Exercise 1.

In this problem, $\sum_a = \sum_{a=1}^k, \sum_{a,b} = \sum_{1\le a,b\le k}, \sum_{a<b} = \sum_{1\le a<b\le k}$.
Question 1. From delta method and saddle point approximation, we have

$$\frac{1}{2^{dk}}\int_{[-1,1]^{dk}}\prod_{a=1}^k \delta(\langle \boldsymbol{x}_0,\boldsymbol{x}_a\rangle - d\mu_a)\prod_{1\le a\le b\le k}\delta(\langle \boldsymbol{x}_a,\boldsymbol{x}_b\rangle - dq_{ab})\prod_{a=1}^k d\boldsymbol{x}_a$$

$$\doteq \frac{1}{2^{dk}}\int_{[-1,1]^{dk}}\int_{\mathbb{R}^{k+k(k+1)/2}}\exp\left(\sum_{a=1}^k -i\lambda_{0a}(\langle \boldsymbol{x}_0,\boldsymbol{x}_a\rangle - d\mu_a)\right.$$

$$\left.+\frac{1}{2}\sum_{1\le a,b\le k}-i\lambda_{ab}(\langle \boldsymbol{x}_a,\boldsymbol{x}_b\rangle - dq_{ab})\right)\prod d\lambda\prod d\boldsymbol{x}_a$$

$$\doteq \frac{1}{2^{dk}}\mathrm{ext}_{\boldsymbol{\lambda},\boldsymbol{\Lambda}}\exp\left(-d(\langle\boldsymbol{\lambda},\boldsymbol{\mu}\rangle + \frac{1}{2}\langle\boldsymbol{\Lambda},\boldsymbol{Q}\rangle)\right)\left(\int_{[-1,1]^k}\exp(\sum_a \lambda_a x_0 x_a + \frac{1}{2}\sum_{a,b}\lambda_{ab}x_a x_b)\prod dx_a\right)^d$$

Therefore, by taking d-limit, we have

$$\mathrm{Ent}(k,\boldsymbol{\mu},\boldsymbol{Q}) = \mathrm{ext}_{\boldsymbol{\lambda},\boldsymbol{\Lambda}} - \langle\boldsymbol{\lambda},\boldsymbol{\mu}\rangle - \frac{1}{2}\langle\boldsymbol{\Lambda},\boldsymbol{Q}\rangle + \log\frac{1}{2^k}\int_{[-1,1]^k}\exp(\sum_a \lambda_a x_0 x_a \frac{1}{2}\sum_{a,b}\lambda_{ab}x_a x_b)\prod dx_a$$

Question 2. With $\mu_a = \mu$, $q_{aa} = q_1$, $q_{ab} = q_0$(for $a \ne b$), we also additionally use replica symmetric ansatz for $\boldsymbol{\lambda},\boldsymbol{\Lambda}$ so that $\lambda_a = \lambda$, $\lambda_{aa} = \lambda_1$, $\lambda_{ab} = \lambda_0$(for $a\ne b$). Then we have

$$\mathrm{ent}(k,\mu,q_0,q_1) = \mathrm{ext}_{\lambda,\lambda_0,\lambda_1} - k\mu\lambda - \frac{1}{2}\left(k(k-1)\lambda_0 q_0 + k\lambda_1 q_1\right) + \log\frac{1}{2^k}\int_{[-1,1]^k}\exp\left(\lambda x_0(\sum_a x_a)\right.$$

$$\left.+\lambda_0\sum_{a<b}x_a x_b + \frac{\lambda_1}{2}\sum_a x_a^2\right)\prod_a dx_a$$

$$= \mathrm{ext}_{\lambda,\lambda_0,\lambda_1} - k\mu\lambda - \frac{1}{2}\left(k(k-1)\lambda_0 q_0 + k\lambda_1 q_1\right) - k\log 2$$

$$+\log\int_{[-1,1]^k}\exp\left(\frac{\lambda_0}{2}(\sum_a x_a)^2 + \lambda x_0\sum_a x_a + \frac{\lambda_1-\lambda_0}{2}\sum_a x_a^2\right)\prod_a dx_a.$$

Question 3. Let $G\sim\mathcal{N}(0,1)$. Then we have

$$\log\int_{[-1,1]^k}\exp\left(\frac{\lambda_0}{2}(\sum_a x_a)^2 + \lambda x_0\sum_a x_a + \frac{\lambda_1-\lambda_0}{2}\sum_a x_a^2\right)\prod_a dx_a$$

$$=\mathbb{E}_G\int_{[-1,1]^k}\exp\left(\sqrt{\lambda_0}(\sum_a x_a)G + \lambda x_0\sum_a x_a + \frac{\lambda_1-\lambda_0}{2}\sum_a x_a^2\right)\prod_a dx_a$$

$$=\mathbb{E}_G\left(\int_{[-1,1]}\exp((\lambda x_0 + \sqrt{\lambda_0}G)x + \frac{\lambda_1 - \lambda 0}{2}x^2)dx\right)^k$$

and thus with reverse replica trick, we conclude

$$T(\mu,q_0,q_1) = \mathrm{ext}_{\lambda,\lambda_0,\lambda_1} - \mu\lambda - \frac{1}{2}\lambda_1 q_1 + \frac{1}{2}\lambda_0 q_0 - \log 2$$

$$+\mathbb{E}_{x_0,G}\log\int_{[-1,1]}\exp\left(\left(\frac{\lambda_1-\lambda_0}{2}\right)x^2 + (\lambda x_0 + \sqrt{\lambda_0}G)x\right)dx$$

Exercise 2.

Question 1.

$$\lim_{n\to\infty} \mathbb{E}[L_n] = \lim_{n\to\infty}\lim_{\beta\to\infty} -\frac{1}{n\beta}\mathbb{E}\log\int_{[-1,1]^d} d\boldsymbol{x}\exp\{-\frac{\beta}{2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|_2^2\}$$
$$= \lim_{n\to\infty}\lim_{\beta\to\infty} -\frac{1}{n\beta}\lim_{k\to 0}\frac{1}{k}\log\int_{[-1,1]^{k\times d}}\prod_{a=1}^{k} d\boldsymbol{x}_a\mathbb{E}\exp\{-\frac{\beta}{2}\sum_{a=1}^{k}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}_a\|_2^2\}$$

By changing the order of limit, we would take the n-limit first and define

$$S(k,\beta) = \lim_{n\to\infty}\frac{1}{n}\log\int_{[-1,1]^{k\times d}}\prod_{a=1}^{k} d\boldsymbol{x}_a\mathbb{E}\exp\{-\frac{\beta}{2}\sum_{a=1}^{k}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}_a\|_2^2\}.$$

By similar calculation from lecture 12, we have

$$S(k,\beta) \doteq \sup_{\boldsymbol{Q},\boldsymbol{\mu}}\lim_{n\to\infty}\frac{1}{n}[\log E(\boldsymbol{Q},\boldsymbol{\mu}) + \log \mathrm{Ent}(\boldsymbol{Q},\boldsymbol{\mu})]$$

where E is exactly the same defined in class, and Ent is defined by

$$\mathrm{Ent}(\boldsymbol{Q},\boldsymbol{\mu}) = \int_{[-1,1]^{dk}}\delta(\bar{\boldsymbol{Q}}(\boldsymbol{x})-\boldsymbol{Q})\delta(\bar{\boldsymbol{\mu}}(\boldsymbol{x})-\boldsymbol{\mu})\prod_{a=1}^{k} d\boldsymbol{x}_a.$$

Note that we already have k-limit of Ent calculated from **Exercise 1**. Thus using

$$q_0 = q,\ q_1 = q + \frac{w}{\beta},\ \lambda_0 = \beta^2\rho,\ \lambda_1 = \beta^2\rho - \beta\nu,\ \lambda = \beta\zeta,$$

We have both n,k-limit of E and Ent resulting as

$$\lim_{k\to 0}\frac{1}{k}S(k,\beta) = \mathrm{ext}_{q,w,\mu,\rho,\nu,\zeta}\,\bar{e} + \bar{\mathrm{ent}},$$

where

$$\bar{e} = -\frac{1}{2}\log\left(1+\frac{w}{\delta}\right) - \frac{\beta}{2(\delta+w)}(1-2\mu+q+\gamma\sigma^2)$$

and

$$\begin{aligned}\bar{\mathrm{ent}} =& \frac{1}{\delta}(-\log 2 + \frac{1}{2}(\beta\nu q + \nu w - \beta\rho w) - \beta\mu\zeta \\ &+ \mathbb{E}_{G,x_0}\log\int_{[-1,1]}\exp\{-\frac{1}{2}\beta\nu x^2 + (\beta\nu x_0 + \sqrt{\rho}\beta G)x\}dx)\end{aligned}$$

Now we multiply $-1/\beta$ and take β-limit, then we have

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}[L_n] = \mathrm{ext}_{q,w,\mu,\rho,\nu,\zeta}&\frac{1}{2(\delta+w)}(1-2\mu+q+\delta\sigma^2) + \frac{1}{\delta}((\mu\zeta + \frac{1}{2}(\rho w - \nu q)) \\ &- \mathbb{E}_{G,x_0}\max_{x\in[-1,1]}(-\frac{1}{2}\nu x^2 + (\nu x_0 + \sqrt{\rho}G)x))\end{aligned}$$

Taking derivative w.r.t q, μ gives

$$\nu = \zeta = \frac{\delta}{\delta + w},$$

and plugging in, we conclude

$$\lim_{n\to\infty} \mathbb{E}[L_n] = \text{ext}_{\rho,\nu} \frac{\nu}{2\delta}(1+\delta\sigma^2) + \frac{\rho}{2}(1/\nu - 1) - \frac{1}{\delta}\mathbb{E}_{G,x_0} \max_{x\in[-1,1]} (-\frac{1}{2}\nu x^2 + (\nu x_0 + \sqrt{\rho}G)x)$$

Note that $\hat{x} = \Pi_{[-1,1]}(x_0 + \frac{\sqrt{\rho}}{\nu}G) = \arg\max_{x\in[-1,1]}(-\frac{1}{2}\nu x^2 + (\nu x_0 + \sqrt{\rho}G)x)$ where $\Pi_A(x)$ is orthogonal projection of x into A. Then we have

$$\lim_{n\to\infty} \mathbb{E}[L_n] = \frac{\nu^*}{2\delta}(1+\delta\sigma^2) + \frac{\rho^*}{2}(1/\nu^* - 1) - \frac{\nu^*}{2\delta}\mathbb{E}_{G,x_0}(-\frac{1}{2}\nu\hat{x}^2 + (\nu x_0 + \sqrt{\rho}G)\hat{x})$$

where ν^*, ρ^* satisfy the stationary equation which are:

$$\frac{\delta}{2}(\frac{1}{\nu} - 1) = \frac{1}{2\sqrt{\rho}}\mathbb{E}[\hat{x}G]$$
$$\frac{1}{2}\mathbb{E}(\hat{x} - x_0)^2 + \delta(\frac{1}{2}\sigma^2 - \frac{\rho}{2\nu^2}) = 0.$$

Question 2. We add additional $h\sum_{i=1}^d \psi(x_i, x_{0,i})$ term in Hamiltonian from the calculation of Question 1 above, and then take d-limit instead of n-limit. Then we will have to multiply δ, and the only difference will be the expectation of the maximum part. Thus we have

$$\lim_{d\to\infty, n/d\to\delta} \frac{1}{d}\sum_{i=1}^d \mathbb{E}[\psi(\hat{x}_i, x_{0,i})] = \partial_h f(h)|_{h=0}$$

where

$$f(h) = \text{ext}_{\rho,\nu} \frac{\nu}{2}(1+\delta\sigma^2) + \frac{\rho\delta}{2}(1/\nu - 1) - \mathbb{E}_{G,x_0} \max_{x\in[-1,1]} (-\frac{1}{2}\nu x^2 + (\nu x_0 + \sqrt{\rho}G)x - h\psi(x, x_0)).$$

By implicit differentiation we have

$$\lim_{d\to\infty, n/d\to\delta} \frac{1}{d}\sum_{i=1}^d \mathbb{E}[\psi(\hat{x}_i, x_{0,i})] = \mathbb{E}_{G,x_0}\psi(\hat{x}, x_0)$$

where

$$\hat{x} = \arg\max_{x\in[-1,1]} -\frac{1}{2}\nu^* x^2 + (\nu^* x_0 + \sqrt{\rho^*}G)x = \Pi_{[-1,1]}(x_0 + \frac{\sqrt{\rho}}{\nu}G)$$

and ν^*, ρ^* satisfies the stationary condition, which are:

$$\frac{\delta}{2}(\frac{1}{\nu} - 1) = \frac{1}{2\sqrt{\rho}}\mathbb{E}[\hat{x}G]$$
$$\frac{1}{2}\mathbb{E}(\hat{x} - x_0)^2 + \delta(\frac{1}{2}\sigma^2 - \frac{\rho}{2\nu^2}) = 0.$$

Then by substituting $\tau^2 = \rho/\nu^2$, we have stationary condition replaced by

$$\frac{\delta}{2}(1 - \nu) = \frac{1}{\tau}\mathbb{E}[\Pi_{[-1,1]}(x_0 + \tau G)G]$$
$$\tau^2 = \sigma^2 + \frac{1}{\delta}\mathbb{E}(\Pi_{[-1,1]}(x_0 + \tau G) - x_0)^2.$$

Note that the second equation solely decided τ, and also $\hat{x}$ only depends on τ not ν. Thus we conclude

$$\lim_{d\to\infty, n/d\to\delta} \frac{1}{d}\sum_{i=1}^{d} \mathbb{E}[\psi(\hat{x}_i, x_{0,i})] = \mathbb{E}\psi(\Pi_{[-1,1]}(x_0+\tau G), x_0) = \delta(\tau^{*2}-\sigma^2),$$

where τ^* solves

$$\tau^2 = \sigma^2 + \frac{1}{\delta}\mathbb{E}(\Pi_{[-1,1]}(x_0+\tau G)-x_0)^2$$

The case where $\psi(x,y)=(x-y)^2$, we have

$$\lim_{d\to\infty, n/d\to\delta} \mathbb{E}\|\hat{\boldsymbol{x}}-\boldsymbol{x}_0\|_2^2/d = \mathbb{E}_G(\Pi_{[-1,1]}(1+\tau^* G)-1)^2$$

where τ^* solves

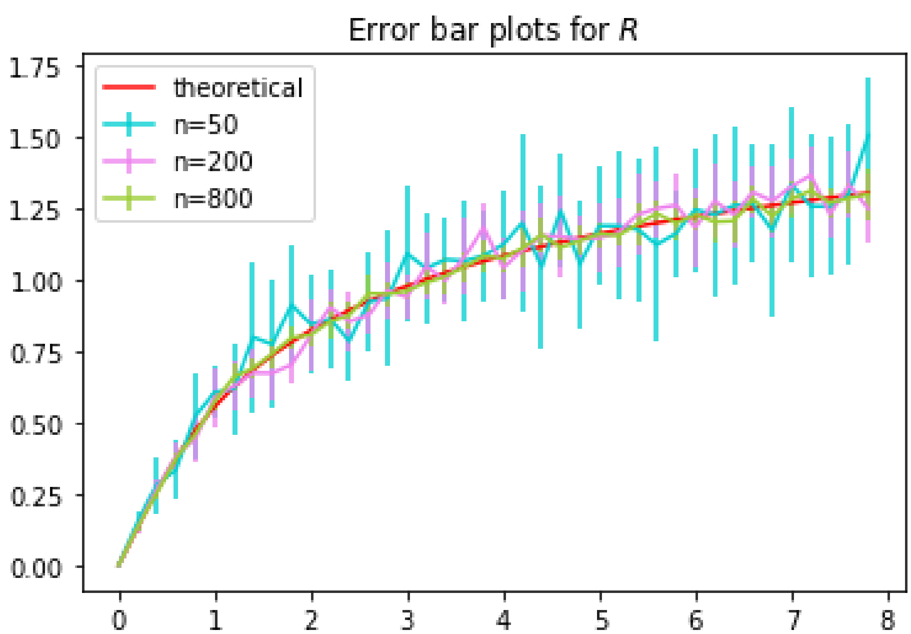
$$\tau^2 = \sigma^2 + \frac{1}{\delta}\mathbb{E}_G(\Pi_{[-1,1]}(1+\tau G)-1)^2$$

Here we replace $x_0 = 1$ by symmetry. To calculate theoretical estimate in **Exercise 4.** we use

$$\mathbb{E}_G(\Pi_{[-1,1]}(1+\tau G)-1)^2 = \tau^2(\Phi(2/\tau)-1/2-2\phi(2/\tau)/\tau)) + 4\Phi(-2/\tau)$$

where Φ, ϕ is cdf and pdf of standard Gaussian distribution.

Exercise 3,4 Codes are attached below.



In [9]:
```
import cvxpy as cp
import numpy as np
import random
import matplotlib.pyplot as plt
import time
import scipy
import scipy.stats
from scipy.stats import norm
from scipy.optimize import fsolve
```

In [2]:
```
d=50
n=2*d
sigmasq=4
ns=10
```

In [3]:
```
sigmasqs=[0.2*i for i in range(40)]
```

In [4]:
```
d=50
n=2*d
er1=[]
stdr1=[]
for i in range(40):
    sigmasq=0.2*i
    rs=[]
    for j in range(ns):
        A=np.random.normal(0,(1/n)**0.5,size=(n,d))
        x0=np.random.randint(2,size=d)
        x0=[2*x-1 for x in x0]
        e=np.random.normal(0,sigmasq**0.5,n)
        y=A@x0+e
        u=cp.Variable(d)
        const=[u<=1,-u<=1]
        l=cp.norm(y-A@u,2)
        prob=cp.Problem(cp.Minimize(l),const)
        prob.solve()
        rs+=[np.linalg.norm(u.value-x0,2)**2/d]
    er1+=[np.mean(rs)]
    stdr1+=[np.std(rs)]
```

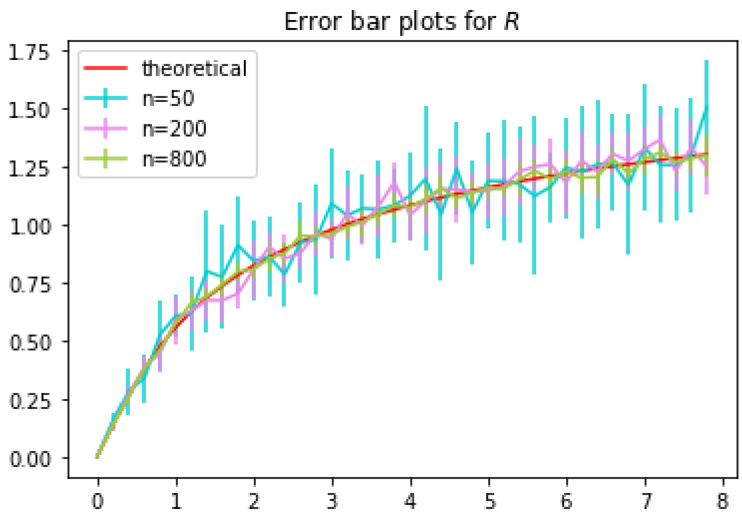
In [5]:
```
d=200
n=2*d
er2=[]
stdr2=[]
for i in range(40):
    sigmasq=0.2*i
    rs=[]
    for j in range(ns):
        A=np.random.normal(0,(1/n)**0.5,size=(n,d))
        x0=np.random.randint(2,size=d)
        x0=[2*x-1 for x in x0]
        e=np.random.normal(0,sigmasq**0.5,n)
        y=A@x0+e
        u=cp.Variable(d)
        const=[u<=1,-u<=1]
        l=cp.norm(y-A@u,2)
        prob=cp.Problem(cp.Minimize(l),const)
        prob.solve()
        rs+=[np.linalg.norm(u.value-x0,2)**2/d]
    er2+=[np.mean(rs)]
    stdr2+=[np.std(rs)]
```

In [6]:
```
d=800
n=2*d
er3=[]
stdr3=[]
for i in range(40):
    sigmasq=0.2*i
    rs=[]
    for j in range(ns):
        A=np.random.normal(0,(1/n)**0.5,size=(n,d))
        x0=np.random.randint(2,size=d)
        x0=[2*x-1 for x in x0]
        e=np.random.normal(0,sigmasq**0.5,n)
        y=A@x0+e
        u=cp.Variable(d)
        const=[u<=1,-u<=1]
        l=cp.norm(y-A@u,2)
        prob=cp.Problem(cp.Minimize(l),const)
        prob.solve()
        rs+=[np.linalg.norm(u.value-x0,2)**2/d]
    print(i)
    er3+=[np.mean(rs)]
    stdr3+=[np.std(rs)]
```

```
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
```

In [11]:
```
def eq(tau,sigma):
    return 2*sigma-2*tau**2+(tau**2)*(norm.cdf(0)-norm.cdf(-2/tau)-2*nor
m.pdf(2/tau)/tau)+4*norm.cdf(-2/tau)
theo=[]
for sigmasq in sigmasqs:
    def taueq(a):
        return eq(a,sigmasq)
    taust=fsolve(taueq,0.5)[0]
    theo+=[2*(taust**2-sigmasq)]
```

In [13]:
```
plt.errorbar(x=sigmasqs,y=er1,yerr=stdr1,color='darkturquoise',label='n=
50')
plt.errorbar(x=sigmasqs,y=er2,yerr=stdr2,color='violet',label='n=200')
plt.errorbar(x=sigmasqs,y=er3,yerr=stdr3,color='yellowgreen',label='n=80
0')
plt.plot(sigmasqs,theo,color='red',label='theoretical')
plt.title('Error bar plots for $R$')
plt.legend()
plt.show()
```



In [ ]: