Lec12

Friday, February 3, 2023 9:01 AM

Transformation to a convex problem

Saturday, January 17, 2009 4:38 PM

What can you do if your initial formulation does not correspond to a convex problem?

- Sometimes your formulation can be converted to an equivalent formulation that is convex!
- Experience helps.

---

① Change of variables

$$\begin{array}{ll} \min & f_0(x) \\ \text{sub to} & f_i(x) \le 0 \\ & h_i(x) = 0 \end{array} \qquad\qquad \begin{array}{ll} \min & f_0(g(z)) \\ \text{sub to} & f_i(g(z)) \le 0 \\ & h_i(g(z)) = 0 \end{array}$$

need $g(\mathrm{dom}\, g) \supseteq D$

Example:

e.g.

$$\boxed{\begin{array}{ll} \min & x^2+y \\ \text{sub to} & \log(y+z) \le x \\ & y^2+z^2 \le 1 \end{array}}$$

ⓐ A constraint

$$\log(y+z) \le x$$

$$\Leftrightarrow \quad \log(y+z) - x \le 0 \qquad \text{not convex}$$

$$\Leftrightarrow \quad \log(e^{y'} + e^{z'}) - x \le 0 \qquad \text{convex}$$

Caveate: Other functions of y & z may not be convex in y' & z'. $(y+z=1)$

ⓑ

$$\min \quad \sqrt{x_1} - \sqrt{x_2}$$

$$\text{sub to} \quad x_1 + x_2 \le 1$$

$$\Leftrightarrow \quad \min \quad x_1' - x_2'$$

sub to $x_1'^2 + x_2'^2 \leq 1$, $x_1' \geq 0$, $x_2' \geq 0$

---

② Transformation of functions

Example

ⓐ a constraint that

$$\frac{x_1}{1+x_2^2} \geq 1$$

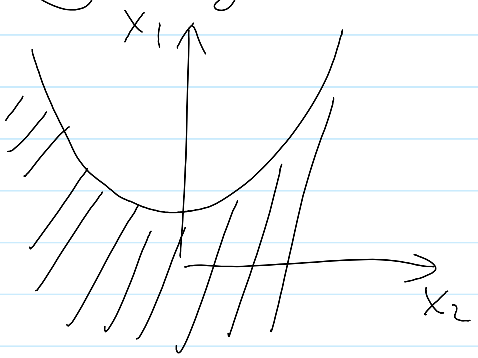
$$\Leftrightarrow \quad 1 - \frac{x_1}{1+x_2^2} \leq 0 \qquad \text{not convex}$$

$$\Leftrightarrow \quad 1 + x_2^2 - x_1 \leq 0 \qquad \text{convex.}$$

- This trick does not really change the constraint set

$$\frac{x_1}{1+x_2^2} \leq 1$$

$$x_1 \leq 1 + x_2^2$$



ⓒ

$$\min \sqrt{x_1 + x_2} \qquad \text{non-convex}$$

$$\Leftrightarrow \min \; x_1 + x_2 \qquad \text{convex}$$

sub to $x_1 + x_2 \geq 0$

---

③ Equality constraints can be converted to inequality constraints, vice versa, if you know additional structure of the problem.

$$\begin{array}{lll} \min & X_1+X_2 & \\ \text{sub to} & X_1^2+X_2^2=1 & \text{non-linear} \end{array}$$

$$\Leftrightarrow \begin{array}{lll} \min & X_1+X_2 & \\ \text{sub to} & X_1^2+X_2^2\le 1 & \text{convex} \end{array}$$

Since the optimal point must lie on the boundary of the new constraint set.

- Another example: combining with simplifying the obj or constraint func.

$$\min\ f_0(g(x)) \Leftrightarrow \begin{array}{ll} \min & f_0(y) \\ \text{sub to} & g(x)=y \end{array}$$

$$\Downarrow \text{ if } f_0 \text{ increasing.}$$

$$\begin{array}{ll} \min & f_0(y) \\ \text{sub to} & g(x)\le y \end{array}$$

Skip the rest

In addition, the following transformation may help you to simplify the search for a convex form. (Even though these transformation alone may not change convexity)

④ Implicit constraints can be made explicit

Suppose $f(x) = \begin{cases} x^T x & Ax=b \end{cases}$

$+\infty$ otherwise

Then

$$\min f(x) \iff \begin{array}{l} \min\ x^T x \\ \text{sub to}\ Ax = b \end{array}$$

⑤ Simplify objective/constraint functions and introduce additional constraints

$$\begin{array}{l} \min\ f_0(A_0x + b_0) \\ \text{sub to}\ f_i(A_ix + b_i) \le 0 \\ \qquad\quad h_i(x) = 0 \end{array} \iff \begin{array}{l} \min\ f_0(y_0) \\ \text{sub to}\ f_i(y_i) \le 0 \\ \qquad\quad A_ix + b_i = y_i \\ \qquad\quad h_i(x) = 0 \end{array}$$

- Conversely, equality constraints can be absorbed into the function

$$\begin{array}{l} \min\ f_0(x) \\ \text{sub to}\ f_i(x) \le 0 \\ \qquad\quad Ax = b \end{array} \iff \begin{array}{l} \min\ f_0(Fz + x_0) \\ \text{sub to}\ f_i(Fz + x_0) \le 0 \end{array}$$

↑ write this constraint as $x = g(z) = F(z) + x_0$

⑥ Epigraph form

$$\begin{array}{l} \min\ f_0(x) \\ \text{sub to}\ f_i(x) \le 0 \end{array} \iff \begin{array}{l} \min\ t \\ \text{sub to}\ f_0(x) - t \le 0 \\ \qquad\quad f_i(x) \le 0 \end{array}$$

⑦ slack variables
(convert inequality constraints to

equality constraints)

$$\begin{array}{ll} \min & f_0(x) \\ \text{sub to} & f_i(x) \leq 0 \end{array} \iff \begin{array}{ll} \min & f_0(x) \\ \text{sub to} & f_i(x) + s_i = 0 \\ & s_i \geq 0 \end{array}$$ if f_i is linear

---

All of these may make an otherwise non-convex problem to a convex problem.

At other times you may have to redefine the problem, choose the right objectives, etc.

Experience really helps. (An art rather than a science.)

More examples follow.

(30)

Geometric programming - skip

Sunday, February 01, 2009 12:55 PM

- There are many known forms of convex problems
  - linear program
  - linear-fractional program
  - quadratic program
  - second order cone program
  - etc
- See Boyd Ch 4.3, 4.4

Geometric program

- Useful when there are products/fractions

Definitions

- function

  $$f(x) = c \cdot x_1^{a_1} \cdot x_2^{a_2} \cdots x_n^{a_n}, \quad c > 0, \; a_i \in \mathbb{R}$$

  is called a monomial function, or simply a monomial.

- A sum of monomials

  $$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}$$

  where $c_k > 0$, is called a posynomial.

- A geometric program is of the form

$$\begin{aligned} \min \quad & f_0(x) \\ \text{sub to} \quad & f_i(x) \le 1 \qquad i = 1, \cdots, m \\ & h_i(x) = 1 \qquad i = 1, \cdots, p \\ & x > 0 \end{aligned}$$

where $f_0, \cdots, f_m$ are posynomials and $h_1, \cdots h_p$ are monomials

- The constraint $x>0$ is always assumed.

Example

ⓐ

$$\begin{aligned} &\max \quad x/y \\ &\text{sub to} \quad 2<x<3 \\ &\qquad x^2+3y/z \le \sqrt{y} \\ &\qquad x/y = z^2 \end{aligned}$$

$$\Leftrightarrow \quad \begin{aligned} &\min \quad x^{-1}y \\ &\text{sub to} \quad 2x^{-1} < 1 \\ &\qquad \frac{x}{3} < 1 \\ &\qquad x^2y^{-\frac{1}{2}} + 3y^{\frac{1}{2}}z^{-1} \le 1 \\ &\qquad xy^{-1}z^{-2} = 1 \end{aligned}$$

Positive-coefficient polynomials ≤ 1

$= 1$

Convert to a Convex Program

- Let $x_i = e^{y_i}$

$$\Rightarrow \quad \begin{aligned} &\min \quad \sum_{k=1}^{K_0} e^{a_{0k}^T y + b_{0k}} \\ &\text{sub to} \quad \sum_{k=1}^{K_i} e^{a_{ik}^T y + b_{ik}} \le 1 \\ &\qquad e^{g_i^T y + h_i} = 1 \end{aligned}$$

- Take log.

$$\Rightarrow \quad \begin{aligned} &\min \quad \log\left[\sum_{k=1}^{K_0} e^{a_{0k}^T y + b_{0k}}\right] \\ &\text{sub to} \quad \log\left[\sum_{k=1}^{K_0} e^{a_{ik}^T y + b_{ik}}\right] \le 0 \\ &\qquad g_i^T y + h_i = 0 \end{aligned}$$

Ex)

$$\begin{aligned} &\min \quad e^{-x'}e^{y'} \\ &\text{sub to} \quad 2e^{-x'} < 1 \\ &\qquad \frac{e^{x'}}{3} < 1 \\ &\qquad \begin{cases} e^{2x'}e^{-\frac{1}{2}y'} + 3e^{\frac{1}{2}y'}e^{-z'} \le 1 \\ e^{x'}e^{-y'}e^{-2z'} = 1 \end{cases} \end{aligned}$$

$$\begin{cases} \log\left[\qquad \right] \le 0 \\ x'-y'-2z' = 0 \end{cases}$$

- i.e., posynomials $\Leftrightarrow$ convex functions
  monomials $\Leftrightarrow$ linear functions
- See Boyd Ch 4.5

㊺

Special cases that are convex

Monday, January 30, 2023 4:33 PM

① Fixed link rate → Pure power control

- We want a certain target rate to be reached for r_l^k $\nwarrow$ given before-hand
  - This target rate may be indep. of the channel k.
  - Good channel → Smaller power
    bad channel → Higher power

- In order for

$$w\lg\left(1+\frac{P_l^k g_{ll}}{\sum_{h\neq l} P_h^k g_{hl}+N_l}\right)\geq r_l \overset{\circ}{=} w\lg(\gamma_l)$$

we just need

$$1+\frac{P_l^k g_{ll}}{\sum_{h\neq l} P_h^k g_{hl}+N_l}\geq \gamma_l \qquad (\gamma_l>1)$$

$$P_l^k g_{ll}\geq(\gamma_l-1)\left[\sum_{h\neq l} P_h^k g_{hl}+N_l\right] \qquad (*)$$

- This is the formulation for power control problems

$$\min \quad \sum_l P_l^k$$

sub to (*)

- A linear program.

Ref: Foschini & Miljanic, "A simple distributed autonomous power control algorithm & its convergence," IEEE Transactions on vehicular Technology v.24. Nov 1993, pp 641-646.

② No interference

- Link rate a function of its own power

$$g_k(P_k) = W \cdot \lg\left(\frac{P_k \cdot g_{kk}}{N_k}\right)$$

  - concave!

Ex: Water-filling for fading channels

- Boyd p245
- Discussed in lecture 1.

$$\max_{p_1, p_2, \dots, p_m} \quad \sum_{k=1}^{M} \rho_k \log\left(1 + \frac{g_k p_k}{N}\right)$$

$$\text{sub to} \quad \sum_{k=1}^{M} \rho_k p_k \le P_0$$

③ No power control

- link rate is fixed
- Can still add ~~the~~ possibility of scheduling

Ex: Opportunistic scheduling for multi-user systems with fading

Ref. Transmission Scheduling for Efficient Wireless Network Utilization, by X. Lin, E.K.P. Chong, and N.B. Shroff, IEEE Infocom, April 2001.

④ Approximation (e.g. High SINR)

Water-filling for fading channels - skip

Saturday, January 24, 2009 6:17 PM

- Boyd p245
- Discussed in lecture 1.
- Single channel / link / user

$$r = \log\left(1 + \frac{g P}{N}\right)$$

  - g : channel gain
  - P : transmission power
  - N : background noise. No interference!
- But the channel gain changes in time.
  - Assume that the channel gain can be in one of K values, $g_1, g_2, \cdots, g_K$
  - $q_k = \Pr[g = g_k]$.
- How do we control the transmission strategy.
  - Specifically, suppose our assigned power must be no greater than P_0 on average.
  - Use same power for all channel gains?

$$\sum_k q_k \log\left(1 + \frac{g_k \cdot P_0}{N}\right).$$

- Or, we can vary the transmission power at each channel gain

- Intuition: If the average transmission power is limited, then we may want to allocate more power for good channels, and less power for bad channels, so that the total transmission rate can be bigger

- Note that this is the opposite to the intuition for power control!

  - May produce larger delay.

- Let P_K be the transmission power when the channel gain is g_K.

  - average power $\sum_{K=1}^{M} \rho_K \cdot P_K$

  - total transmission rate

$$\sum_{K=1}^{M} \rho_K \log\left(1+\frac{g_K \cdot P_K}{N}\right)$$

$$\max_{P_1, P_2, \cdots, P_m} \sum_{K=1}^{M} \rho_K \log\left(1+\frac{g_K P_K}{N}\right)$$

$$\text{sub to} \quad \sum_{K=1}^{M} \rho_K P_K \le P_0 \quad , \quad P_K \ge 0$$

  - A convex problem

Solution :

- Since the objective is increasing in P_k, the solution must be at

$$\sum_{k=1}^{M} \gamma_k P_k = P_0 .$$

- Assume in addition that $P_k > 0$.

- The normal vector is $[\gamma_1, \cdots, \gamma_M]$

- The gradient of $-\sum_k \gamma_k \log\left(1+\frac{\gamma_k P_k}{N}\right)$

is

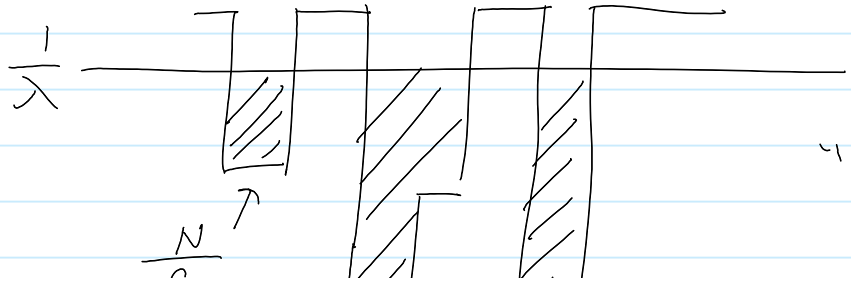
$$\left[-\frac{\gamma_k \cdot \frac{\gamma_k}{N}}{1+\frac{\gamma_k P_k}{N}}\right]_{k=1,\cdots,M}$$

$$\Rightarrow \quad \frac{\gamma_k \cdot \frac{\gamma_k}{N}}{1+\frac{\gamma_k P_k}{N}} = \lambda \gamma_k$$

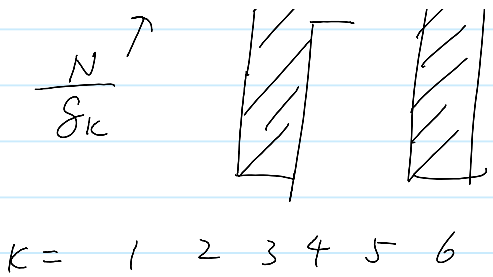
$$\Rightarrow \quad 1+\frac{\gamma_k}{N} \cdot P_k = \frac{\gamma_k}{N} \cdot \frac{1}{\lambda}$$

$$P_k = \left(\frac{1}{\lambda} - \frac{N}{\gamma_k}\right)$$

- λ is chosen such that $\sum P_k = P_0$



"water-filling"



Water-filling

— $\delta_k \uparrow \quad \frac{1}{\delta_k} \downarrow \quad P_k \uparrow$

— What if $\frac{1}{\lambda} - \frac{N}{\delta_k} < 0$?

(Wait when we discuss duality.)

⑩

Opportunistic scheduling - Long - skip

Saturday, January 24, 2009 10:23 PM

Fading + Multi-user

- The channel can be in one of K states
- Assume two users for simplicity

If the system is at state k, and user i ($\in 1,2$) is chosen for transmission, the payoff to the chosen user is U_i^k. The payoff is zero for the other user.

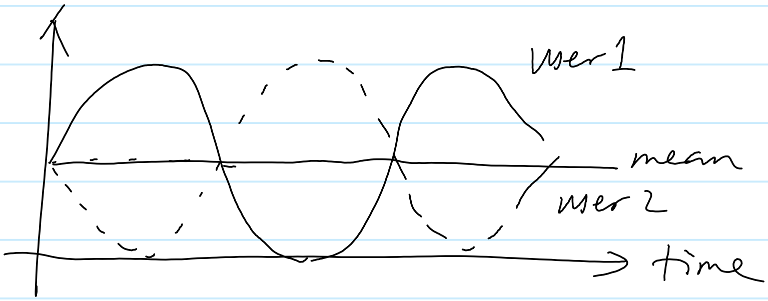
- Power is otherwise fixed before-hand.

Ⓠ How to choose which user to transmit?

- If we choose each user uniformly randomly with prob. 1/2

$$\frac{1}{2} E[U_1^k] + \frac{1}{2} E[U_2^k]$$

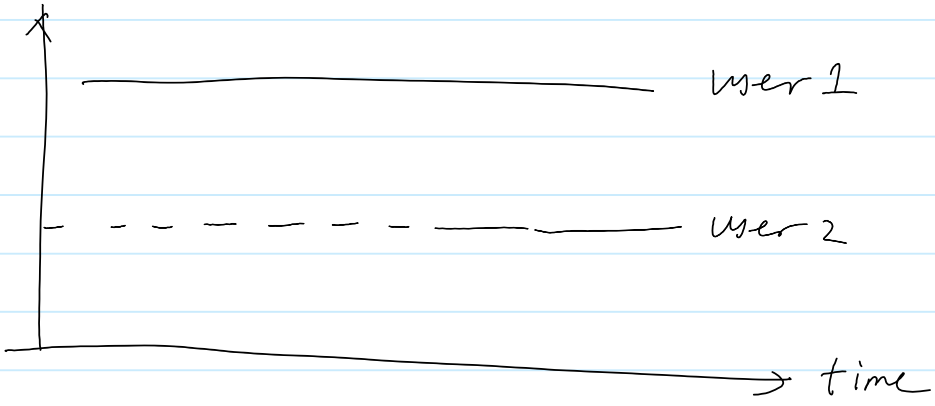
- average performance of each user.



- If we choose the user with better rate

  - achieve higher overall performance.
  - improvement for both users

- Should we always pick the user with higher rate



  - The worse user could be starved

  - Need to consider the tradeoff btw performance & fairness.

- Let $q_K = Pr\{\text{channel is at state } k\}$

- Power is fixed, so the rate is completely determined by the channel state

- U_i^k : utility

- $Q_i^k = \begin{cases} 1 & \text{if the policy decides that user } i \text{ should be chosen at state } k \\ 0 & \text{o/w.} \end{cases}$

- The expected payoff to user i is

$$\sum_k f_k Q_i^k \cdot U_i^k$$

- The system payoff

$$\sum_i \sum_k f_k Q_i^k U_i^k$$

- Need a way to balance system payoff vs. individual payoff.

$\Rightarrow$ Impose a fairness constraint.

ⓐ Temporal fairness

- Fraction of time user i is scheduled

$$\sum_k f_k Q_i^k$$

- Fairness constraint in time : user i should be scheduled at least r_i

fraction of time.

$$\max \quad \sum_i \sum_k p_k Q_i^k u_i^k$$

$$\text{Sub to} \quad \sum_k p_k Q_i^k \geq r_i \quad \forall i$$

$$\sum_i Q_i^k = 1 \quad \forall k$$

$$Q_i^k \in [0, 1]$$

- Not a convex problem. Instead, relax $Q_i^k \in [0, 1]$

  Q_i^k : user i should be chosen with prob. Q_i^s if the state is k

- A linear problem.

Other types of fairness constraints:

ⓐ Temporal Constraint

ⓑ Utility Constraint

ⓒ fraction of utility constraint.

etc.

HW: Formulate into convex problems

Ref. Transmission Scheduling for Efficient Wireless Network Utilization, by X. Lin, E.K.P. Chong, and N.B. Shroff, IEEE Infocom, April 2001.

Approximation

Sunday, February 01, 2009 12:38 PM

High-SINR Approximation

$$g_l(\vec{p}) = \log\left(1 + \frac{p_l}{\sum_{j\neq l} p_j + N_l}\right) \simeq \log \frac{p_l}{\sum_{j\neq i} p_j + N_l}$$

- For simplicity, assume that each user has a link, and no fading.

$$\Rightarrow \quad \max \quad \sum_l r_l$$

$$\text{sub to} \quad r_l \leq \log \frac{p_l}{\sum_{j\neq l} p_j + N_l}, \quad l=1,\cdots,L$$

$$p_l \leq p_{max}$$

- change of variable

$$p_l = e^{Q_l}$$

$$\Rightarrow \quad \max \quad \sum_l r_l$$

$$\text{sub to} \quad r_l \leq Q^l - \log\left[\sum_{j\neq l} e^{Q_j} + N_l\right]$$

$$l=1,\cdots,L.$$

Ref: M. Chiang, "Balancing Transport and Physical Layer in Multi-hop wireless Networks, Joint Optimal Congestion & Power Control," IEEE JSAC, Vol. 23, No. 1, pp 104-116, 2005.

- Pitfalls
  - Solution may not lead to $\overset{\text{high}}{\wedge}$ S2NR
  - As a results, the quality of the solution may be worse than scheduling / time-interleaving

Example:

ⓐ "High-S2NR approximation":

- both users transmit at the same time. at rate

$$\log\left(1+\frac{P}{P+N}\right)$$

- Total rate

$$2\log\left(1+\frac{P}{P+N}\right) \simeq 2\log 2$$

ⓑ Time-interleaving

- One user transmit at a time
- Total rate

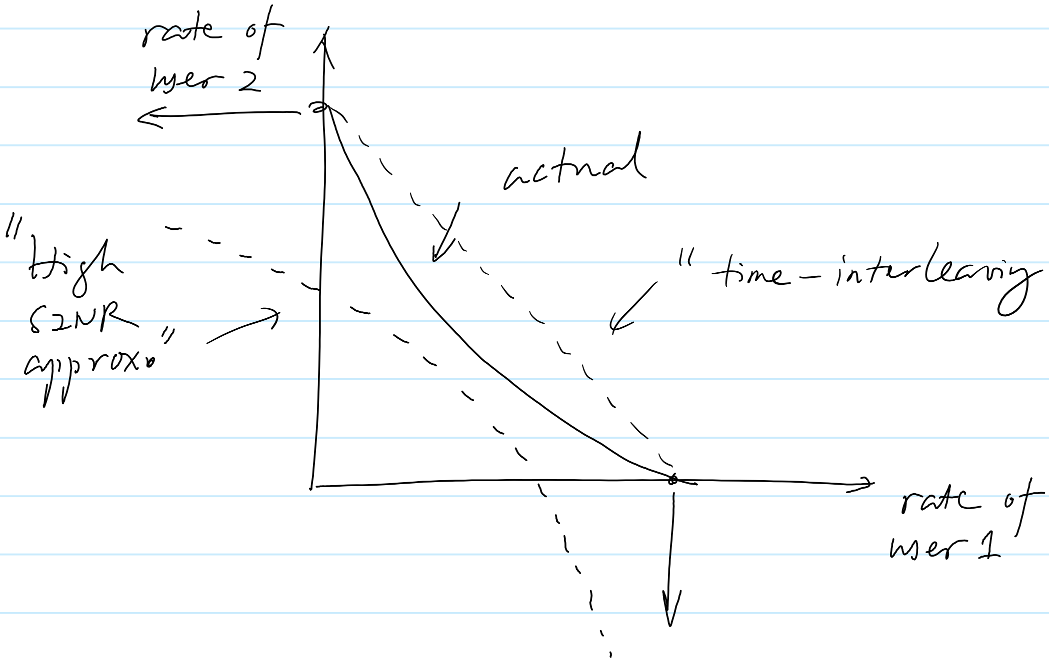
$$\log\left(1+\frac{P}{N}\right) \simeq \log\frac{P}{N}$$

$$\log\left(1+\frac{P}{N}\right) \simeq \log\frac{P}{N}$$

- Average rate per-user

$$\frac{1}{2}\log\left(1+\frac{P}{N}\right)$$

When N is small, option ⓑ is often much better!



Two lessons

① Convex approximations are not always better.

② Need to use your domain knowledge to pick the right formulation/approximation.

㉟

Linear Fractional

Sunday, February 15, 2015 9:50 PM

Linear-fractional programming

- Boyd p151
- $$\min \ \frac{c^Tx+d}{e^Tx+f} \qquad (1)$$

  $$\text{sub to} \quad e^Tx+f>0$$
  $$Gx \le h$$
  $$Ax = b$$

- Can be converted to a convex program
  - Let
    $$y = \frac{x}{e^Tx+f} \qquad z = \frac{1}{e^Tx+f} > 0$$
  - We want a one-to-one mapping between
    $$x \Leftrightarrow (y, z), \ z>0$$
    - For every x with $e^Tx+f>0$, the (y, z) can be obtained above
    - and we must have $e^Ty+fz=1$
    - Given (y,z), we may not be able to find x (since there are 2 equations

find x (since there are 2 equations but only 1 unknown)

- We may set
$$x = \frac{y}{z}$$
- But we need
$$z = \frac{1}{e^T \frac{y}{z} + f}$$
$$\Leftrightarrow \quad e^T y + f z = 1$$

- Thus, there is a 1-to-1 mapping between
$$x \ ; \quad e^T x + f > 0$$
and
$$(y, z) \ ; \quad z > 0 \ \& \ e^T y + f z = 1$$
- (1) can be converted to

$$\begin{aligned} \min \quad & c^T y + d z \\ \text{sub to} \quad & z > 0 \\ & e^T y + f z = 1 \\ & G y \le h z \\ & A y = b z \end{aligned}$$

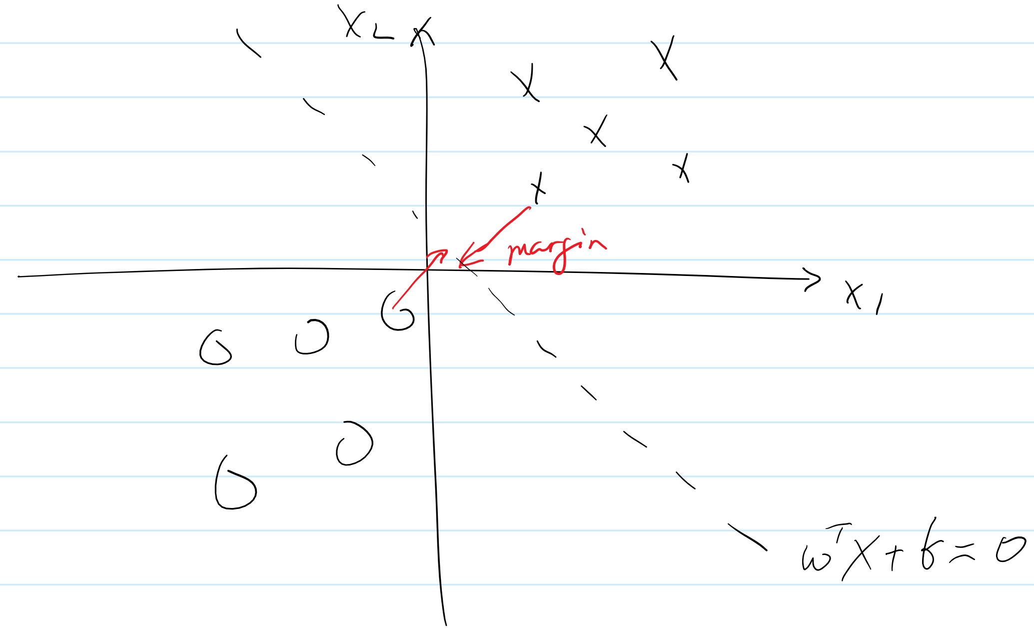
- (Can also show that replacing $z > 0$ by $z \geq 0$ will not change the objective value.)

- Basically $y=0$ corresponds to $x \to +\infty$

SVM for classification problems

Wednesday, February 1, 2023 12:15 PM

- Suppose that we are given a classification problem
  - (x_i, t_i) , $i=1,\dots,n$
    indep. variable ↑ ↑ label
  - $t_i \in \{-1, 1\}$ — 2 classes



- We want to learn a linear model that separates them

$$y_i = w^T x_i + b$$

- We want

$$y_i > 0 \iff t_i = 1 \qquad (*)$$

$$y_i < 0 \Leftrightarrow t_i = -1$$

- If there exists such (w,b) that satisfies $(*)$, the data is said to be linearly separable.

- We may still have many choices for (w,b) though

  $\Rightarrow$ Find the one with the highest margin

  $\Rightarrow$ "robust"

- For each x_i, the distance to the plane

$$w^T x + b = 0$$

is

$$\frac{|w^T x_i + b|}{\|w\|_2}$$

  - We call this the "margin"

- Our goal is

$$\max_{w,b} \min_{i} \frac{|w^T x_i + b|}{\|w\|_2}$$

- Can we get a convex optimization problem?

First glance :

First glance:

- $|\cdot|$ is convex
- Pointwise minimum of convex func?
- Maximization of a convex func?
- Division by $\|w\|_2$?

On the other hand:

- Pointwise minimum of linear function is concave
- Maximization of a concave function is good!
- Use the linear-fractional trick for the fraction.

Specifically:

- Remove the absolute-value sign using t_i

$$\frac{t_i(w^T x_i + b)}{\|w\|_2}$$

- We therefore want

$$\max_{w,b} \min_i \frac{t_i(w^T x_i + b)}{\|w\|_2}$$

$$\|w\|_2$$

$$= \max_{w,b} \frac{1}{\|w\|_2} \cdot \min_{i=1,\dots,n} t_i(w^T x_i + b)$$

- Would have been fine if without $\frac{1}{\|w\|_2}$

Dealing with $\frac{1}{\|w\|_2}$:

- Let

$$u = \frac{w}{\|w\|_2}, \qquad v = \frac{b}{\|w\|_2}$$

  - The objective becomes

$$\min_{i=1,\dots,n} t_i(u^T x_i + v) \quad \text{— concave}$$

  - Need additional constraints

$$\|u\|_2 = 1$$

    - However, this is not a convex constraint!
    - Replace by $\|u\|_2 \le 1$. (Why does it work?)

$$\Rightarrow \quad \max_{u,v} \min_{i=1,\dots,n} t_i(u^T x_i + v)$$

$$\text{sub to} \quad \|u\|_2 \le 1$$

- The book uses a different form:

- The book uses a different form:

$$\min \quad \|u\|_2$$

$$\text{sub to} \quad \min_{i=1,\cdots,n} \; t_i(u x_i + v) \geq 1$$

$$\Downarrow$$

$$t_i(u x_i + v) \geq 1 \quad \text{for all } i=1,\cdots,n$$

See Boyd P425 for the non-separable case.