$$\text{s.t. } x_1^2 + x_2^2 = 2$$

4 Lagrange theory

$$\begin{aligned} & \min U(x) \\ \text{s.t.} \ : \quad & h_i(x) = 0 \\ g_j(x) \ \leq \ & 0 \end{aligned}$$

$$L(x, \lambda, \mu) = U(x) + \sum \lambda_i h_i(x) + \sum \mu_j g_j(x)$$

then at the optimal point $\nabla_x L(x, \lambda, \mu) = 0$ and $\nabla_\lambda L(x, \lambda, \mu) = 0$.

$\nabla_\lambda L(x, \lambda, \mu) = 0$ implies that $h(x_i) = 0$, which it must.

consider $L(x, \lambda) = U(x) + \sum \lambda_i h_i(x)$, for all feasible x, we have $L(x, \lambda) = U(x)$, so working with U or L should be the same (for feasiable x).

It turns out that at the optimal point, $\mu_j = 0$ or $g_j(x) = 0$. So again, U and L are the same,

For constraints $g_i(x)$, if $\mu_i > 0$, then the constraint is active. note that if $\mu_j > 0$, then $g_j(x) = 0$, on the other hand, if $g_j(x) = 0$, then it does not mean that $\mu_j = 0$ (and hence, might not be active)

Note, the constraint is active implies that the constraint is reducing the ability to decrease the objective function any further.

5 Examples

$$\min \frac{1}{2}\left(x_1^2+x_2^2+x_3^2\right)$$
$$\text{s.t.} x_1+x_2+x_3 \leq -3$$
$$L(x,\mu) = \frac{1}{2}\left(x_1^2+x_2^2+x_3^2\right)+\mu\left(x_1+x_2+x_3+3\right)$$
$$\nabla L = \begin{pmatrix} x_1+\mu & x_2+\mu & x_3+\mu \end{pmatrix}$$

So we know that $x_i = x_j$.

If $\mu = 0$, then

$$\nabla L = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} = 0$$
$$x_1 = x_2 = x_3 = 0$$

but this does not satisfy the constraint.

or $\mu > 0$, in which case

$$\nabla L = \begin{pmatrix} x_1+\mu & x_2+\mu & x_3+\mu \end{pmatrix} = 0$$
$$x_1 = x_2 = x_3 = -\mu$$

Since the constraint is active, $x_1+x_2+x_3 = -3\mu = -3$, so $\mu = 1$

6 Examples

$$\begin{aligned} \min -\left(x_1 x_2 + x_2 x_3 + x_1 x_3\right) \\ \text{s.t.} x_1 + x_2 + x_3 &= 3 \end{aligned}$$

$$L\left(x, \lambda\right) = -\left(x_1 x_2 + x_2 x_3 + x_1 x_3\right) + \lambda \left(x_1 + x_2 + x_3 - 3\right)$$

$\nabla_x L = 0$

$$\begin{aligned} -x_2 - x_3 + \lambda &= 0 \\ -x_1 - x_3 + \lambda &= 0 \\ -x_1 - x_2 + \lambda &= 0 \end{aligned}$$

and $\nabla_\lambda L = 0$

$$x_1 + x_2 + x_3 = 3$$

this is four equations and four unknowns, which can be solved, $x_i = 1$ and $\lambda = 2$

homework: page 298 3.1.1

7 Necessary and sufficient conditions (equality constraints)

$\nabla_x L(x, \lambda, \mu) = 0$ and $\nabla_\lambda L(x, \lambda, \mu) = 0$.are necessary conditions, the optimal point must satisfy these. But, if these are satisfied, that does not mean that the it si a solution.

however, if $\nabla_x L(x, \lambda, \mu) = 0$ and $\nabla_\lambda L(x, \lambda, \mu) = 0$ and

$$y' \nabla^2_{xx} L(x^*, \lambda^*)\, y > 0 \text{ for all } y \text{ with } \nabla h_i(x)'\, y = 0,$$

then x is optimal.

If there are inequality constraints, the

$$\begin{aligned} y' \nabla^2_{xx} L(x^*, \lambda^*)\, y &> 0 \\ \text{for all } y \text{ with } \nabla h_i(x)'\, y &= 0, \\ \text{and } \nabla g_j(x)'\, y &= 0 \text{ for } j \text{ active.} \end{aligned}$$

8 Sensitvity

$$
\begin{aligned}
p(u,v) &= \min U(x) \\
\text{s.t.} h(x) &= u \\
g(x) &= v
\end{aligned}
$$

then

$$
\begin{aligned}
\nabla_u p(u,v) &= -\lambda(u,v) \\
\nabla_v p(u,v) &= -\mu(u,v)
\end{aligned}
$$

Note: if $\mu = 0$, then changing v does not impact the solution. This means that $g(x) = 0$ is not active, i.e., it is not impacting the solution

9 e.g.

$$\begin{aligned} & \min - \left(x_1 x_2 + x_2 x_3 + x_1 x_3\right) \\ \text{s.t.} x_1 + x_2 + x_3 &= 3 + u \end{aligned}$$

Found $\lambda = 2$ for $u = 0$.
redo $\nabla_x L = 0$

$$\begin{aligned} -x_2 - x_3 + \lambda &= 0 \\ -x_1 - x_3 + \lambda &= 0 \\ -x_1 - x_2 + \lambda &= 0 \\ x_1 + x_2 + x_3 &= 3 + u \end{aligned}$$

so $x_i = x_j$ and $x_i = (3 + u)/3$. So

$$\begin{aligned} p(u) &= -\left(x_1 x_2 + x_2 x_3 + x_1 x_3\right) \\ &= -\left(3\left((3+u)/3\right)^2\right) \\ &= -(3+u)^2/3 \end{aligned}$$

$$\begin{aligned} dp/du &= -2(3+u)/3 \\ dp/du|_{u=0} &= -2 = -\lambda \end{aligned}$$

10 Dual Problem

$$
\begin{aligned}
&\min f(c) \\
g_j(x) &\leq 0 \\
h_i(x) &= 0 \\
x &\in X
\end{aligned}
$$

Lagrangian

$$L(x, \lambda, \mu) = f(x) + \sum \lambda_i h_i(x) + \sum \mu_j g(x)$$

Define the dual function

$$q(\lambda, \mu) = \inf_{x \in X} L(x, \lambda, \mu)$$

The dual problem is

$$
\begin{aligned}
&\max q(\lambda, \mu) \\
\mu &\geq 0
\end{aligned}
$$

Note that there are no constraints. But there is both an inf and max

11 Dual Problem (II).

Thm: if the primal has a solution, then the dual does, and the value of the solutions are the same.

Thm: Let λ^*, μ^* be optimal, then x^* is optimal if and only if it is feasible and

$$x^* \in \arg\min L(x, \lambda^*, \mu^*)$$

Thm: $\mu_j^* g_j(x^*) = 0$, complementary slackness

Thm: $L(x^*, \lambda, \mu) \leq L(x^*, \lambda^*, \mu^*) \leq L(x^*, \lambda^*, \mu^*)$ saddle point

12 The Dual Problem - linear constraints

Primal problem

$$\begin{aligned} & \min f(x) \\ \text{s.t.} e_i' x &= d_i \text{ for } i = 1, ..m \\ a_j' x &\leq b_j \text{ for } j = 1, ..., r \\ x &\in X \end{aligned}$$

Lagrangian

$$L(x, \lambda, \mu) = f(x) + \sum \lambda_i (e_i' x - d_i) + \sum \mu_j \left(a_j' x - b_j\right)$$

Define the dual function

$$q(\lambda, \mu) = \inf_{x \in X} L(x, \lambda, \mu)$$

The dual problem is

$$\begin{aligned} & \max q(\lambda, \mu) \\ \mu &\geq 0 \end{aligned}$$

Note that there are no constraints. But there is both an inf and max

13 Dual of a Linear program

$$\min c'x$$
$$e_i'x = d_i \text{ for } i = 1, ..m$$
$$q(\lambda) = \inf_{x \geq 0} \sum_j \left(c_j - \sum_i \lambda_i e_{i,j} \right) x_j + \sum_i \lambda_i d_i$$

Dual problem:

$$\max q(\lambda).$$

Now, if $c_j - \sum \lambda_i e_{i,j} \geq 0$, then the infimum is attained for $x = 0$ and then $q(\lambda) = \sum \lambda_i d_i$. On the other hand, if $c_j - \sum \lambda_i e_{i,j} < 0$, for some j, we can pick x_j are large as desired and make $q(\lambda) = -\infty$. Thus, the max will surely have $c_j - \sum \lambda_i e_{i,j} \geq 0$. Hence the dual problem becomes

$$\max q(\lambda)$$
$$c_j - \sum \lambda_i e_{i,j} \geq 0$$

or

$$\max \sum \lambda_i d_i$$
$$c_j - \sum \lambda_i e_{i,j} \geq 0$$

which is the dual of the linear problem we saw long ago.

14 Network flow optimization

Let x_i be the flow over link i and the cost is $f_i(x_i)$, then

$$\begin{aligned} & \min \sum f_i(x_i) \\ a_j x - e_j &= 0 \text{ for all } j \text{ (conservation)} \end{aligned}$$

or

$$\begin{aligned} & \min \sum f_i(x_i) \\ \sum_i a_{j,i} x_i - e_j &= 0 \text{ for all } j \end{aligned}$$

Dual

$$
\begin{aligned}
q(\lambda) &= \inf_x \sum_i f_i(x_i) + \sum_j \lambda_j \left(\sum_i a_{j,i} x_i - e_j \right) \\
&= \inf_x \sum_i \mathrm{f}_i(x_i) + \sum_i \mathrm{x}_i \sum_j \lambda_j \mathrm{a}_{j,i} \text{-} \sum_j \lambda_j \mathrm{e}_j \\
&= \inf_x \sum_i \left(\mathrm{f}_i(x_i) + \mathrm{x}_i \sum_j \lambda_j \mathrm{a}_{j,i} \right) - \sum_j \lambda_j \mathrm{e}_j \\
&= \inf_x \sum_i \left(\mathrm{f}_i(x_i) + \mathrm{x}_i \sum_j \lambda_j \mathrm{a}_{j,i} - \frac{1}{n} \sum_j \lambda_j \mathrm{e}_j \right) \\
&= \sum_i q_i(\lambda) \\
q_i(\lambda) &= \inf_x \mathrm{f}_i(x_i) + \mathrm{x}_i \sum_j \lambda_j \mathrm{a}_{j,i} - \frac{1}{n} \sum_j \lambda_j \mathrm{e}_j
\end{aligned}
$$

So n seperate problems need to be solved, instead of a single large problem..

15 Network optimization

Let x_j be the flow along path i and Let R be the routing matrix so that $y_i = \sum R_{i,j} x_j$ is the flow across link j.

$$\begin{aligned} \min \sum_j U(x_j) \\ Rx \ \leq \ K \end{aligned}$$

e.g., $U(x_i) = -\log(x_i)$.
or

$$\begin{aligned} \min \sum_j U(x_j) \\ \sum_j R_{i,j} x_j \ \leq \ K_i \text{ for all } i \end{aligned}$$

Dual:

$$\begin{aligned} q(\mu) &= \inf \sum_j U(x_j) + \sum_i \mu_i \left(\sum_j R_{i,j} x_j - K_i \right) \\ &= \inf_x \sum_j \left(U(x_j) + x_j \sum_i R_{i,j} \mu_i - \sum_i \mu_i K_i \right) \\ &= \sum q_j(\mu) - \sum_i \mu_i K_i \\ q_j(\mu) &= \inf_{x_j} \left(U(x_j) + x_j \sum_i R_{i,j} \mu_i \right) \end{aligned}$$

dual problem

$$\max q(\mu)$$

.

16 Network optimization (II)

Now consider

$$\inf_{x_j} \left(U\left(x_j\right) + x_j \sum_i R_{i,j}\mu_i \right)$$

First, note that $\sum_i R_{i,j}\mu_i$ is the sum of the μ_i along the route used by flow i. If we interperate μ_i as the price of link i, then $\sum_i R_{i,j}\mu_i$ is the cost along the route used by flow j.

Note that flow j can solve this problem by getting the cost along the route it uses.

This cost is indicated back to the source as a loss probability or marking probability, then $\sum_i R_{i,j}\mu_i$ is the marking probability and $x_j \sum_i R_{i,j}\mu_i$ is the drop rate for flow j.

TCP is an interative solver of the above function.

On the other hand,

$$\begin{aligned} & \max_{\mu} \inf_{x} \sum_j \left(U\left(x_j\right) + x_j \sum_i R_{i,j}\mu_i - \sum_i \mu_i K_i \right) \\ = \; & \max_{\mu} \inf_{x} \sum_i \mu_i \left(\sum_j R_{i,j}x_j - K_i \right) + \sum_j U\left(x_j\right) \end{aligned}$$

$\sum_j R_{i,j}x_j$ total flow across link l. $\left(\sum_j R_{i,j}x_j - K_i\right)$ the amount over the capacity..

17 Optimal TCP/IP

Once x is known (for a given μ), we can solve $\max q(\mu)$.We can use an iterative solver of the form:

$$\begin{aligned} x_i(k+1) &= F_i\left(x_i(k), \sum_i R_{i,j}\mu_i\right) \text{ for each source } i \\ \mu_j(k+1) &= G_j\left(\mu_j(k), \sum_j R_{i,j}x_j\right) \text{ for each router } j \end{aligned}$$

$$\mu_j(k+1) = \begin{array}{c} \mu_j(k) + \gamma\left(\sum_j R_{i,j}x_j^* - K_i\right) \text{ if } \mu_j(k) > 0 \\ \mu_j(k) + \gamma\left(\sum_j R_{i,j}x_j^* - K_i\right)^+ \text{ if } \mu_j(k) = 0 \end{array}$$

where x^* is the optimal flow rates (note $\mu \geq 0$). But in this case, μ_j is the queue occupancy!

We can add some extra state

$$\begin{aligned} x_i(k+1) &= F_i\left(x_i(k), \sum_j R_{i,j}\mu_j\right) \text{ for each source } i \\ \mu_j(k+1) &= G_j\left(\mu_j(k), \sum_i R_{i,j}x_i(k), v_j(k)\right) \text{ for each router } j \\ v_j(k+1) &= H_j\left(\mu_j(k), \sum_i R_{i,j}x_i(k), v_j(k)\right) \end{aligned}$$

Different TCP and AQM are different F, G, and H.