EE 364 a

Wentao Jiang
wentao@Stanford
~~06165416~~ 06165416

3.13. $D_{KL}(u,v) = f(u) - f(v) - \nabla f(v)^T(u-v)$, $f(v) = \sum_i v_i \log v_i$

Because the negative entropy is strictly convex,

$\therefore$ $f(u) \geq f(v) + \nabla f(v)^T(u-v)$ and the equality is at $u=v$.

i.e., $D_{KL}(u,v) > 0$ and $D_{KL}(u,v) = 0$ if only if $u=v$.

3.36 Conjugate of f is $f^*(y) = \sup_{x \in \mathbf{dom} f} (y^T x - f(x))$

(a) Consider $n=2$ first. $f^*(y) = \sup_{x \in R^2} (x_1 y_1 + x_2 y_2 - \max(x_1, x_2))$

If $x_1 \geq x_2$, $f^*(y) = \sup_{\substack{x \in R^2 \\ x_1 \geq x_2}} (x_1 y_1 + x_2 y_2 - x_1)$

$= \sup_{x_1 \geq x_2} (-x_1(1-y_1) + x_2 y_2)$

$\therefore$ If $\begin{cases} y_2 > 0 \\ y_1 - 1 > 0 \end{cases} \Rightarrow f^*(y) = \infty$ If $\begin{cases} y_2 < 0 \\ y_1 - 1 < 0 \end{cases} \Rightarrow f^*(y) = \infty$

If y_1 or $y_2 < 0$, we could chose $\begin{cases} x_1 < 0 \\ x_2 = 0 \end{cases}$ or $\begin{cases} x_1 = 0 \\ x_2 < 0 \end{cases}$ so that $f^*(y) = \infty$. This easily generalize to R^n.

Since $f^*(y) = \sup_{x \in R^n} (x^T y - \max_i(x)) \leq \sup_x [\max(x)(\sum_i y_i - 1)]$ for $y \succeq 0$

where the equal sign happens for $x = x_0 \cdot \vec{1}$

$\therefore f^*(y) = \infty$ if $\sum_i y_i \neq 1$ and $f^*(y) \leq 0$ for $\sum_i y_i = 1$

$\therefore f^*(y) = \begin{cases} 0 & y \succeq 0 \text{ and } \sum_i y_i = 1 \\ \infty & \text{other vise.} \end{cases}$

(b) For $r=1$, it's the same as (a).

For general r, similarly $f^*(y) = \sup_x (x^T y - \sum_i^r x_{[i]})$

if $\exists k$, $y_k < 0$, then $\exists x$ such ~~$\leq \sup_x [\max(x)(\sum_i y_i$~~

that $x_k < 0$, $x_i = 0$ for $i \neq k$ leads to $f^*(y) = \infty$.

For $y \succeq 0$, let's look at $x = x_0 \cdot \vec{1}$ first.

$f^*(y) = \sup_x (x_0(\sum_i y_i - r))$

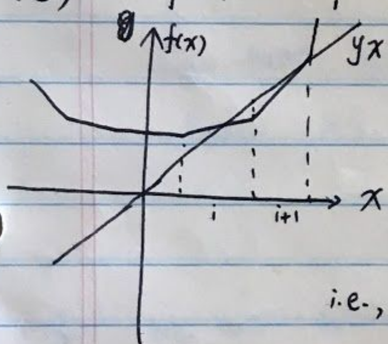
Hence if $\sum_i y_i > r$ or $\sum_i y_i < r$, $f^*(y) = \infty$

If $\sum_i y_i = r$, if $\exists k$, $y_k > 1$, then use the same x such that $x_k > 0$, $x_i = 0$ for $i \neq k$ leads to $f^*(y) = \infty$

$\therefore 0 \preceq y \preceq 1$. For y in this range, $x^T y \le \sum_i x_{[i]}$, i.e. $f^*(y) \le 0$ $\therefore$ $f^*(y) = 0$ for $\sum_i y_i = r$ when $x = x_0 \cdot \vec{1}$.

$$\therefore f^*(y) = \begin{cases} 0 & 0 \preceq y \preceq 1, \ \sum_i y_i = r \\ \infty & \text{otherwise.} \end{cases}$$

(c) $f(x) = \max_i (a_i x + b_i)$ is a piecewise-linear function



From the left figure, it's easy to see if $y > \max(a)$ or $y < \min(a)$, $f^*(y) = \infty$

For $a_i \le y < a_{i+1}$, maximum occure ~~at the~~ between segment i and $i+1$, where $a_i x + b_i = a_{i+1} x + b_{i+1}$

i.e., $x = \dfrac{b_{i+1} - b_i}{a_{i+1} - a_i}$

$$f^*(y) = xy - a_i x - b_i = \frac{b_{i+1} - b_i}{a_{i+1} - a_i}(y - a_i) - b_i$$

i.e., $f^*(y)$ is a piece wise-linear function as given above.

(d) $f(x) = x^p$, $f^*(y) = \sup_x (xy - x^p)$ $\quad h(x,y) = xy - x^p$ is differentiable

$\partial_x h(x,y) = y - p x^{p-1}$. For $y < 0$, $\partial_x h < 0$. $f^*(y) = h(0,y) = 0$

For $y > 0$, $\partial_x^2 h = -p(p-1)x^{p-2} < 0$, $\partial_x h = 0$ gives $x_0 = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$

$\therefore f^*(y) = h(x_0, y) = (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}}$

$$\text{i.e., } f^*(y) = \begin{cases} 0 & y < 0 \\ (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}} & y \ge 0 \end{cases}$$

For $p < 0$, if $y > 0$, $\partial_x h > 0$, $f^*(y) = h(\infty, y) = \infty$

If $y < 0$, $\partial_x^2 h = -p(p-1)x^{p-2} < 0$, $\partial_x h = 0$ gives $x_0 = \left(\frac{y}{p}\right)^{\frac{1}{p-1}}$

$$\therefore f^*(y) = \begin{cases} (p-1)\left(\frac{y}{p}\right)^{\frac{p}{p-1}} & y \le 0 \\ \infty & y \ge 0 \end{cases}$$

4.9 Make substitution $y = Ax$

$\therefore$ The problem is minimize $c^T A^{-1} y$

subject to $y \preceq b$

If $c^T A^{-1} \npreceq 0$, then $c^T A^{-1} y$ is unbounded below.

If $c^T A^{-1} \preceq 0$, then $p^* = c^T A^{-1} b$

$\therefore$ i.e., $p^* = \begin{cases} c^T A^{-1} b & A^{-T} c \preceq 0 \\ -\infty & \text{otherwise.} \end{cases}$

4.11

(a) l_∞ norm selects the maximum component(s).

Hence it's equivalent to minimizing y

subject to $Ax - b \preceq y \cdot \vec{1}$

$Ax - b \succeq -y \cdot \vec{1}$

(b) We are minimizing $\sum_i |Ax - b|_i$

i.e., minimize $\sum_i y_i$

subject to $Ax - b \succeq -y$

$Ax - b \preceq y$

(c) Similar to (b), it's equivalent to minimize $\sum_i y_i$

subject to $Ax - b \succeq -y$

$Ax - b \preceq y$

$-1 \preceq x \preceq 1$

(d) Similar to (c), it's equivalent to minimize $\sum_i y_i$

subject to $-y \preceq x \preceq y$

$-1 \preceq Ax - b \preceq 1$

(e) From (a) and (b), it's equivalent to minimize $\sum_i y_i + z$

subject to $-y \preceq Ax - b \preceq y$

$-z \preceq x \preceq z$, where $z \in R$.

5.13 (a) The Lagrangian is

$$L(x,\lambda,\nu) = c^T x + \lambda^T (Ax-b) + \sum_i \nu_i (1-x_i) x_i$$
$$= x^T(-\mathrm{diag}(\nu))x + (c+A^T\lambda+\nu)^T x - \lambda^T b$$

which we minimize over x to get the dual function:

$$g(\lambda,\nu) = \begin{cases} -\infty & \nu \not\preceq 0 \\ -\frac{1}{4}\sum_i \frac{[(c+A^T\lambda+\nu)_i]^2}{\nu_i^2} - \lambda^T b & \nu \prec 0 \end{cases}$$

∴ The resulting dual problem is maximize

$$-\frac{1}{4}\sum_i \frac{[(c+A^T\lambda+\nu)_i]^2}{\nu_i^2} - \lambda^T b$$

subject to $\nu \succeq 0$, $\lambda \succeq 0$

(b) The dual function of the LP relaxation is

$$L(x,\alpha,\beta,\gamma) = c^T x + \alpha^T(Ax-b) - \beta^T x + \gamma^T(x-\vec{1})$$
$$= (c^T + \alpha^T A - \beta^T + \gamma^T)x - \alpha^T b - \gamma^T\cdot\vec{1}$$

which is affine ∴ minimize over x:

$$g(\alpha,\beta,\gamma) = \begin{cases} -\alpha^T b - \gamma^T\cdot\vec{1} & c + A^T\alpha - \beta + \gamma = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

∴ The dual problem is maximize $-\alpha^T b - \gamma^T\cdot\vec{1}$

subject to $c + A^T\alpha - \beta + \gamma = 0$

$\alpha \succeq 0$, $\beta \succeq 0$, $\gamma \succeq 0$

I can't see how they are equivalent.

A3.8 epi f is $f(x) \le t$. Denote $y = Ax+b$ and $M = P_0 + \cdots + x_n P_n$

$$\therefore f(x) \le t \Rightarrow y^T M^{-1} y \le t$$

Since $\mathrm{dom}\, f = \{x \in R^n \mid M \succ 0\}$, $f(x) > 0$ ~~$t>0$ if epi f~~

~~$t>0$ if epi $f \neq \emptyset$.~~

$\therefore$ From A5.5, consider $X = \begin{bmatrix} M & y \\ y^T & t \end{bmatrix}$

$$\therefore X \succ 0 \Leftrightarrow \begin{cases} M \succ 0 \\ s = t - y^T M^{-1} y > 0 \end{cases}$$

i.e., the symmetric matrix $F(x,t)$ we are looking for

is $F(x,t) = \begin{bmatrix} P_0 + x_1 P_1 + \cdots + x_n P_n & Ax+b \\ (Ax+b)^T & t \end{bmatrix}$.

A3.26 Consider $y_j = \sqrt{x_j}$, then $f_i(x) = \frac{1}{2} \sum_k (P_i)_{kk} y_k + \sum_{jk} (P_i)_{jk} \sqrt{y_j y_k}$

$$+ \sum_k (q_i)_k \sqrt{y_k} + r_i$$

Since $\sqrt{y_k}$ and $\sqrt{y_j y_k}$ are concave, and $(P_i)_{jk} < 0$, $q_i \prec 0$.

$\therefore f_i$ is a convex function of y

$\therefore$ Now it's a convex problem.

A4.5 The Lagrangian is $L(x, \lambda, \mu) = \frac{1}{2} x^T x - x^T a + \frac{1}{2} a^T a$

$$+ \lambda (\vec{1}^T x - 1) - \mu^T x.$$

We only consider $a \succ 0$

and $x \succ 0$, which is sufficient for general a.

$L(x, a, \mu) = \frac{1}{2} x^T x + (\lambda \cdot \vec{1} - a - \mu)^T x + \frac{1}{2} a^T a - \lambda$

Minimize over x gives $g(a, \mu) = -\frac{1}{2} \sum_i (\lambda - a_i - \mu_i)^2 + \frac{1}{2} a^T a - \lambda$

at $x = \mu + a - \lambda$.

$\therefore$ The dual problem is maximize $g(\lambda, \mu)$

subject to $\lambda \ge 0$, $\mu \succeq 0$.

If the original constraint is $\|x\|_1 = 1$, then $\lambda \in R$.

For optimal $x = x^*$, from the KKT condition, if $x_i = \mu_i + a_i - \lambda \neq 0$,

then $\mu_i = 0$, $\therefore x_i = a_i - \lambda$, if $x_i \neq 0$.

P5 i.e., if we know the non zero components of x and λ, the problem is solved.

The dual problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\sum_i (\lambda - a_i - \mu_i)^2 + \lambda \\ \text{subject to} & \mu \succeq 0 \end{array}$$

$$\Leftrightarrow \begin{array}{ll} \text{minimize} & \frac{1}{2}\|\nu - \lambda\|_2 + \lambda \\ \text{subject to} & \nu \succeq a \end{array}$$

A 16.9 p_t: price, u_t: usage, q_t: energy stored, c_t: charging

(a) The optimization problem is

$$\begin{array}{ll} \text{minimize} & p^T(u+c) \\ \text{subject to} & -D \preceq c \preceq C \quad \text{and} \quad u+c \succeq 0 \\ & 0 \preceq q \preceq Q \\ & q_{i+1} = q_i + c_i \\ & q_1 = q_T + c_T \end{array}$$

The problem is how to implement the last two constraints. It turns out that they can be straight forward typed in MATLAB. See attached code and plots.

The endpoint is limited by the charge/discharge rate, when the capacity is enough, the optimal strategy is to charge as much as possible during low price intervals and use it during high price time intervals. The benefit is directly limited by the total energy stored, which is limited by the change/discharge rate.

Code for A 16.9:

```
randn('seed', 1);
T = 96; % 15 minute intervals in a 24 hour period
t = (1:T)';
p = exp(-cos((t-15)*2*pi/T)+0.01*randn(T,1));
u = 2*exp(-0.6*cos((t+40)*pi/T) -0.7*cos(t*4*pi/T)+0.01*randn(T,1));
figure;
plot(t/4, p);
hold on
plot(t/4,u,'r');

%
Q = 35; C = 3; D = -3;
%%
cvx_begin
variable c(T)
variable q(T)
minimize(p'*(u+c))
subject to
0<=q<=Q;
D<=c<=C;
sum(c)==0;
q(2:T)==q(1:T-1) + c(1:T-1);
q(1)==q(T)+c(T);
u+c>=0;
cvx_end

%%
plot(t/4, c);
plot(t/4, q);
legend({'p','u','c','q'});

%% total cost versus capacity
vs_C3D3 = [];
vs_C1D1 = [];
Qs = [1:10:200];
for ii = 1:length(Qs)
    Q = Qs(ii);
    [v, c, q] = cvx_solve_c_q(Q, 3, -3);
    vs_C3D3(ii) = v;
    [v, c, q] = cvx_solve_c_q(Q, 1, -1);
```

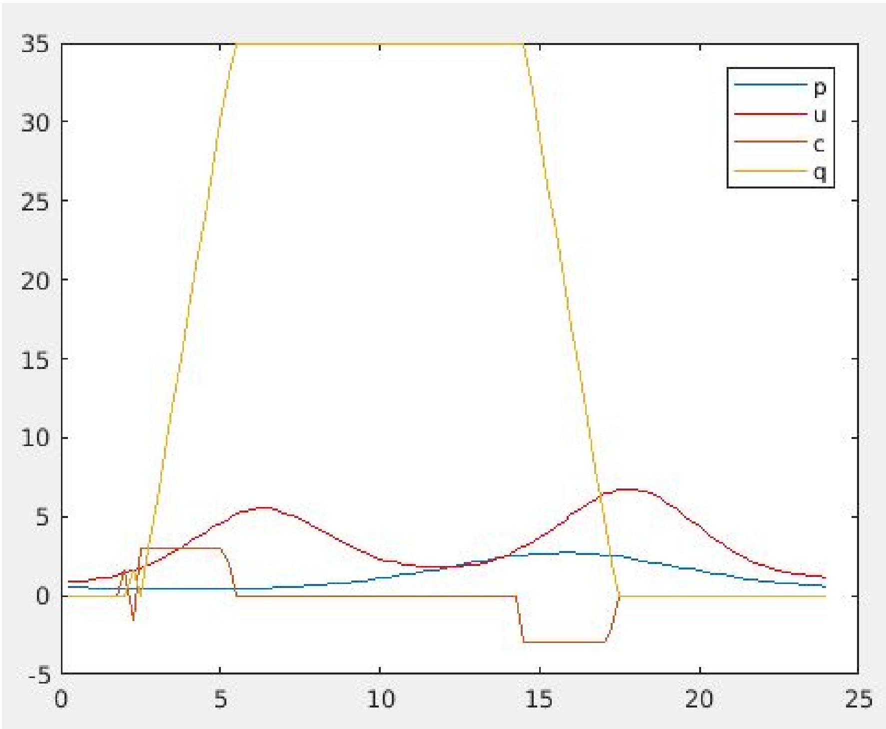
```
    vs_C1D1(ii) = v;
end

figure;
plot(Qs, vs_C3D3);
hold all;
plot(Qs, vs_C1D1);
xlabel('storage capacity');
ylabel('min total cost');
legend({'C=3,D=-3','C=1,D=-1'})

%%
function [cvx_optval, c, q] = cvx_solve_c_q(Q, C, D)
randn('seed', 1);
T = 96; % 15 minute intervals in a 24 hour period
t = (1:T)';
p = exp(-cos((t-15)*2*pi/T)+0.01*randn(T,1));
u = 2*exp(-0.6*cos((t+40)*pi/T) -0.7*cos(t*4*pi/T)+0.01*randn(T,1));

cvx_begin
variable c(T)
variable q(T)
minimize(p'*(u+c))
subject to
0<=q<=Q;
D<=c<=C;
sum(c)==0;
q(2:T)==q(1:T-1) + c(1:T-1);
q(1)==q(T)+c(T);
u+c>=0;
cvx_end
end
```

Plot of p, u, c and q:



Total cost versus storage capacity: