Exercise 5.18 [Boyd & Vandenberghe, 2004]

Ans: Without any sophisticated thoughts the problem is equivalent to the feasibility problem

$$\begin{array}{rl} \text{minimize} & 1 \\ \text{subject to} & \inf\{a^T x \mid Ax \preceq b\} > \gamma > \sup\{a^T y \mid Cy \preceq d\} \\ \text{variable} & a, \gamma \end{array}$$

Consider s^* as the optimal value of the below problem

$$\begin{array}{rl} \text{minimize} & a^T x \\ \text{subject to} & Ax \preceq b \\ \text{variable} & x \end{array}$$

The Lagrangian $L(x, \lambda) = a^T x + \lambda^T (Ax - b)$ gives the Lagrange dual function

$$g(\lambda) = \begin{cases} -\lambda^T b, & A^T \lambda + a = 0 \\ -\infty, & \text{else} \end{cases}$$

We should be able to assume the polyhedron $\mathcal{P}_1$ has nonempty interior, that is, Slater's condition for this problem, hence strong duality gives $s^* = -\lambda^T b$ for some $\lambda \succeq 0$ and $A^T \lambda + a = 0$. Similarly if t^* is the optimal value of the below problem

$$\begin{array}{rl} \text{maximize} & a^T y \\ \text{subject to} & Cy \preceq d \\ \text{variable} & y \end{array}$$

The Lagrangian $L(y, \mu) = -a^T y + \mu^T (Cy - d)$ gives the Lagrange dual function

$$g(\mu) = \begin{cases} -\mu^T d, & C^T \mu - a = 0 \\ -\infty, & \text{else} \end{cases}$$

With the same assumption that $\mathcal{P}_2$ has nonempty interior, $t^* = -\mu^T d$ for some $\mu \succeq 0$ and $C^T \mu - a = 0$. The original problem can be viewed as

$$\begin{array}{rl} \text{minimize} & 1 \\ \text{subject to} & -\lambda^T b > \gamma > -\mu^T d \\ & \lambda \succeq 0, \mu \succeq 0 \\ & A^T \lambda + a = 0 \\ & C^T \mu - a = 0 \end{array}$$

which is an LP feasibility problem now. □

Additional Exercise 4.25 [Boyd & Vandenberghe, 2017]

Ans: The first equality follows from

$$D_1 A D_2 u = diag(u) diag(Ax)^{-1} A diag(u)^{-1} diag(x) u = diag(u) diag(Ax)^{-1} Ax = u$$

To derive a equivalent convex problem, first take logarithm to the objective and constraints,

$$\begin{aligned} \text{minimize} \quad & -\sum_{i=1}^{n} \alpha_i \log(Ax)_i \\ \text{subject to} \quad & \sum_{i=1}^{n} \alpha_i \log x_i = 0 \\ & x \succ 0 \end{aligned}$$

The Lagrangian and the optimality conditions are

$$\begin{aligned} L(x,\lambda,\nu) &= -\alpha^T \log(Ax) - \lambda^T x + \nu \alpha^T \log x \\ \nabla_x L(x,\lambda,\nu) &= -A^T diag(Ax)^{-1}\alpha - \lambda + \nu diag(x)^{-1}\alpha = 0 \end{aligned}$$

Here $\log(\cdot)$ denotes the element-wise logarithm on vectors. The complementary slackness on $x \succ 0$ ensures that $\lambda = 0$. The condition above gives

$$-diag(x) A^T diag(Ax)^{-1}\alpha + \nu\alpha = 0$$

If we sum over the entries of this vector equality, we get

$$\begin{aligned} 0 &= -\sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} a_{ji} \frac{\alpha_j}{y_j} \right) + \nu \sum_{i=1}^{n} \alpha_i \\ &= -\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\alpha_j a_{ji} x_i}{y_j} + \nu \sum_{i=1}^{n} \alpha_i \\ &= -\sum_{j=1}^{n} \frac{\alpha_j}{y_j} \sum_{i=1}^{n} a_{ji} x_i + \nu \sum_{i=1}^{n} \alpha_i \\ &= -\sum_{j=1}^{n} \frac{\alpha_j}{y_j} y_j + \nu \sum_{i=1}^{n} \alpha_i \\ &= (-1+\nu) \sum_{i=1}^{n} \alpha_i \end{aligned}$$

and $\nu = 0$. This gives us $A^T diag(y)^{-1}\alpha = diag(x)^{-1}\alpha$, or, $A^T diag(y)^{-1} diag(u) v = diag(x)^{-1} diag(u) v$ since $\alpha_i = u_i v_i$. The equality we are aiming for is

$$(diag(u) diag(y)^{-1} A diag(u)^{-1} diag(x))^T v = v$$

now naturally follows. $\square$

Additional Exercise 4.26 [Boyd & Vandenberghe, 2017]

Ans: (a) Use the fact that $\|\cdot\|_2^* = \|\cdot\|_2, \|\cdot\|_1^* = \|\cdot\|_\infty$ and the Legendre transform of the norms, we derive the Lagrange dual function

$$\begin{aligned}
g(\nu) &= \inf_{x,y} L(x,y,\nu) \\
&= \inf_{x,y} \|y\|_2 + \gamma\|x\|_1 + \nu^T(Ax - b - y) \\
&= -\nu^T b + \left(\inf_y \|y\|_2 - \nu^T y\right) + \left(\inf_x \gamma\|x\|_1 + \nu^T Ax\right) \\
&= -\nu^T b - \left(\sup_y y^T\nu - \|y\|_2\right) - \gamma\left(\sup_x x^T\left(-\frac{1}{\gamma}A^T\nu\right) - \|x\|_1\right) \\
&= \begin{cases} -\nu^T b, & \|\nu\|_2^* \le 1, \left\|-\frac{1}{\gamma}A^T\nu\right\|_1^* \le 1 \\ -\infty, & \text{else} \end{cases} \\
&= \begin{cases} -\nu^T b, & \|\nu\|_2 \le 1, \left\|A^T\nu\right\|_\infty \le \gamma \\ -\infty, & \text{else} \end{cases}
\end{aligned}$$

(b) From the aspects of the equivalent problem, $y^* = Ax^* - b \neq 0$ and $r = y/\|y\|_2$. The KKT condition suggests

$$\nabla_y L(x^*, y^*, \nu^*) = \frac{y^*}{\|y^*\|_2} - \nu = 0$$

Hence we see $\nu^* = r$ here. Since ν^* is feasible for the dual problem, we have $\|A^T r\|_\infty = \|A^T\nu^*\|_\infty \le \gamma$. The other KKT condition requires

$$\begin{aligned}
\nabla_y L(x^*, y^*, \nu^*) &= \gamma sgn(x^*) + A^T\nu^* \\
&= \gamma sgn(x^*) + A^T r = 0
\end{aligned}$$

Hence $r^T Ax^* + \gamma\|x^*\|_1 = (\nabla_x L(x^*, y^*, \nu^*))^T x^* = 0$.

(c) The KKT condition on $\nabla_y L$ from part (b) asserts

$$a_i^T r = -\gamma sgn(x_i^*)$$

where a_i denotes the i-th column of A. Suppose $\|a_i\|_2 < \gamma$, then by Cauchy's inequality $a_i^T r < \gamma$ (since r is defined to be an unit vector) and the above equality cannot hold unless $sgn(x_i^*) = 0$, *i.e.* $x_i^* = 0$. $\square$

Additional Exercise 6.5 [Boyd & Vandenberghe, 2017]

Ans: (a) The maximum likelihood estimate of x is

$$x = \arg\max \prod_{i=1}^{m} p(\phi^{-1}(y_i) - a_i^T x)$$

where $p(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}}$ is the p.d.f. of the IID's $v_i \sim \mathcal{N}(0, \sigma^2)$. To find a convex representation of this problem we can take logarithm to the objective function and get $f_0(x) = \sum_{i=1}^{m} \left(-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(\phi^{-1}(y_i) - a_i^T x)^2}{2\sigma^2} \right)$, or we can drop the constant $-\frac{m}{2}\log(2\pi\sigma^2)$ and let

$$\begin{aligned} f_0(x) &= -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (\phi^{-1}(y_i) - a_i^T x)^2 \\ &= -\frac{1}{2\sigma^2} \|z - A^T x\|_2^2 \end{aligned}$$

where $A = [a_1, \cdots, a_m]$ is the collection of the coefficient vectors. This is a least square minimization problem except that z is another variable with some inequality constraint inscribed due to $\alpha \le \phi'(u) \le \beta$. In particular, from mean value theorem, $\frac{y_i - y_j}{z_i - z_j} = \phi'(\xi_{ij})$ for any $i, j = 1, \cdots, m$, we get the induced inequalities

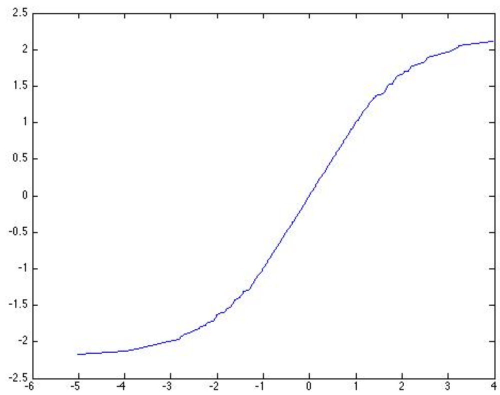
$$\frac{y_i - y_j}{\beta} \le z_i - z_j \le \frac{y_i - y_j}{\alpha}$$

for $y_i > y_j$. Now assume y is sorted that $y_{k+1} \ge y_k$, the problem to be solved can be formulated as below

$$\begin{array}{ll} \text{minimize} & \|A^T x - z\|_2^2 \\ \text{subject to} & \dfrac{y_{k+1} - y_k}{\beta} \le z_{k+1} - z_k \le \dfrac{y_{k+1} - y_k}{\alpha} \text{ for } k = 1, \cdots, m-1 \end{array}$$

Note that if it happens $y_{k+1} = y_k$, then $z_{k+1} = z_k$ will be enforced!

(b) The following plot is the optimal solution z versus y.



Attached is the code to inscribe the problem and call CVX.

```
D = zeros(299,300);
for k=1:299
    D(k,k) = -1;
    D(k,k+1) = 1;
end

cvx_begin
variables x(4) z(300)
minimize norm(A*x - z)
subject to
    (1/beta)*D*y <= D*z
    D*z <= (1/alpha)*D*y
cvx_end
```

Here's some numerics from the result.

```
Calling SDPT3 4.0: 899 variables, 305 equality constraints
   For improved efficiency, SDPT3 is solving the dual problem.
------------------------------------------------------------


 num. of constraints = 305
 dim. of socp   var  = 301,   num. of socp blk  =  1
 dim. of linear var  = 598
 20 linear variables from unrestricted variable.
 *** convert ublk to lblk
*******************************************************************
   SDPT3: Infeasible path-following algorithms
*******************************************************************
 version  predcorr  gam  expon  scale_data
    NT      1      0.000   1        0
it pstep dstep pinfeas dinfeas  gap      prim-obj      dual-obj    cputime
-------------------------------------------------------------------
 0|0.000|0.000|8.2e+00|5.9e+01|7.1e+05| 5.037923e+03  0.000000e+00| 0:0:00| chol  1
 1|0.885|0.977|9.4e-01|1.6e+00|2.4e+04| 4.812382e+03 -1.693377e+01| 0:0:01| chol  1
 2|0.684|0.627|3.0e-01|6.6e-01|1.1e+04| 3.665563e+03 -2.905354e+01| 0:0:01| chol  1
 3|0.923|0.369|2.3e-02|4.3e-01|5.3e+03| 1.970508e+03 -3.497857e+01| 0:0:01| chol  1
 4|0.874|0.338|2.9e-03|2.9e-01|3.0e+03| 1.229351e+03 -3.654966e+01| 0:0:01| chol  1
 5|0.973|0.730|7.7e-05|7.9e-02|6.8e+02| 3.949011e+02 -3.017530e+01| 0:0:01| chol  1
 6|0.604|0.204|3.1e-05|6.3e-02|4.8e+02| 3.024580e+02 -2.742389e+01| 0:0:01| chol  1
 7|1.000|0.218|7.8e-07|4.9e-02|2.6e+02| 1.689486e+02 -2.408657e+01| 0:0:01| chol  1
 8|0.730|0.535|3.8e-07|2.3e-02|1.3e+02| 9.771440e+01 -1.625992e+01| 0:0:01| chol  1
 9|1.000|0.261|6.7e-08|1.7e-02|6.7e+01| 4.575721e+01 -1.366559e+01| 0:0:01| chol  1
```

```
10|0.823|0.455|2.7e-08|9.2e-03|3.6e+01| 2.368631e+01 -1.011086e+01| 0:0:01| chol  1
11|0.967|0.351|6.3e-09|6.0e-03|1.8e+01| 9.156954e+00 -7.802653e+00| 0:0:01| chol  1
12|1.000|0.384|2.5e-09|3.7e-03|1.0e+01| 4.238214e+00 -5.807395e+00| 0:0:01| chol  1
13|0.736|0.354|1.2e-09|2.4e-03|7.1e+00| 2.320164e+00 -4.583437e+00| 0:0:01| chol  1
14|0.563|0.288|6.9e-10|1.7e-03|5.5e+00| 1.608416e+00 -3.780771e+00| 0:0:01| chol  1
15|0.727|0.281|2.2e-10|1.2e-03|3.6e+00| 2.221933e-01 -3.263875e+00| 0:0:01| chol  1
16|0.900|0.283|5.2e-10|8.7e-04|2.4e+00|-5.236924e-01 -2.851653e+00| 0:0:01| chol  1
17|0.908|0.329|3.1e-10|5.9e-04|1.6e+00|-9.046204e-01 -2.467141e+00| 0:0:01| chol  1
18|0.742|0.475|1.2e-10|3.1e-04|8.7e-01|-1.177404e+00 -2.033419e+00| 0:0:01| chol  1
19|0.807|0.238|6.2e-11|2.4e-04|6.4e-01|-1.289185e+00 -1.913648e+00| 0:0:01| chol  1
20|0.875|0.389|1.9e-10|1.4e-04|3.8e-01|-1.377674e+00 -1.752529e+00| 0:0:01| chol  1
21|1.000|0.325|5.8e-11|9.7e-05|2.5e-01|-1.421602e+00 -1.665739e+00| 0:0:01| chol  1
22|0.640|0.547|5.2e-11|4.4e-05|1.2e-01|-1.440453e+00 -1.558271e+00| 0:0:01| chol  1
23|0.789|0.149|2.5e-11|3.7e-05|1.0e-01|-1.447220e+00 -1.545533e+00| 0:0:01| chol  1
24|1.000|0.291|3.7e-11|2.7e-05|7.2e-02|-1.452773e+00 -1.523698e+00| 0:0:01| chol  1
25|0.926|0.436|9.0e-11|2.3e-05|4.2e-02|-1.458300e+00 -1.498800e+00| 0:0:01| chol  1
26|1.000|0.901|4.0e-13|1.0e-05|7.8e-03|-1.460940e+00 -1.468540e+00| 0:0:01| chol  1
27|1.000|0.774|1.0e-13|2.0e-06|2.4e-03|-1.462750e+00 -1.465152e+00| 0:0:01| chol  1
28|0.908|0.936|1.6e-13|6.0e-07|7.3e-04|-1.463201e+00 -1.463925e+00| 0:0:01| chol  1
29|1.000|0.972|1.6e-13|1.8e-07|7.2e-05|-1.463545e+00 -1.463615e+00| 0:0:01| chol  1
30|1.000|0.965|3.7e-13|1.8e-08|7.6e-06|-1.463575e+00 -1.463583e+00| 0:0:02| chol  1
31|1.000|0.974|1.5e-12|1.9e-09|7.9e-07|-1.463579e+00 -1.463579e+00| 0:0:02| chol  1
32|0.998|0.988|9.6e-13|1.9e-10|1.5e-08|-1.463579e+00 -1.463579e+00| 0:0:02|
  stop: max(relative gap, infeasibilities) < 1.49e-08
-------------------------------------------------------------------
 number of iterations   = 32
 primal objective value = -1.46357893e+00
 dual   objective value = -1.46357895e+00
 gap := trace(XZ)       = 1.54e-08
 relative gap           = 3.92e-09
 actual relative gap    = 3.34e-09
 rel. primal infeas (scaled problem)   = 9.61e-13
 rel. dual     "        "       "      = 1.95e-10
 rel. primal infeas (unscaled problem) = 0.00e+00
 rel. dual     "        "       "      = 0.00e+00
 norm(X), norm(y), norm(Z) = 9.4e+00, 2.4e+01, 8.6e+00
 norm(A), norm(b), norm(C) = 5.2e+01, 2.0e+00, 1.0e+01
 Total CPU time (secs)  = 1.58
 CPU time per iteration = 0.05
 termination code       =  0
 DIMACS: 9.6e-13  0.0e+00  5.5e-10  0.0e+00  3.3e-09  3.9e-09
-------------------------------------------------------------------
```

```
------------------------------------------------------------
Status: Solved
Optimal value (cvx_optval): +1.46358

>> max((D*y) ./ (D*z))

ans =

    1.0000

>> min((D*y) ./ (D*z))

ans =

     0

>> x

x =

    0.4819
   -0.4657
    0.9364
    0.9297
```

Additional Exercise 7.1 [Boyd & Vandenberghe, 2017]

Ans: (a) Suppose $a_i^T Q a_i \leq 1$ for $i = 1, \cdots, p$. For $x \in \mathbb{R}^n$ with $x^T Q^{-1} x \leq 1$, or, say, $\|Q^{-1/2}x\|_2^2 \leq 1$, Cauchy's inequality gives

$$|a_i^T x| = |(Q^{1/2}a_i)^T (Q^{-1/2}x)| \leq \|Q^{1/2}a_i\|_2 = \sqrt{a_i^T Q a_i} \leq 1$$

for $i = 1, \cdots, p$. On the other hand suppose $|a_i^T x| \leq 1$ for $i = 1, \cdots, p$ whenever $x^T Q^{-1} x \leq 1$. Introduce a new variable $y = Q^{-1/2}x$, then the assumed condition states

$$\|y\|_2^2 \leq 1 \Rightarrow |(Q^{1/2}a_i)^T y| \leq 1 \text{ for } i = 1, \cdots, p$$

If $a_j^T Q a_j > 1$ for some j, then take $y_j = Q^{1/2}a_j / \|Q^{1/2}a_j\|_2$, we can violate the above predicate since $\|y_j\|_2^2 = 1$ yet

$$|(Q^{1/2}a_j)^T y_j| = \|Q^{1/2}a_j\| = \sqrt{a_j^T Q a_j} > 1$$

(b) Denote $A = [a_1, \cdots, a_p] \in \mathbb{R}^{n \times p}$ be the collection of coefficients of the constraints. We first derive the Lagrange dual function from the Lagrangian and the Legendre transform of log-determinant,

$$\begin{aligned}
L(Q, \lambda) &= \log\det Q^{-1} + \sum_{i=1}^{p} \lambda_i (a_i^T Q a_i - 1) \\
&= \log\det Q^{-1} + diag(\lambda) : (A^T Q A) - \mathbf{1}^T \lambda \\
&= -\mathbf{1}^T \lambda - \left((-A diag(\lambda) A^T) : Q - \log\det Q^{-1} \right) \\
g(\lambda) &= \inf_Q L(Q, \lambda) \\
&= -\mathbf{1}^T \lambda - \log\det(-A diag(\lambda) A^T)^{-1} + n \\
&= n - \mathbf{1}^T \lambda - \log\det \left(-\sum_{i=1}^{p} \lambda_i a_i a_i^T \right)^{-1}
\end{aligned}$$

under the (domain) assumption $-A diag(\lambda) A^T = -\sum_{i=1}^{p} \lambda_i a_i a_i^T \prec 0$. The dual problem can be formulated now,

$$\begin{array}{ll}
\text{maximize} & n - \mathbf{1}^T \lambda - \log\det \left(-A diag(\lambda) A^T \right)^{-1} \\
\text{subject to} & A diag(\lambda) A^T \succ 0 \\
& \lambda \succeq 0
\end{array}$$

Note here that (c) The KKT condition suggests

$$\begin{aligned}
\nabla_Q L(Q, \lambda) = 0 \\
&= -Q^{-1} + A diag(\lambda) A^T
\end{aligned}$$

Hence $Q = (Adiag(\lambda)A^T)^{-1}$. Now for $x \in C$, *i.e.* $|a_i^T x| \le 1$ for $i = 1, \cdots, p$,

$$x^T Q^{-1} x = x^T Adiag(\lambda)A^T x = \sum_{i=1}^{p} \lambda_i z_i^2$$

where $z = A^T x \preceq \mathbf{1}$ since $|z_i| = |a_i^T x| \le 1$. On the other hand, due to the strong duality, $d^* = p^*$,

$$n - \mathbf{1}^T \lambda - \log\det Q = \log\det Q^{-1}$$

$$n - \mathbf{1}^T \lambda = \log\det Q + \log\det Q^{-1} = \log(\det Q \det Q^{-1}) = \log\det I = 0$$

therefore,

$$x^T Q^{-1} x = \sum_{i=1}^{p} \lambda_i z_i^2 \le \sum_{i=1}^{p} \lambda_i = \mathbf{1}^T \lambda = n$$

$\square$