This homework is due at **2pm, March 4, 2019**.

Problem 1

Let $f : \mathbb{R}^p \to \mathbb{R}$. Its *gradient* is a p-dimensional vector given by

$$\nabla f(x) := \left(\frac{\partial f}{\partial x^{(1)}}(x), \dots, \frac{\partial f}{\partial x^{(p)}}(x) \right), \quad \forall x \in \mathbb{R}^p,$$

where $x^{(i)}$ denotes the i-th entry of the vector x. Its *Hessian* is a matrix in $\mathbb{R}^{p\times p}$ given by

$$\left[\nabla^2 f(x) \right]^{(i,j)} := \frac{\partial^2 f}{\partial x^{(i)} \partial x^{(j)}}(x), \quad \forall x \in \mathbb{R}^p,$$

for all $1 \le i, j \le p$, where $\left[\nabla^2 f(x) \right]^{(i,j)}$ denotes the (i, j)-th entry of the matrix $\nabla^2 f(x)$.
Define

$$g(x) := \frac{1}{2} \langle x, Ax \rangle, \quad \forall x \in \mathbb{R}^p,$$

for some symmetric matrix $A \in \mathbb{R}^{p\times p}$.

1. (10 points) **Show that**

$$\nabla g(x) = Ax, \quad \forall x \in \mathbb{R}^p.$$

2. (15 points) **Show that**

$$\nabla^2 g(x) = A, \quad \forall x \in \mathbb{R}^p.$$

Problem 2

Let $\mathscr{X} \subset \mathbb{R}^{p\times p}$ be the set of positive semi-definite matrices of unit trace.

1. (10 points) **Show that for each $X \in \mathscr{X}$, we can write**

$$X = \sum_{i=1}^{p} \lambda_i \Pi_i,$$

   where $\lambda_1, \dots, \lambda_p$ are non-negative real numbers summing up to one, and Π_i are rank-1 matrices of unit trace.

2. (15 points) Let $Y \in \mathbb{R}^{p\times p}$ which may not be in $\mathscr{X}$. Define $f(X) := \mathrm{Tr}(YX)$, the trace of YX. Define

$$f^\star = \min_X \left\{ f(X) \mid X \in \mathscr{X} \right\}.$$

   Show that there exists some rank-1 matrix $X^\star \in \mathscr{X}$, such that $f^\star = f(X^\star)$.

Problem 3

Let ξ be a non-negative real-valued random variable, and η be a real-valued random variable.

1. (10 points) **Show that**

$$\mathsf{P}(\xi \ge t) \le \frac{\mathsf{E}\,\xi}{t}, \quad \forall t > 0,$$

   where $\mathsf{P}(\xi \ge t)$ denotes the probability that $\xi \ge t$, and $\mathsf{E}\,\xi$ denotes the expectation of ξ.

Problem 3 continued on the next page...

2. (15 points) Suppose that η is a subgaussian random variable of parameter $\lambda > 0$, i.e.,

$$\mathsf{E} \exp\left(\lambda \eta^2\right) \leq 2.$$

Show that

$$\mathsf{P}\left(|\eta| \geq t\right) \leq 2 \exp\left(-\lambda t^2\right), \quad \forall t \geq 0.$$

Problem 4

Let $\xi_1, \ldots, \xi_n$ be a sequence of independent Bernoulli random variable of expectation $\theta^\natural \in [0,1]$, i.e.,

$$\mathsf{P}(\xi_i = 1) = 1 - \mathsf{P}(\xi_i = 0) = \theta^\natural, \quad 1 \leq i \leq n.$$

The maximum-likelihood estimator of θ is given by any minimizer of the function

$$f_n(\theta) := -\frac{1}{n} \sum_{i=1}^{n} \left[\xi_i \log\theta + (1 - \xi_i) \log(1 - \theta) \right],$$

on the interval $[0,1]$.

1. (10 points) **Show that indeed, the maximum-likelihood estimator, which we denote by $\hat{\theta}_n$, is given by**

$$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^{n} \xi_i.$$

2. (15 points) **Show that as $n \to +\infty$, it holds that $\hat{\theta}_n$ converges to $\theta^\natural$ in probability, i.e.,**

$$\lim_{n \to +\infty} \mathsf{P}\left(\left|\hat{\theta}_n - \theta^\natural\right| \geq t\right) = 0, \quad \forall t > 0.$$

   Do not simply cite the law of large numbers.