Homework 1

Due Thursday, 8/27/20 on Canvas.

This homework covers basic ideas in linear algebra, probability and statistics that will be needed throughout the course.

1. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times m}$. Show that

$$\mathrm{Tr}(\boldsymbol{AB}) = \mathrm{Tr}(\boldsymbol{BA}).$$

2. Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\mathrm{Im}(\boldsymbol{A}) = \{\boldsymbol{y} \in \mathbb{R}^m : \boldsymbol{y} = \boldsymbol{Ax} \in \mathbb{R}^n \text{ for some } \boldsymbol{x} \in \mathbb{R}^n\}$ be the image of $\boldsymbol{A}$, $r(\boldsymbol{A}) = \dim\{\mathrm{Im}(\boldsymbol{A})\}$ be the rank of $\boldsymbol{A}$ and $\ker(\boldsymbol{A}) = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{Ax} = \boldsymbol{0}\}$ be the kernel of $\boldsymbol{A}$. Recall that $\mathrm{Im}(\boldsymbol{A}) \subseteq \mathbb{R}^m$ and $\ker(\boldsymbol{A}) \subseteq \mathbb{R}^n$ are vector subspaces.

   (a) For any $\boldsymbol{G} \in \mathbb{R}^{n\times q}$, show that $\mathrm{Im}(\boldsymbol{AG}) \subseteq \mathrm{Im}(\boldsymbol{A})$ and $\ker(\boldsymbol{G}) \subseteq \ker(\boldsymbol{AG})$.

   (b) If $\boldsymbol{G}$ has full row-rank (i.e. $\boldsymbol{G}$ is surjective), show that $\mathrm{Im}(\boldsymbol{AG}) = \mathrm{Im}(\boldsymbol{A})$.

   (c) For any matrix $\boldsymbol{B} \in \mathbb{R}^{m\times n}$, show that $r(\boldsymbol{A}+\boldsymbol{B}) \le r(\boldsymbol{A}) + r(\boldsymbol{B})$.

3. Let $W \subset \mathbb{R}^n$ be a vector subspace of $\mathbb{R}^n$. Define $W^\perp = \{\boldsymbol{v} \in \mathbb{R}^n : \boldsymbol{v}^T\boldsymbol{w} = 0 \text{ for all } \boldsymbol{w} \in W\}$ to be W's orthogonal complement.

   (a) Show that $W \cap W^\perp = \{\boldsymbol{0}\}$ and $(W^\perp)^\perp = W$.

   (b) Show that for any matrix $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\ker(\boldsymbol{A}^T) = \mathrm{Im}(\boldsymbol{A})^\perp$ and $\mathrm{Im}(\boldsymbol{A}^T) = \ker(\boldsymbol{A})^\perp$.

   (c) Conclude that the Fredholm Alternative holds:

$$\mathbb{R}^n = \mathrm{Im}(\boldsymbol{A}^T) \cup \ker(\boldsymbol{A}), \quad \mathrm{Im}(\boldsymbol{A}^T) \cap \ker(\boldsymbol{A}) = \boldsymbol{0}$$
$$\mathbb{R}^m = \mathrm{Im}(\boldsymbol{A}) \cup \ker(\boldsymbol{A}^T), \quad \mathrm{Im}(\boldsymbol{A}) \cap \ker(\boldsymbol{A}^T) = \boldsymbol{0}$$

   (d) Use part (c) to show that for all $\boldsymbol{x} \in \mathbb{R}^m$, there exists $\boldsymbol{x}_0 \in \mathrm{Im}(\boldsymbol{A})$, $\boldsymbol{x}_1 \in \ker(\boldsymbol{A}^T)$ such that $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_1$ and $\boldsymbol{x}_0^T\boldsymbol{x}_1 = 0$. Conclude that $\|\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}_0\|_2^2 + \|\boldsymbol{x}_1\|_2^2$.

4. Application of the CLT: Earth's current population is 7.594 billion, 49.6% of which are female. Assume that each child's biological sex is independent of the biological sexes of all other children.

   (a) Test the null hypothesis that the probability a child is born female is exactly 50%. What is your conclusion?

   (b) Using the data given in the problem statement and assuming that biological sexes are i.i.d random variables, calculate a 95% confidence interval for f, the probability that a child is born female.

   (c) Explain how you can use the confidence interval computed in part (b) to test the null hypothesis from part (a) at a significance level of $\alpha = 0.05$.

5. Let T_ν be a t-distribution with $\nu > 0$ degrees of freedom and χ^2_ν be a chi-squared distribution with ν degrees of freedom.

   (a) Using the fact that $T_\nu \overset{\mathcal{D}}{=} \frac{N(0,1)}{\sqrt{\chi^2_\nu/\nu}}$ (where the numerator and denominator are independent), show that

   $$P(T_\nu \leq -t \text{ or } T_\nu \geq t) = P\left(F_{1,\nu} \geq t^2\right)$$

   for all $t > 0$, where $F_{1,\nu}$ is the f-distribution with $1, \nu$ degrees of freedom.

   (b) Use the definition of a chi-squared random variable and the strong law of large numbers to show that

   $$n^{-1}\chi^2_n \overset{a.s.}{\to} 1 \quad \text{as } n \to \infty.$$

   Use this to conclude that

   $$\lim_{n\to\infty} P(T_n \leq t) = P\{N(0,1) \leq t\}$$

   for all $t \in \mathbb{R}$. That is, the t-distribution resembles a normal distribution for large degrees of freedom. (**Hint**: $\chi^2_n \overset{\mathcal{D}}{=} \sum_{i=1}^{n} Z_i^2$, where $Z_i \overset{i.i.d}{\sim} N(0,1)$.)