Problem 1 (Kurtosis of a sum). Recall that the kurtosis of a random variable $X$ with zero mean is

$$kurt(X) = E[X^4] - 3(E[X^2])^2. \quad (1.1)$$

One way to understand the importance of the factor 3 in (1.1) is to consider a sum of two independent random variables. Let $X$ and $Y$ be two independent random variables with zero mean and unit variance, i.e.

$$E[X] = 0, \quad E[X^2] = 1 \quad (1.2)$$
$$E[Y] = 0, \quad E[Y^2] = 1. \quad (1.3)$$

Show that

$$kurt(X + Y) = kurt(X) + kurt(Y). \quad (1.4)$$

Can you see the importance of factor 3?

Hint: Use binomial formula and linearity of expectation.

Problem 2 (Kurtosis of normal distribution). Another way to see the importance of the factor 3 is to consider the kurtosis of normal distribution. We will prove that if $X$ is normally distributed with 0 mean, then $kurt(X) = 0$. To compute the kurtosis, we need the fourth moment $E[X^4]$. To compute it, we use very powerful and general technique of moment-generating functions. The moment-generating function of random variable $Y$ is

$$M_Y(t) = E[e^{tY}], \quad t \in \mathbb{R}. \quad (2.1)$$

One important feature of moment-generating functions is that if we know $M_Y$, we can easily compute the $n$th moment of $Y$ by differentiating $M_Y$ $n$ times and evaluating the derivative at origin. In other words,

$$\frac{d^n M_Y}{dt^n}(0) = E[Y^n], \quad (2.2)$$

where $\frac{d^n M_Y}{dt^n}(0)$ is the $n$th derivative of $M_Y$ evaluated at origin. (Here we assume that the derivative exists.)

The moment-generating function for normally distributed $X$ with 0 mean and variance $\sigma^2$ is

$$M_X(t) = \exp(\sigma^2 t^2 / 2). \quad (2.3)$$

Use (2.3) to compute $E[X^4]$ and conclude that $kurt(X) = 0$.

Problem 3 (Traces and eigenvalues). Let $A = Q\Lambda Q^T \in \mathbb{R}^{n \times n}$ be a matrix and its eigendecomposition. You can assume that $Q$ is orthogonal and that the eigenvalues are real. Show that

$$\text{tr}(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i = \text{tr}(\Lambda). \quad (3.1)$$
**Problem 4** (Laplacian is positive semi-definite). Let $G = (V, E)$ be an undirected graph and let $A$ be its adjacency matrix and $\Delta$ its degree matrix. Let $L = \Delta - A$ be the Laplacian of $G$. Recall that the incidence matrix $P$ of $G$ (for some fixed but arbitrary ordering of the edges) is the $|V|$-by-$|E|$ matrix with

\[
p_{ij} = \begin{cases} 
1 & \text{if edge } j \text{ starts from node } i \\
-1 & \text{if edge } j \text{ ends to node } i \\
0 & \text{otherwise.}
\end{cases}
\] (4.1)

Show that

\[L = PP^T\] (4.2)
and conclude that the Laplacian is positive semi-definite.

**Problem 5** (Normalized cut). Show that the solution for the relaxed normalized cut is obtained by taking the $k$ least eigenvectors of the symmetric normalized Laplacian $L^*$ similarly as the ratio cut is solved by taking the $k$ least eigenvectors of the Laplacian.

*Hint:* Express normalized cut using the symmetric Laplacian by re-writing the equation

\[J_{nc}(C) = \sum_{i=1}^{k} \frac{c_i^T L c_i}{c_i^T \Delta c_i},\]

using the facts that $\Delta = \Delta^{1/2} \Delta^{1/2}$, $\Delta^{1/2} \Delta^{-1/2} = I$, and $\Delta = \Delta^T$ (as $\Delta$ is diagonal).

**Problem 6** (Nuclear norm). Let $A \in \mathbb{R}^{n \times m}$ be an arbitrary matrix, and let $(\sigma_i)_{i=1}^{\min\{m,n\}}$ be its singular values. The nuclear norm of $A$, denoted $\|A\|_*$, is defined as

\[\|A\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i.\] (6.1)

Show that

\[\|A\|_* \geq \|A\|_F.\] (6.2)