HW5

Carl Mueller
CSCI 5254 - Convex Optimization

April 7, 2018

6.1

Proposition 1.

$$log(x+1) \leq x \tag{1}$$
$$-log(x+1) \geq x \tag{2}$$
$$x > -1 \tag{3}$$

Proposition 2.

$$-log(1-x) \text{ is convex} \tag{4}$$

Proposition 3.

$$\phi(||u||_\infty) = -a^2 log(1 - \frac{||u||_\infty}{a^2}) \tag{5}$$

Left inequality working backwards:

$$||u||_2^2 \leq -a^2 \sum_{i=1}^{m} log(1 - \frac{u_i^2}{a^2})$$
$$\sum_{i=1}^{m} \frac{|u_i|^2}{a^2} \leq -\sum_{i=1}^{m} log(1 - \frac{u_i^2}{a^2})$$

Given proposition 1:

$$\sum_{i=1}^{m} \frac{|u_i|^2}{a^2} \leq -\sum_{i=1}^{m} log(1 - \frac{u_i^2}{a^2})$$
$$\text{true when:}$$
$$-\frac{u_i^2}{a^2} \geq -1$$

Right inequality:

$$\text{Given: } u_i^2 \leq ||u_i||_\infty^2$$

$$\sum_{i=1}^{m} -log(1 - \frac{u_i^2}{a^2}) \leq \sum_{i=1}^{m} -log(1 - \frac{||u_i||_\infty^2}{a^2}) \text{ given proposition 1}$$

$$\sum_{i=1}^{m} -log(1 - \frac{u_i^2}{a^2}) \leq \frac{u_i^2}{||u||_2^\infty} \sum_{i=1}^{m} -log(1 - \frac{||u_i||_\infty^2}{a^2}) \text{ given } \frac{u_i^2}{||u||_2^\infty} \geq 1$$

$$-a^2 \sum_{i=1}^{m} log(1 - \frac{u_i^2}{a^2}) \leq -a^2 \frac{u_i^2}{||u||_2^\infty} \sum_{i=1}^{m} log(1 - \frac{||u_i||_\infty^2}{a^2})$$

$$-a^2 \sum_{i=1}^{m} log(1 - \frac{u_i^2}{a^2}) \leq \frac{u_i^2}{||u||_2^\infty} \phi(||u||_\infty)$$

6.9

To show convexity, the following level set must be convex:

$$S_\alpha = \{\, t_i \mid \max_{i=1,\ldots,k} |\frac{p(t_i)}{q(t_i)} - y_i| \leq \alpha \,\}$$

Due to absolute value, following inequalities must hold:

$$-\alpha q(t_i) \leq y_i q(t_i) - p(t_i) \leq \alpha q(t_i)$$

This is represent two inequalities that define a polyhedron and is therefore convex. Since the level set is convex, the original minimzation problem is at least quasiconvex.

7.3

Proposition 1.

$$P(x|y=1) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-z^2}{2}} dz \tag{6}$$

$$P(x|y=0) = 1 - \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{\frac{-z^2}{2}} dz \tag{7}$$

Ordering probability terms in order of $y = 1$ and $y = 0$, our total probability is:

$$p(a, b) = \prod_{i=1}^{q} P_i(a^T u_i + b|y = 1) \prod_{i=q+1}^{m} (1 - P_i(a^T u_i + b|y = 0))$$

The negative log likelihood:

$$l(a, b) = \sum_{i=1}^{q} -log(P_i(a^T u_i + b|y = 1)) + \sum_{i=q+1}^{m} -log(1 - P_i(a^T u_i + b|y = 0))$$

The negative log likelihood is a convex function, so minimizing this function is a convex optimization problem.

7.4

a)

Proposition 1.

$$\text{Sample mean: } u = \frac{1}{N}\sum_{k=1}^{N} y_k \tag{8}$$

$$\text{Covariance: } Y = \frac{1}{N}\sum_{k=1}^{N}(y_k - u)(y_k - u)^T \tag{9}$$

$$-\frac{N}{2}nlog(2\pi) - \frac{N}{2}log(det(R)) - \frac{1}{2}R^{-1}\sum_{k=1}^{N}(y_k - a)(y_k - a)^T$$

$$= -\frac{N}{2}nlog(2\pi) - \frac{N}{2}log(det(R)) - \frac{1}{2}R^{-1}\sum_{k=1}^{N}(y_k y_k^T - a y_k^T - y_k a^T + aa^T)$$

$$= -\frac{N}{2}nlog(2\pi) - \frac{N}{2}log(det(R)) - \frac{1}{2}R^{-1}(\sum_{k=1}^{N} y_k y_k^T - \sum_{k=1}^{N} a y_k^T - \sum_{k=1}^{N} y_k a^T + \sum_{k=1}^{N} aa^T))$$

$$= -\frac{N}{2}nlog(2\pi) - \frac{N}{2}log(det(R)) - \frac{1}{2}R^{-1}(\sum_{k=1}^{N} y_k y_k^T - \sum_{k=1}^{N} a y_k^T - \sum_{k=1}^{N} y_k a^T + Naa^T)$$

Substitute sample mean:

$$= -\frac{N}{2}nlog(2\pi) - \frac{N}{2}log(det(R)) - \frac{1}{2}R^{-1}\sum_{k=1}^{N} y_k y_k^T - Nay^T - Nua^T + Naa^T$$

$$= R^{-1}\sum_{k=1}^{N}(y_k - a)(y_k - a)^T - R^{-1}N(a-u)(a-u)^T$$

$$= -\frac{N}{2}nlog(2\pi) - \frac{N}{2}log(det(R)) - \frac{1}{2}(NR^{-1}Y + R^{-1}N(a-u)(a-u)^T)$$

$$= -\frac{N}{2}nlog(2\pi) - \frac{N}{2}log(det(R)) - \frac{1}{2}(Ntr(R^{-1}Y) + N(a-u)R^{-1}(a-u)^T)$$

Set the gradient to zero to see a and R optimal values.

$$\nabla_a l(R, a) = -2R^{-1}(a-u) = 0$$

$$\therefore a = u$$

$$\nabla_R l(R, a) = -R^{-1} + R^{-1}(Y - (a-u)(a-u)^T)R^{-1} = 0$$

$$R = Y + (a-u)(a-u)^T$$

$$R = Y + (0)(0)^T$$

$$\therefore R = Y$$

7.8

Express sign function as a probability where we order values with $y > 1$ followed by $y < 0$:

$$\prod_{i=1}^{k} prob(a_i^T x + b_i + v_i > 0) \prod_{i=k+1}^{m} prob(a_i^T x + b_i + v_i < 0)$$

Since a_i and b_i are known values, the only RV is the noise term. We can epxress v_i as an expression of $a_i^T x + b_i$. P represents the cumulative density function of v_i. We can represent the probability as follows:

$$\prod_{i=1}^{k} P(-a_i^T x - b_i) \prod_{i=k+1}^{m} 1 - P(-a_i^T x - b_i)$$

Log likelihood below is concave so if we maximize, we obtain a convex problem:

$$l(x) = \sum_{i=1}^{k} log(P(-a_i^T x - b_i)) + \sum_{i=k+1}^{m} log(1 - P(-a_i^T x - b_i))$$

7.9

Given

$$y_i = f(a_i^T x + b_i + v_i), i = 1, \ldots, m$$

We know that a_i and b_i are knowns, so lets expression the random variable v_i as an expression of all other terms. We assume that f is an invertible function.

$$v_i = f^{-1}(y_i) - a_i^T x - b_i$$

The probability of observing $y_i, \ldots, y_m$ is:

$$\prod_{i=1}^{m} prob(f^{-1}(y_i) - a_i^T x - b_i)$$

$$l(x, f) = \sum_{i=1}^{m} log(prob(f^{-1}(y_i) - a_i^T x - b_i))$$

This log probability is concave w.r.t x and f. Thus maximizing generates a convex optimization problem.

Additional Exercises:

3.9

a)

Given:

$$z = [\Re x, \Im x]$$

Setup a system of equations using the vector breakdown of x for its $\Re$ and $\Im$ components:

$$||x||_2^2 = ||z||_2^2$$
$$\begin{bmatrix} \Re A & -\Im A \\ \Im A & \Re A \end{bmatrix} \begin{bmatrix} \Re x \\ \Im x \end{bmatrix} = \begin{bmatrix} \Re b \\ \Im b \end{bmatrix}$$

This becomes the optimization problem:

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & ||z||_2 \\ \text{subject to} \quad & \begin{bmatrix} \Re A & -\Im A \\ \Im A & \Re A \end{bmatrix} \begin{bmatrix} \Re x \\ \Im x \end{bmatrix} = \begin{bmatrix} \Re b \\ \Im b \end{bmatrix} \end{aligned}$$

b)

Define the second order cone:

$$K_i = \{\, (z, t) \mid ||z||_2 \le t \,\}$$

The SOCP:

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & |z||_2 \\ & \begin{bmatrix} \Re A & -\Im A \\ \Im A & \Re A \end{bmatrix} \begin{bmatrix} \Re x \\ \Im x \end{bmatrix} = \begin{bmatrix} \Re b \\ \Im b \end{bmatrix} \end{aligned}$$

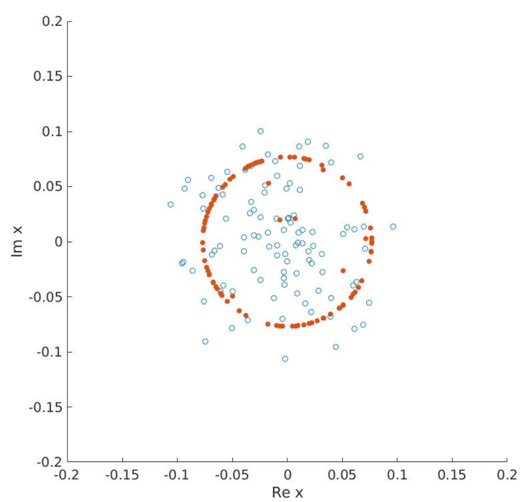
c)

Code:

```
randn('state',0);
m = 30; n = 100;
Are = randn(m,n); Aim = randn(m,n);
bre = randn(m,1); bim = randn(m,1);
A = Are + i*Aim;
b = bre + i*bim;

Atot = [Are -Aim; Aim Are];
btot = [bre; bim];
z_2 = Atot'*inv(Atot*Atot')*btot;
x_2 = z_2(1:100) + i*z_2(101:200);
```

```
cvx_begin
    variable x(n) complex
    minimize( norm(x) )
    subject to
    A*x == b;
cvx_end

cvx_begin
    variable xinf(n) complex
    minimize( norm(xinf,Inf) )
    subject to
    A*xinf == b;
cvx_end

figure(1)
scatter(real(x),imag(x)), hold on,
scatter(real(xinf),imag(xinf),[],'filled'), hold off,
axis([-0.2 0.2 -0.2 0.2]), axis square,
xlabel('Re x'); ylabel('Im x');
```

Results: The red dots represent the infinity norm.



4.1

a)

Code:

```
M = [1 -1/2; -1/2 2];
m = [-1 0]';
A = [1 2; 1 -4; 5 76];
```

```
b = [-2 -3 1]';
delta = .1

cvx_begin
    variable x(2)
    dual variable y
    minimize(quad_form(x, M)+m'*x)
    subject to
        y: A*x <= b;
cvx_end
p_star = cvx_optval
y
x
```

Results:
p_star = 8.2222
y =
1.8994
3.4684
0.0931
x =
-2.3333
0.1667

<u>KKT Conditions</u>
Primal:

$$x_1^* + 2x_2^* \leq u_1$$
$$x_1^* + -4x_2^* \leq u_2$$
$$5x_1^* + 76x_2^* \leq 1$$

Dual:

$$\lambda_1^*, \lambda_2^*, \lambda_3^* \geq 0$$

Complementary Slackness:

$$\lambda_1^*(x_1^* + 2x_2^* - u_1) = 0$$
$$\lambda_2^*(x_1^* + -4x_2^* - u_2) = 0$$
$$\lambda_3^*(5x_1^* + 76x_2^* - 1) = 0$$

First Order Conditions:

$$4x_2^* - x_1^* + 2\lambda_1^* - 4\lambda_2^* + 76\lambda_3^* = 0$$
$$2x_1^* - x_2^* - 1 + \lambda_1^* + \lambda_2^* + 5\lambda_2^* = 0$$

b)

Code:

```
M = [1 -1/2; -1/2 2];
m = [-1 0]';
A = [1 2; 1 -4; 5 76];
b = [-2 -3 1]';

cvx_begin
    variable x(2)
    dual variable y
    minimize(quad_form(x, M)+m'*x)
    subject to
        y: A*x <= b;
cvx_end
p_star = cvx_optval

array = [0 -1 1];
table = [];
delta = 0.1;

for i = array
    for j = array
        p_pred = p_star - [y(1) y(2)]*[i; j]*delta;
        cvx_begin
            variable x(2)
            minimize(quad_form(x,M)+m'*x)
            subject to
                A*x <= b+[i;j;0]*delta
        cvx_end
        p_exact = cvx_optval;
        table = [table; i*delta j*delta p_pred p_exact]
    end
end
```

Results:

d_1	d_2	p^*_{pred}	p^*_{exact}
0	0	8.2222	8.2222
0	-0.1000	8.5691	8.7064
0	0.1000	7.8754	7.9800
-0.1000	0	8.4122	8.5650
-0.1000	-0.1000	8.7590	8.8156
-0.1000	0.1000	8.0653	8.3189
0.1000	0	8.0323	8.2222
0.1000	-0.1000	8.3791	8.7064
0.1000	0.1000	7.6854	7.7515

We can see that $p^*_{pred} \leq p^*_{exact}$ for all pertubations.