1. (a). The Lagrangian is $L(x, y, v) = \max_{i=1,\dots,m} y_i + \sum_{i=1}^{m} v_i (a_i^T x + b_i - y_i)$

The dual function is thus $g(v) = \inf_{\substack{x \in R^n \\ y \in R^m}} (\max_{i=1,\dots,m} y_i - v^T y + v^T A x + v^T b)$, where a_i^T is i^{th} row of A

$$= b^T v + \inf_x (v^T A x) + \inf_y (\max_{i=1,\dots,m} y_i - v^T y)$$

$\inf_x (v^T A x)$ is bounded if $A^T v = 0$, and $\inf_y (\max_{i=1,\dots,m} y_i - v^T y)$ is bounded if $\mathbb{1}^T v = 1$, $v \geq 0$

Thus, $g(v) = \begin{cases} b^T v, & \text{if } A^T v = 0, \mathbb{1}^T v = 1, v \geq 0 \\ -\infty, & \text{otherwise} \end{cases}$

Therefore, the dual problem is

$$\begin{aligned} \max \quad & b^T v \\ \text{s.t.} \quad & A^T v = 0 \\ & \mathbb{1}^T v = 1 \\ & v \geq 0 \end{aligned}$$

(b). Let $t \geq \max_{i=1,\dots,m}(a_i^T x + b_i)$. The primal problem can thus be formulated as a LP:

$$\begin{aligned} \min \quad & t \\ \text{s.t.} \quad & Ax + b - t\mathbb{1} \leq 0 \end{aligned}$$

The dual is

$$\begin{aligned} \max \quad & b^T z \\ \text{s.t.} \quad & A^T z = 0 \\ & \mathbb{1}^T z = 1 \\ & z \geq 0 \end{aligned}$$

which is exactly the same as the result in (a).

(c). piece-wise linear:

Primal

$$\min \max_{i=1,\dots,m} (a_i^T x + b_i)$$

Dual

$$\begin{aligned} \max\ & b^T \nu \\ \text{s.t. } & A^T \nu = 0 \\ & \mathbb{1}^T \nu = 1 \\ & \nu \succeq 0 \end{aligned}$$

Geometric programming approximation:

$$\min \log\left(\sum_{i=1}^m e^{a_i^T x + b_i}\right)$$

$$\begin{aligned} \max\ & b^T \nu + \left(-\sum_{i=1}^m \nu_i \log \nu_i\right) \\ \text{s.t. } & A^T \nu = 0 \\ & \mathbb{1}^T \nu = 1 \\ & \nu \succeq 0 \end{aligned}$$

We can see that both problems are convex optimization and Slater's condition is met, so, we have strong duality $p^*_{pwl} = d^*_{pwl}$, $p^*_{gp} = d^*_{gp}$. In addition, we notice that the constraints of the dual problem are identical in both problems and their objective function is very similar.

Consider the function $-\sum_{i=1}^m \nu_i \log \nu_i$ with $\begin{cases} \mathbb{1}^T \nu = 1 \\ \nu \succeq 0 \end{cases}$. It's easy to see that it is bounded by $0 \le -\sum_{i=1}^m \nu_i \log \nu_i \le \log m$.

Thus, $0 \le d^*_{gp} - d^*_{pwl} \le \log m$ and thus $0 \le p^*_{gp} - p^*_{pwl} \le \log m$

(d). Similarly, let's write down the dual problem of this new primal problem:

$$\begin{aligned} \max\ & b^T \nu - \frac{1}{\gamma} \sum_{i=1}^m \nu_i \log \nu_i \\ \text{s.t. } & A^T \nu = 0 \\ & \mathbb{1}^T \nu = 1 \\ & \nu \succeq 0 \end{aligned}$$

Then similarly we know $0 \le -\frac{1}{\gamma}\sum_{i=1}^m \nu_i \log \nu_i \le \frac{1}{\gamma}\log m$

Thus, similarly, $0 \le p^*_{new} - p^*_{pwl} \le \frac{1}{\gamma}\log m$, which means $p^*_{new} \to p^*_{pwl}$ as $\gamma \to \infty$

2. (a). First of all, we claim $\sum_{k=1}^{m} a_k a_k^T \succ 0$ because $a_1, \cdots, a_m$ is assumed to span R^n.

Let $M_i = \begin{bmatrix} \sum_{k=1}^{m} a_k a_k^T & a_i \\ a_i^T & 1 \end{bmatrix}$. Since $\sum_{k=1}^{m} a_k a_k^T \succ 0$, $M_i \succeq 0 \Leftrightarrow 1 - a_i^T (\sum_{k=1}^{m} a_k a_k^T)^{-1} a_i \geq 0$

In other words, if we can show $M_i \succeq 0$, then $a_i^T X_{sim} a_i \leq 1$, for $i = 1, \cdots, m$.

Hence, X_{sim} is feasible.

Let $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in R^{n+1}$ where $v_1 \in R^n$, $v_2 \in R$ be any non-zero vector

$$v^T M_i v = \sum_{k=1}^{m} v_1^T a_k a_k^T v_1 + 2 \cdot v_2 a_i^T v_1 + v_2^2 = \sum_{\substack{k=1 \\ k \neq i}}^{m} (a_k^T v_1)^2 + (a_i^T v_1 + v_2)^2 \geq 0$$

Thus, $M_i \succeq 0$, $i = 1, \cdots, m$. Thus, X_{sim} is feasible.

(b). Consider the dual objective function when λ takes the form of $t\mathbb{1}$:

$$\begin{aligned} g(\lambda) &= \log\det\left(\sum_{i=1}^{m} t\, a_i a_i^T\right) - mt + n \\ &= \log\det\left(t \cdot \sum_{i=1}^{m} a_i a_i^T\right) - mt + n \\ &= \log\left(t^n \det\left(\sum_{i=1}^{m} a_i a_i^T\right)\right) - mt + n \\ &= \log\det\left(\sum_{i=1}^{m} a_i a_i^T\right) + n\log t - mt + n \end{aligned}$$

Now we maximize the dual function w.r.t. t to find the best lower bound:

$$\frac{d g(\lambda)}{dt} = \frac{n}{t} - m = 0 \Rightarrow t^* = \frac{n}{m}$$

Thus the best lower bound is $g(\lambda) = \log\det\left(\sum_{i=1}^{m} a_i a_i^T\right) + n\log\frac{n}{m}$

The primal objective function value at $X = X_{sim}$ is $-\log\det\left(\sum_{i=1}^{m} a_i a_i^T\right)^{-1}$

so the duality gap associated with X_{sim} and λ is $n\log\frac{m}{n}$

The volume of the ellipsoid associated with X is $e^{-\frac{O}{2}}$ where $O = -\log\det X$

Then the volumn of ellipsoid associated with $X_{sim} \leq \left(\frac{m}{n}\right)^{\frac{n}{2}} \cdot V_{optimal}$

3. (a). We know $f_0, f_1, \dots, f_m$ are convex function. $\max\{0, f_i(x)\}$, $i=1,\dots,m$ is also a convex function because of point-wise maximum operation preserves convexity. $\alpha \max\{0, f_i(x)\}$, $i=1,\dots,m$ is convex because positive scaling preserves convexity. $\max\limits_{i=1,\dots,m} \alpha \max\{0, f_i(x)\}$ is also convex because, again, point-wise maximum preserves convexity. Finally, addition ~~ad~~ of two convex function also preserves convexity so $\phi(x) = f_0 + \max\limits_{i=1,\dots,m} \alpha \max\{0, f_i(x)\}$ is a convex function.

(b). The Lagrangian $L(x, y, \lambda, \mu) = f_0(x) + \alpha y + \sum_{i=1}^{m} \lambda_i (f_i(x) - y) + \mu(-y)$

$$= f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + (\alpha - \mu - \mathbb{1}^T \lambda) y$$

The dual function $g(\lambda, \mu) = \inf\limits_{x} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)\right) + \inf\limits_{y} \left((\alpha - \mu - \mathbb{1}^T \lambda) y\right)$

$$= \underbrace{g(\lambda)}_{\text{dual function of original problem}}, \quad \alpha - \mu - \mathbb{1}^T \lambda = 0$$

Hence, the dual problem can be written as

$$\begin{aligned} \max \quad & g(\lambda) \\ \text{s.t.} \quad & \mathbb{1}^T \lambda \le \alpha \\ & \lambda \succeq 0 \end{aligned}$$

(c). ~~λ^* is~~ If $\mathbb{1}^T \lambda^* < \alpha$, then λ^* is also optimal of the dual problem in (b). In the optimal solution of the primal problem (the penalty version), $y^* = 0$ because of complementary slackness. So the optimal solution of the primal problem (the penalty version) satisfies $f_i(x^*) \le 0$, $i = 1, \dots, m$. Hence, x^* is also a feasible solution of the orignal problem. Therefore, x^* is also an optimal solution of the original problem.

4. (a). The problem is a convex optimization problem because objective function e^{-x} is convex in (x,y) and constraints is $\{(x,y) \mid x=0, y>0\}$ which is also a convex set.

The optimal value is $p^* = e^{-0} = 1$

(b). The Lagrangian $L(x,y,\lambda) = e^{-x} + \frac{\lambda x^2}{y}$

The dual function $g(\lambda) = \inf_{x,y>0} \left(e^{-x} + \frac{\lambda x^2}{y}\right) = \begin{cases} 0, & \text{if } \lambda \geq 0 \\ -\infty, & \text{if } \lambda < 0 \end{cases}$

Hence, the dual problem is

$$\begin{array}{ll} \max & 0 \\ \text{s.t.} & \lambda \geq 0 \end{array}$$

with optimal value $d^* = 0$ and $p^* - d^* = 1$

(c). No, because we cannot find a point (x,y) from D such that $\frac{x^2}{y} < 0$.

(d). $p^*(u) = \begin{cases} +\infty \text{ (infeasible)}, & \text{if } u<0 \\ 1, & \text{if } u=0 \\ 0, & \text{if } u>0 \end{cases}$

$p^*(u) \geq 1 - \lambda^* u$ does not hold when $u>0$ since $lhs = 0$ but $rhs = 1$

5. The KKT condition for this problem is

$X \succ 0$, $Xs = y$, $X^{-1} = I + \frac{1}{2}(vs^T + sv^T)$, where $v \in \mathbb{R}^n$ is the Lagrangian multiplier for the inequality constraints.

Now, if we have $\tilde{X}$ and $\tilde{v}$ that satisfy this condition then we can express $\tilde{X}$ interms of s and y:

$$\tilde{X}^{-1} = I + (1 + y^T y) ss^T - ys^T - sy^T$$

This can be verified to be the inverse of X^* given above:

$$\tilde{X}^{-1} X^* = \left(I + yy^T - \frac{1}{s^T s} ss^T\right) + (1 + y^T y)(ss^T + sy^T - ss^T) - (ys^T + yy^T - ys^T) - \left(sy^T + (y^T y) sy^T - \frac{1}{s^T s} ss^T\right)$$

$$= I$$

6. For any feasible point x, we have $0 \geq f_i(x) \geq f_i(x^*) + \nabla f_i(x^*)^T(x-x^*)$, $i=1,\cdots,m$

Using $\lambda_i^* \geq 0$, we have
$$\begin{aligned} 0 &\geq \sum_{i=1}^{m} \lambda_i^* \left(f_i(x^*) + \nabla f_i(x^*)^T (x-x^*)\right) \\ &= \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*)^T (x-x^*) \\ &= 0 + (x-x^*)^T \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) \\ &= -\nabla f_0(x^*)^T (x-x^*)^T \end{aligned}$$

Hence, we have $\nabla f_0(x^*)^T (x-x^*)^T \geq 0$ for any feasible x.