Recitation 6 Solutions

These problems will cover Markov's inequality, covariance matrices, PCA and convergence.

1. (Random Vector)

   (a) The covariance matrix equals

   $$\Sigma_{\tilde{x}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.25 & 0.25 \\ 0 & 0.25 & 0.25 \end{bmatrix}, \tag{1}$$

   so $\text{Var}(\tilde{x}_1) = 1$, $\text{Var}(\tilde{x}_2) = 0.25$, $\text{Var}(\tilde{x}_3) = 0.25$.

   (b) The maximum variance in any direction is given by the largest eigenvalue, which is equal to 1. There cannot be another direction with higher variance.

   (c) Let $\tilde{y} := a_1\tilde{x}_1 + a_2\tilde{x}_2 + a_3\tilde{x}_3$. By linearity of expectation, $\text{E}(\tilde{y}) = 0$. By Chebyshev's inequality, if $\text{Var}(\tilde{y}) = 0$ then $\text{P}(\tilde{y} \neq 0) = 0$, which is exactly what we want. According to the eigendecomposition, the variance is zero in the direction of the third eigenvector, so setting $a = 0$, $b = 1/\sqrt{2}$, and $c = -1/\sqrt{2}$ does the trick.

2. (Not centering) We have

   $$\begin{aligned} \text{E}(\tilde{x}\tilde{x}^T) &= \text{E}\left[(c(\tilde{x}) + \mu)(c(\tilde{x}) + \mu)^T\right] && (2) \\ &= \text{E}\left[c(\tilde{x})c(\tilde{x})^T\right] + \text{E}\left[c(\tilde{x})\mu^T\right] + \text{E}\left[\mu c(\tilde{x})^T\right] + \text{E}(\mu\mu^T) && (3) \\ &= \Sigma_{\tilde{x}} + \mu\mu^T && (4) \\ &= I + \mu\mu^T. && (5) \end{aligned}$$

   Let $u_1 := \mu / \left|\left|\mu\right|\right|_2$, and let $u_2, \ldots, u_d$ orthonormal vectors orthogonal to u_1, we have

   $$\begin{bmatrix} u_1 & u_2 & \cdots & u_d \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_d \end{bmatrix}^T = I, \tag{6}$$

   because it is an orthonormal set, so the matrix is orthogonal. This implies

   $$\begin{aligned} \text{E}(\tilde{x}\tilde{x}^T) &= I + \mu\mu^T && (7) \\ &= \begin{bmatrix} u_1 & u_2 & \cdots & u_d \end{bmatrix} \begin{bmatrix} \left|\left|\mu\right|\right|_2^2 + 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_d \end{bmatrix}^T && (8) \end{aligned}$$

   so the first eigenvalue equals $\left|\left|\mu\right|\right|_2^2 + 1$ and the corresponding eigenvector is collinear with the mean.

3. (Markov's inequality)

(a)

$$\sum_{x=1}^{n} x p_{\tilde{x}}(x) = \sum_{x<a} x p_{\tilde{x}}(x) + \sum_{x \geq a} x p_{\tilde{x}}(x) \tag{9}$$

$$\geq \sum_{x<a} x p_{\tilde{x}}(x) + a \sum_{x \geq a} p_{\tilde{x}}(x) \tag{10}$$

$$\geq a \sum_{x \geq a} p_{\tilde{x}}(x). \tag{11}$$

(b) For $a > 0$

$$\mathrm{P}(\tilde{x} \geq a) = \sum_{x \geq a} p_{\tilde{x}}(x) \tag{12}$$

$$\leq \frac{\sum_{x=1}^{n} x p_{\tilde{x}}(x)}{a} \quad \text{by the equation} \tag{13}$$

$$= \frac{\mathrm{E}(\tilde{x})}{a}. \tag{14}$$

(c) By Markov's inequality,

$$\mathrm{P}(\tilde{x} > 1000) \leq \frac{\mathrm{E}(\tilde{x})}{1000} = \frac{1}{2}$$

4. (Sample median as an estimator of the median)
Let's denote the sample median by $\tilde{y}(n)$ and the median of an iid sequence of random variables as γ. We want to show that for any $\epsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}(|\tilde{y}(n) - \gamma| \geq \epsilon) = 0. \tag{15}$$

We will prove that

$$\lim_{n\to\infty} \mathrm{P}(\tilde{y}(n) \geq \gamma + \epsilon) = 0. \tag{16}$$

The same argument allows to establish

$$\lim_{n\to\infty} \mathrm{P}(\tilde{y}(n) \leq \gamma - \epsilon) = 0. \tag{17}$$

If we order the set $\{\tilde{x}(1), \ldots, \tilde{x}(n)\}$, then $\tilde{y}(n)$ equals the $(n+1)/2$th element if n is odd and the average of the $n/2$th and the $(n/2+1)$th element if n is even. The event $\tilde{y}(n) \geq \gamma + \epsilon$ therefore implies that at least $(n+1)/2$ of the elements are larger than $\gamma + \epsilon$.
For each individual $\tilde{x}(i)$, the probability that $\tilde{x}(i) > \gamma + \epsilon$ is

$$p := 1 - F_{\tilde{x}(i)}(\gamma + \epsilon) = 1/2 - \epsilon' \tag{18}$$

where we assume that $\epsilon' > 0$. If this is not the case then the cdf of the iid sequence is flat at γ and the median is not well defined. The number of random variables in the set $\{\tilde{x}(1), \ldots, \tilde{x}(n)\}$ which are lager than $\gamma + \epsilon$ is distributed as a binomial random variable $\tilde{b}_n$ with parameters n and p. As a result, we have

$$\mathrm{P}\left(\tilde{y}(n) \geq \gamma + \epsilon\right) \leq \mathrm{P}\left(\frac{n+1}{2} \text{ or more samples are greater or equal to } \gamma + \epsilon\right) \tag{19}$$

$$= \mathrm{P}\left(\tilde{b}_n \geq \frac{n+1}{2}\right) \tag{20}$$

$$= \mathrm{P}\left(\tilde{b}_n - np \geq \frac{n+1}{2} - np\right) \tag{21}$$

$$\leq \mathrm{P}\left(|\tilde{b}_n - np| \geq n\epsilon' + \frac{1}{2}\right) \tag{22}$$

$$\leq \frac{Var\left(\tilde{b}_n\right)}{\left(n\epsilon' + \frac{1}{2}\right)^2} \text{ by Chebyshev's inequality} \tag{23}$$

$$= \frac{np(1-p)}{n^2\left(\epsilon' + \frac{1}{2n}\right)^2} \tag{24}$$

$$= \frac{p(1-p)}{n\left(\epsilon' + \frac{1}{2n}\right)^2} \tag{25}$$

which converges to zero as $n \to \infty$. This establishes (15) and therefore, sample median converges to the median of an iid sequence of random variables.