Appendix F: A Komlós-type lemma from probability theory

These notes provide a formulation and proof for an elementary lemma from probability theory which is extremely useful in many optimisation problems involving convexity. Recall that L^0 denotes the vector space of all (equivalence classes of, for the equivalence relation of equality P-a.s.) random variables on a given probability space $(\Omega, \mathcal{F}, P)$ and taking values in $\mathbb{R}$. For a sequence $(Y_n)_{n\in\mathbb{N}}$ in L^0, we denote for $m \in \mathbb{N}$ by $\operatorname{conv}(Y_m, Y_{m+1}, \dots)$ the set of all (finite) convex combinations of $(Y_k)_{k\geq m}$, i.e. all Y of the form $Y = \sum_{k=m}^{\infty} \lambda_k Y_k$ with the $\lambda_k \geq 0$ satisfying $\sum_{k=m}^{\infty} \lambda_k = 1$ and at most finally many $\lambda_k \neq 0$.

Lemma F.1. *For any sequence $(Y_n)_{n\in\mathbb{N}}$ of nonnegative random variables, there exists a sequence $(\widetilde{Y}_n)_{n\in\mathbb{N}}$ with $\widetilde{Y}_n \in \operatorname{conv}(Y_n, Y_{n+1}, \dots)$ for all n and $\widetilde{Y}_n \to Y$ P-a.s. for some random variable Y taking values in $[0, +\infty]$.*

Proof. It is clear that the sequence

$$\alpha_n := \inf\left\{E[e^{-Y}] \,\middle|\, Y \in \operatorname{conv}(Y_n, Y_{n+1}, \dots)\right\}$$

increases to some $\alpha \leq 1$. Take a sequence $(Y'_n)_{n\in\mathbb{N}}$ with $Y'_n \in \operatorname{conv}(Y_n, Y_{n+1}, \dots)$ and $E[e^{-Y'_n}] \leq \alpha_n + \frac{1}{n}$ for all n. For $\varepsilon > 0$, define the set

$$B_\varepsilon := \left\{(x, y) \in [0, \infty)^2 \,\middle|\, |x - y| \geq \varepsilon \text{ and } x \wedge y \leq \frac{1}{\varepsilon}\right\}$$

(a picture will help to illustrate the situation). Since the mapping $z \mapsto e^{-z}$ is convex, we always have

$$e^{-(x+y)/2} \leq \frac{1}{2}(e^{-x} + e^{-y}).$$

For $(x, y) \in B_\varepsilon$, a calculation gives

$$e^{-(x+y)/2} - \frac{1}{2}(e^{-x} + e^{-y}) \leq -\delta \qquad \text{for some } \delta = \delta(\varepsilon) > 0,$$

and therefore

$$e^{-(x+y)/2} \le \frac{1}{2}(e^{-x} + e^{-y}) - \delta I_{B_\varepsilon}(x,y).$$

Choosing $x := Y'_m$ and $y := Y'_n$ yields for $n \neq m$ that

$$\begin{aligned}
\alpha_m &\le E\left[e^{-(Y'_m+Y'_n)/2}\right] \\
&\le \frac{1}{2}\left(E\big[e^{-Y'_m}\big] + E\big[e^{-Y'_n}\big]\right) - \delta P[(Y'_m, Y'_n) \in B_\varepsilon] \\
&\le \frac{1}{2}\left(\alpha_m + \frac{1}{m} + \alpha_n + \frac{1}{n}\right) - \delta P[(Y'_m, Y'_n) \in B_\varepsilon],
\end{aligned}$$

and so we obtain that

$$\lim_{n,m\to\infty} P[(Y'_m, Y'_n) \in B_\varepsilon] = 0.$$

Considering the separate cases $|x-y| < \varepsilon$ or $x \wedge y > \frac{1}{\varepsilon}$ or $(x,y) \in B_\varepsilon$ leads to the estimate

$$|e^{-x} - e^{-y}| \le \varepsilon + 2e^{-1/\varepsilon} + 2I_{B_\varepsilon}(x,y).$$

This gives in turn that

$$\left|E\big[e^{-Y'_m} - e^{-Y'_n}\big]\right| \le \varepsilon + 2e^{-1/\varepsilon} + 2P[(Y'_m, Y'_n) \in B_\varepsilon]$$

so that $(e^{-Y'_n})_{n\in\mathbb{N}}$ is a Cauchy sequence in $L^1(P)$ and hence convergent in $L^1(P)$. Therefore this sequence has a subsequence $(e^{-\widetilde{Y}_n})_{n\in\mathbb{N}}$ which converges P-a.s., and then the sequence $(\widetilde{Y}_n)_{n\in\mathbb{N}}$ is also P-a.s. convergent and has $\widetilde{Y}_n \in \mathrm{conv}(Y_n, Y_{n+1}, \dots)$ like for Y'_n. **q.e.d.**

Remark. If one has extra properties for the original sequence $(Y_n)_{n\in\mathbb{N}}$, one can also say more about the limit Y. For example, if all the Y_n are bounded by some constant, the same is true for the $\widetilde{Y}_n$ and hence also for Y.