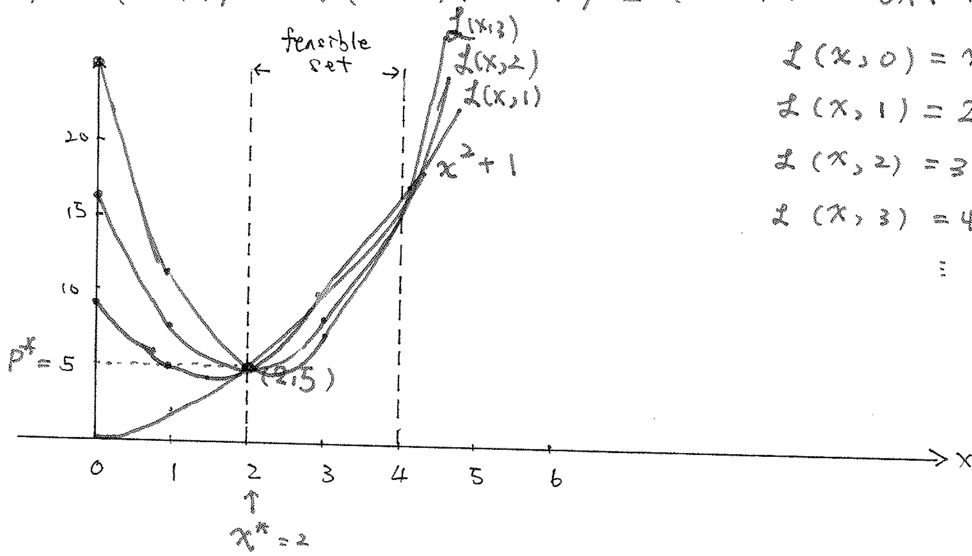
5.1

(a) minimize x^2+1 subject to $(x-2)(x-4)\le 0$. Since $(x-2)(x-4)\le 0$, we have $2\le x\le 4$ which is equal to the feasible set. And $2\le x\le 4 \Leftrightarrow 4\le x^2\le 16 \Leftrightarrow 5\le x^2+1\le 17$, which means that the optimal value $p^*=5$ while the optimal point $x^*=2$.

(b) $\mathcal{L}(x,\lambda) = (x^2+1)+\lambda(x-2)(x-4) = (1+\lambda)x^2 - 6\lambda x + (1+8\lambda)$.



$\mathcal{L}(x,0) = x^2+1$

$\mathcal{L}(x,1) = 2x^2-6x+9$, $\min \mathcal{L}(x,1) = \frac{9}{2}$

$\mathcal{L}(x,2) = 3x^2-12x+17$, $\min \mathcal{L}(x,2) = 5$

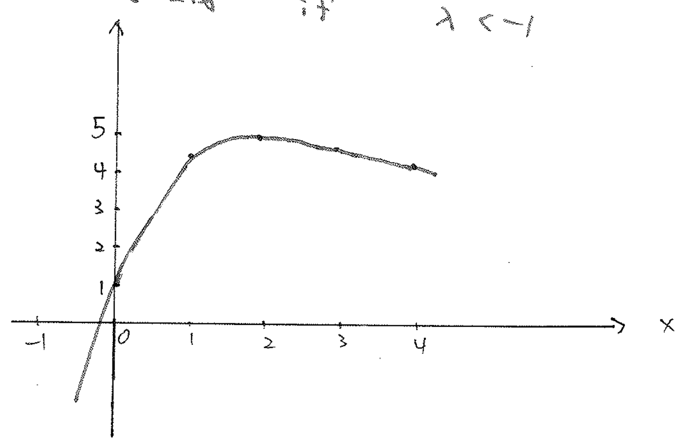
$\mathcal{L}(x,3) = 4x^2-18x+25$, $\min \mathcal{L}(x,3) = \frac{19}{4}$

$\vdots$

We can see that the minimum value of $\mathcal{L}(x,\lambda)$ over x is always less than p^*. We observe that $\min \mathcal{L}(x,\lambda)$ increases as λ from 0 to 2, reaches the maximum value at $\lambda=2$, and $\min \mathcal{L}(x,\lambda)$ decreases when $\lambda>2$. Hence, we have $p^* = \mathcal{L}(x,\lambda)$ when $\lambda=2$.

And we have $g(\lambda) = \inf_{x\in D}\left((1+\lambda)x^2 - 6\lambda x + (1+8\lambda)\right)$. When $\lambda<-1$, $g(\lambda)$ can go to $-\infty$. When $\lambda\ge -1$, we have $\nabla_x \mathcal{L}(x,\lambda) = 2(1+\lambda)x - 6\lambda = 0 \Leftrightarrow x = \frac{3\lambda}{1+\lambda}$. Thus, we get

$$g(\lambda) = \begin{cases} \frac{-9\lambda^2}{1+\lambda}+1+8\lambda & \text{if } \lambda\ge -1 \\ -\infty & \text{if } \lambda<-1 \end{cases}$$



We can see that the dual function g is concave, and its value is equal to $p^*=5$ when $\lambda=2$, and less for other λ.

(c) The Lagrange dual problem is: maximize $\frac{-9\lambda^2}{1+\lambda} + 1 + 8\lambda$ subject to $\lambda \geq 0$.

Let the objective function as $h(\lambda) = \frac{-9\lambda^2}{1+\lambda} + 1 + 8\lambda$. $\frac{dh}{d\lambda} = \frac{-\lambda^2 - 2\lambda + 8}{(1+\lambda)^2}$, $\frac{d^2h}{d\lambda^2} = \frac{-18\lambda - 18}{(1+\lambda)^4}$

Set $\frac{dh}{d\lambda} = 0$ we have $\lambda = 2$ and since $\frac{d^2h}{d\lambda^2} \leq 0$ when $\lambda \geq 0$, which implies the dual optimum occurs at $\lambda = 2$ with $d^* = 5$. Since Slater's constraint qualification is satisfied, strong duality holds for this example with $p^* = d^* = 5$.

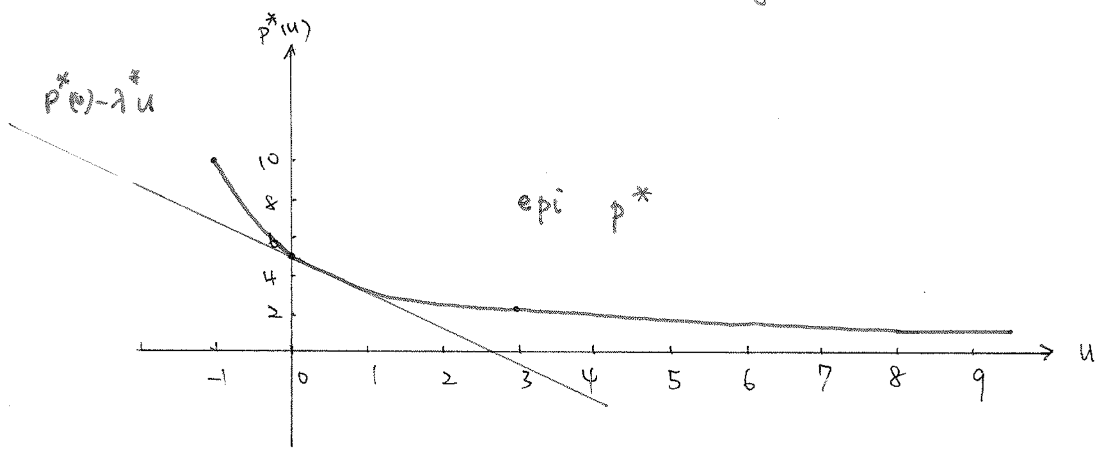
(d) minimize x^2+1 subject to $(x-2)(x-4) \leq u$. Since $\inf_x \left((x-2)(x-4) \right) = -1$, this constraint is infeasible when $u < -1$. When $u \geq -1$, we have $(x-2)(x-4) \leq u \Leftrightarrow$

$$x^2 - 6x + 8 \leq u \Leftrightarrow x^2 - 6x + 9 \leq 1+u \Leftrightarrow (x-3)^2 \leq 1+u \Leftrightarrow 3-\sqrt{1+u} \leq x \leq 3+\sqrt{1+u}.$$

When $-1 \leq u \leq 8$, the optimal point is $x^*(u) = 3 - \sqrt{1+u}$. When $u \geq 8$, the optimum is the unconstrained minimum of x^2+1, which means $x^*(u) = 0$.

Hence, we have:

$$p^*(u) = \begin{cases} \infty & \text{if } u < -1 \\ 11 + u - 6\sqrt{1+u} & \text{if } -1 \leq u \leq 8 \\ 1 & \text{if } u \geq 8 \end{cases}$$



$$\frac{dp^*(0)}{du} = 1 - \frac{6}{2\sqrt{1+u}}\Bigg|_{u=0} = 1 - 3 = -2 = -\lambda^*$$

5.3:

For $\lambda = 0$, $g(\lambda) = \inf_{x \in D} c^T x = -\infty$.

For $\lambda > 0$, $g(\lambda) = \inf_{x \in D} \left(c^T x + \lambda f(x) \right) = \lambda \inf_{x \in D} \left(f(x) - \left(\frac{c}{-\lambda}\right)^T x \right) = -\lambda f^*\left(\frac{-c}{\lambda}\right)$

When $\lambda > 0$, $-g(\lambda)$ is the perspective function of f^*. Since f^* is convex even if f is not convex, we know $-g(\lambda)$ is convex which implies $g(\lambda)$ is concave.

The dual problem is maximize $-\lambda f^*\left(\frac{-c}{\lambda}\right)$ subject to $\lambda \geq 0$.

5.12

Introduce new variables y_i and equality constraints $y_i = b_i - a_i^T x$. Then we can rewrite the problem as: minimize $-\sum_{i=1}^m \log y_i$ subject to $y = b - Ax$, where $A \in \mathbb{R}^{m \times n}$ has a_i^T as its ith row.

The Lagrangian is: $\mathcal{L}(x, y, \nu) = -\sum_{i=1}^m \log y_i + \nu^T(y - b + Ax)$.

The dual function is: $g(\nu) = \inf_{x, y \in D} \left(-\sum_{i=1}^m \log y_i + \nu^T(y - b + Ax) \right)$

$\nu^T A x$ can go to $-\infty$ unless $\nu^T A = 0$. The terms in y are unbounded below if $\nu \ngtr 0$, and achieve there minimum when $y_i = \frac{1}{\nu_i}$.

Then we get the dual function as: $g(\nu) = \begin{cases} \sum_{i=1}^m \log \nu_i + m - b^T \nu & \text{if } A^T \nu = 0, \nu > 0 \\ -\infty & \text{otherwise} \end{cases}$

And the dual problem is

$$\text{maximize} \quad \sum_{i=1}^m \log \nu_i + m - b^T \nu$$

$$\text{subject to} \quad A^T \nu = 0$$

A3.3

(a) norm([x + 2*y, x - y]) is correctly identify as convex, but the equality constraints are only valid when both left and right hand sides are affine. Since the norm of a vector is zero if and only if the vector is zero. Hence, we can rewrite the constraints as $x + 2*y == 0$ and $x - y == 0$, which are equal to $x == 0$ and $y == 0$.

(b) The inner square is acceptable since square is convex and its expr = $x+y$ is an affine func. However, the outer square is not acceptable because square does not acept a convex agument. We can use square_pos instead since it accepts convex argument, then we have:

$$\text{square\_pos}(\text{square}(x+y)) \leq x-y$$

Or we can also denote the constraint as: $(x+y)^4 \leq x-y$

(c) $\frac{1}{x}$ and $\frac{1}{y}$ are convex if $x, y \in \mathbb{R}_{++}$. Thus, we have to rewrite the constraints as

inv_pos(x) + inv_pos(y) <= 1. The inv_pos function has domain on $\mathbb{R}_{++}$, at the same time include $x > 0$ and $y > 0$.

(d) norm() is convex if norm() is increasing / decreasing in its arguments and its exprs are convex/concave. We know that the exprs max(x,1) and max(y,2) are convex, but norm() are not always increasing. We have to make the exprs as affine functions to satisfy the DCP rules in norm(). Let's set $\max(x,1) \leq u$, $\max(y,2) \leq v$, then we have:

norm([u, v]) <= 3*x + y

max(x, 1) <= u

max(y, 2) <= v.

This works since norm() is convex and monotonic over $(u,v) \in \mathbb{R}^2_+$, $u \in [1,\infty)$ and $v \in [2,\infty)$.

(e) Greater - than >= constraints, where the left-hand expression is concave. However, xy is not concave on $x >= 0$ and $y >= 0$. We can rewrite the constraints as: $x >=$ inv_pos(y).

(f) The original constraint expression fails since we divide a convex function by a concave function. We can rewrite the constraint as = quad_over_lin $(x+y$, sqrt$(y)) <= x-y+5$.

Since quad_over_lin is monotone decreasing in the second argument, so we can put a concave function in its second argument, and sqrt is concave.

(g) The function x^3+y^3 is convex for $x \geq 0$ and $y \geq 0$, but x^3 or y^3 is not convex when $x < 0$ or $y < 0$. Hence, we can rewrite the constraint as:

pow_pos $(x, 3)$ + pow_pos $(y, 3)$ <= 1, $x >= 0$, $y >= 0$.

(h) Since xy in the sqrt is not convex, which makes CVX to reject the statement. We know that $\sqrt{xy - z^2} = \sqrt{x(y - z^2/x)}$. And we know that geo_mean() is concave if it is increasing in argument i and expr$_i$ is concave. We see that geo_mean increases according to x and $y - z^2/x$, and x, $y - \frac{z^2}{x}$ are concave. Hence, we can rewrite the constraint as:

x + z <= 1 + geo_mean([x, y- quad_over_lin(z,x)])

x >= 0

y >= 0.

A3.10

(a) Since $\mathbb{E}c = c_0$, we have $\mathbb{E}c^Tx = c_0^Tx$. Hence, the problem becomes as a LP as:

minimize c_0^Tx, subject to $Ax \preceq b$.

(b) We have
$$\mathrm{var}(c^Tx) = \mathbb{E}\left(c^Tx - \mathbb{E}c^Tx\right)^2 = \mathbb{E}\left(c^Tx - c_0^Tx\right)^2$$
$$= \mathbb{E}\left((c-c_0)^Tx\right)^2 = \mathbb{E}\ x^T(c-c_0)(c-c_0)^Tx$$
$$= x^T\,\mathbb{E}(c-c_0)(c-c_0)^Tx = x^T\Sigma x$$

Then the risk-sensitive cost can be solved by minimizing the follow:

$$\text{minimize} \quad c_0^Tx + \gamma x^T\Sigma x$$
$$\text{subjecto to} \quad Ax \preceq b$$

which is a convex quadratic problem in x since $\gamma \geq 0$ and $\Sigma \succeq 0$.

(c) If $\gamma < 0$, then the objective function $c_0^Tx + \gamma x^T\Sigma x$ becomes concave.

Thus, it is not a convex optimization problem.

(d) We know that c^Tx is a random variable, normally distribution with mean c_0^Tx and variance $x^T\Sigma x$. Thus, we have $\mathrm{prob}(c^Tx \geq \beta) = \Phi\left(\frac{\beta - c_0^Tx}{\|\Sigma^{\frac{1}{2}}x\|}\right)$ where $\Phi(t) = \frac{1}{\sqrt{2\pi}}\int_x^\infty e^{\frac{-u^2}{2}}du$.

Φ is monotone decreasing, so $\mathrm{prob}(c^Tx \geq \beta) \leq \alpha \Leftrightarrow \dfrac{\beta - c_0^Tx}{\|\Sigma^{\frac{1}{2}}x\|} \geq \Phi^{-1}(\alpha)$

$$\Leftrightarrow \Phi^{-1}(\alpha)\|\Sigma^{\frac{1}{2}}x\| + c_0^Tx \leq \beta.$$

If $\alpha \le 0.5$, we have $\Phi^{-1}(\alpha) \ge 0$, so this is a convex constraint over x.

The original problem can be rewritten as:

$$\text{minimize } \beta \ , \text{ subject to } \Phi^{-1}(\alpha)\|\Sigma^{\frac{1}{2}}x\| + C_0^T x \le \beta, \quad Ax \preceq b.$$

The problem is a second order cone programming since the objective function is linear, $\Phi^{-1}(\alpha)\|\Sigma^{\frac{1}{2}}x\| + C_0^T x \le \beta$ is a second order cone constraint, and $Ax \preceq b$ is a linear inequality constraint.

If the $\alpha > 0.5$, we then get a risk-seeking problem. If $\alpha > 0.5$, then $\Phi^{-1}(\alpha) < 0$ so $\beta < \mathbb{E}\, c^T x = C_0^T x$. There are two way to decrease $\text{prob}(c^T x \ge \beta)$ if $\beta < \mathbb{E}\, c^T x$:

One is that we can decrease the expected value such that the PDF shifts to left.
The other is that we can increase the variance, which is a risk-seeking choice.

A3.32:

(a) We have to prove the hint first. If $u > 0$, then $1(u>0) = 1$ and $1+\lambda u > 1$, so $(1+\lambda u)_+ > 1$. If $u \le 0$, then $1(u>0) = 0$ and $(1+\lambda u)_+ \ge 0$. Hence, we have $(1+\lambda u)_+ \ge 1(u>0)$ for all $u \in \mathbb{R}$. Let $u = f_i(x)$, we have $(1+\lambda f_i(x))_+ \ge 1(f_i(x) > 0)$ for all i. If we sum up both sides of the inequality from $i=1$ to m, we get

$$\sum_{i=1}^{m} 1(f_i(x) > 0) \le \sum_{i=1}^{m} (1+\lambda f_i(x))_+ \quad \text{---} ①$$

The constraint $\sum_{i=1}^{m}(1+\lambda f_i(x))_+ \le m-k$ in the original problem combines with ①, we get

$\sum_{i=1}^{m} 1(f_i(x) > 0) \le m - k$. Thus, $f_i(x) > 0$ for at most $m-k$ value of i,

which means that $f_i(x) \le 0$ for at least k values of i.

(b)

If $\lambda > 0$ then $(\lambda u)_+ = \lambda (u)_+$ for all $u \in \mathbb{R}$, which implies that $(1+\lambda f_i(x))_+ = \lambda\left(\frac{1}{\lambda} + f_i(x)\right)_+$.

Then the constraint can be rewritten as:

$$\sum_{i=1}^{m} (1+\lambda f_i(x))_+ \le m-k \iff \sum_{i=1}^{m} \lambda\left(\frac{1}{\lambda}+f_i(x)\right)_+ \le m-k$$

$$\iff \sum_{i=1}^{m} \left(\frac{1}{\lambda}+f_i(x)\right)_+ \le \frac{1}{\lambda}(m-k)$$

Set $\frac{1}{\lambda} = \mu$, then the problem can be expressed as:

$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && \sum_{i=1}^{m} (\mu + f_i(x))_+ \le \mu(m-k) \\ &&& \mu > 0 \end{aligned}$$

The function $(\cdot)_+ = \max\{\cdot, 0\}$ is convex and nondecreasing, and $\mu + f_i(x)$ is convex in μ and x, so $(\mu + f_i(x))_+$ is convex and it is a convex optimization problem. After solving the problem and getting the optimal value μ^*, the optimal value λ is $\lambda^* = \frac{1}{\mu^*}$.

(c) The optimal value of λ is $\lambda^* = 282.98$. The objective value is -8.45. The constraints satisfied is 66. If we take the tentative solution, choose the k constraints with the smallest values of $f_i(x)$ and minimize $f_0(x)$ subject to these k constraints, we get the objective value $= -8.80$ (by using CVXOPT solver).

python implementation:

```
import cvxpy as cvx
from satisfy_some_constraints_data import *
x = cvx.Variable(n)
mu = cvx.Variable()
constraints = [cvx.sum(cvx.pos(mu + A*x - b))<=(m-k)*mu,\
               mu >=0]
problem = cvx.Problem(cvx.Minimize(c.T*x), constraints)
problem.Solve(Solver= cvx.CVXOPT)
least_violated = np.argsort(A.dot(x.value)-b)[:k]
constraints = [A[least_violated]*x<=b[least_violated]]
problem = cvx.Problem(cvx.Minimize(c.T*x), constraints)
problem.solve(solver = cvx.CVXOPT)
```

A4.3:

The Lagrangian is:

$$\mathcal{L}(x, \lambda_1, \lambda_2) = \sum_{k=1}^{n} x_k \log \frac{x_k}{y_k} + \lambda_1^T (Ax - b) + \lambda_2^T (\mathbb{1}^T x - 1)$$

$$= \sum_{k=1}^{n} x_k \log \frac{x_k}{y_k} + \sum_{k=1}^{n} a_k^T \lambda_1 x_k - b^T \lambda_1 + \lambda_2 \sum_{k=1}^{n} x_k - \lambda_2 \quad \text{— ①}$$

Minimizing over x_k gives us the conditions:

$$\nabla \mathcal{L}_{x_k} = \log \frac{x_k}{y_k} + 1 + a_k^T \lambda_1 + \lambda_2 = 0 \Leftrightarrow x_k = y_k e^{-(a_k^T \lambda_1 + \lambda_2 + 1)} \quad \text{— ②}$$

Putting ② into ① gives us the Lagrange dual function:

$$g(\lambda_1, \lambda_2) = \sum_{k=1}^{n} y_k e^{-(a_k^T \lambda_1 + \lambda_2 + 1)} \log e^{-(a_k^T \lambda_1 + \lambda_2 + 1)} + \sum_{k=1}^{n} a_k^T \lambda_1 y_k e^{-(a_k^T \lambda_1 + \lambda_2 + 1)}$$

$$- b^T \lambda_1 + \lambda_2 \sum_{k=1}^{n} y_k e^{-(a_k^T \lambda_1 + \lambda_2 + 1)} - \lambda_2$$

$$= -b^T \lambda_1 - \lambda_2 - \sum_{k=1}^{n} y_k e^{-(a_k^T \lambda_1 + \lambda_2 + 1)} \quad \text{— ③}$$

The dual problem is: maximize $-b^T \lambda_1 - \lambda_2 - \sum_{k=1}^{n} y_k e^{-(a_k^T \lambda_1 + \lambda_2 + 1)}$

The dual problem can be simplified if we optimize over λ_2 by setting $\nabla_{\lambda_2} g = 0$, then we have:

$$-1 + \sum_{k=1}^{n} y_k e^{-(a_k^T \lambda_1 + \lambda_2 + 1)} = 0 \Leftrightarrow \lambda_2 = -1 + \log \sum_{k=1}^{n} y_k e^{-a_k^T \lambda_1} \quad \text{— ④}$$

Substituting ④ into ③ gives us:

$$g(\lambda_1) = -b^T \lambda_1 + 1 - \log \sum_{k=1}^{n} y_k e^{-a_k^T \lambda} - \sum_{k=1}^{n} y_k e^{-(a_k^T \lambda_1 - 1 + \log \sum_{k=1}^{n} y_k e^{-a_k^T \lambda_1} + 1)}$$

$$= -b^T \lambda_1 - \log \sum_{k=1}^{n} y_k e^{-a_k^T \lambda_1}$$

If we set $-\lambda_1 = z$, then we get

$$g(\lambda_1) = b^T z - \log \sum_{k=1}^{m} y_k e^{a_k^T z}$$

Hence, the Lagrange dual problem is:

$$\text{maximize} \quad b^T z - \log \sum_{k=1}^{m} y_k e^{a_k^T z}.$$