3.2. Alarajat

Olkoot prosessit $\{Y_n\}$, $\{\bar{X}_n\}$ ja funktiot Λ, Λ^* kuten lauseessa 3.1. Seuraavassa Λ:ista oletetaan aina riittävästi sen takaamiseksi, että $\Lambda(t) > -\infty$, $\forall t$ (sovelletaan lemmaa 3.1).

Lause 3.2. Oletetaan, että $0 \in \mathring{D}(\Lambda)$ ja että $\nabla\Lambda(0)$ on olemassa. Silloin

$$\bar{X}_n \xrightarrow{\exp} \nabla\Lambda(0), \quad n \to \infty.$$

Todistus. Olkoon $\varepsilon > 0$. Lauseen 3.1 nojalla Λ^* on hyvä vf ja

$$\limsup n^{-1} \log \mathbb{P}(\bar{X}_n \in B(\nabla\Lambda(0), \varepsilon)^c)$$

$$\leq -\inf\{\Lambda^*(x) \mid |x - \nabla\Lambda(0)| \geq \varepsilon\}.$$

Koska $\Lambda^*(x) \to \infty$, kun $|x| \to \infty$ (sillä Λ^*:in tasojoukot ovat kompakteja), voidaan minimointi edellä rajoittaa eräiseen kompaktiin joukkoon. Koska Λ^* on api, voidaan määrätä $x_0 \in \mathbb{R}^d$, $|x_0 - \nabla\Lambda(0)| \geq \varepsilon$:

$$\Lambda^*(x_0) = \inf\{\Lambda^*(x) \mid |x - \nabla\Lambda(0)| \geq \varepsilon\}.$$

Lemman 3.12 nojalla (otetaan $\ell(t) = \Lambda(t) - \langle t, \nabla\Lambda(0)\rangle$)

$$\Lambda^*(x_0) > \Lambda^*(\Lambda'(0)) = 0.$$

Nähdään, että

$$\mathbb{P}(|\bar{X}_n - \nabla\Lambda(0)| \geq \varepsilon) \leq e^{-n\Lambda^*(x_0)/2},$$

kun n on riittävän suuri (jos $\Lambda^*(x_0) = +\infty$, saadaan ylärajaksi $e^{-n\alpha}$, $\alpha > 0$ mielivaltainen). $\square$

Lauseen 3.2 tulos on ylärajatulos, kuten todistuksesta nähdään. Siitä seuraa myös yksinkertainen alarajatulos:

olkoon $G \subseteq \mathbb{R}^d$ avoin ja $\nabla\Lambda(0) \in G$. Silloin

$$\liminf_{n} n^{-1} \log \mathbb{P}(\bar{X}_n \in G) \geq - \inf_{x \in G} \Lambda^*(x).$$

Nimittäin vasen puoli on $= 0$ lauseen nojalla ja

$$\Lambda^*(\Lambda'(0)) = \langle 0, \Lambda(0) \rangle - \Lambda(0) = 0,$$

kts. seuraus 3.1. Siis myös oikea puoli on $= 0$.

<u>Määritelmä 3.1.</u> Olkoon ξ $\mathbb{R}^d$-arvoinen satunnaisvektori ja

$$f(t) = \log \mathbb{E}\{e^{\langle t, \xi \rangle}\}, \quad t \in \mathbb{R}^d.$$

Jos $f(t) < \infty$, niin <u>ξ:n konjugaattijakauma P_t parametrilla t</u> määritellään ehdosta

$$P_t(B) = \mathbb{E}(e^{\langle t, \xi \rangle - f(t)} 1(\xi \in B)), \quad B \in \mathcal{B}^d.$$

Jos ξ:n alkuperäinen jakauma on P, niin

$$\frac{dP_t}{dP}(x) = e^{\langle t, x \rangle - f(t)}, \quad x \in \mathbb{R}^d.$$

(Radon-Nikodymin derivaatta). Vastaavia odotusarvoja merkitään $\mathbb{E}_t$:llä,

$$\mathbb{E}_t(h(\xi)) = \int_{\mathbb{R}^d} h(x)\, dP_t(x)$$

$$= \int_{\mathbb{R}^d} h(x) e^{\langle t, x \rangle - f(t)} dP(x) = \mathbb{E}(h(\xi) e^{\langle t, \xi \rangle - f(t)}).$$

Seuraavassa tulee olemaan $\Lambda(t_0) < \infty$ eräälle kiinteälle $t_0 \in \mathbb{R}^d$. Tällöin

$$\Lambda_n(t_0) = n^{-1} \log \mathbb{E}\{ e^{\langle t_0, Y_n \rangle} \} < \infty,$$

kun n on suuri. Voidaan siis tehdä konjugaattimuunnoksia Y_n:ille. Käytetään tällöin merkintöjä $\mathbb{P}_{t_0}$, $\mathbb{E}_{t_0}$ (Y_n:it saadaan luonnollisesti määriteltyä samassa tn-kentässä jakaumamuunnoksen jälkeenkin). Käytetään myös merkintöjä Λ_{n,t_0}, Λ_{t_0}, $\Lambda^*_{t_0}$. Esimerkiksi

$$(3.14)\quad \Lambda_{n,t_0}(t) = n^{-1} \log \mathbb{E}_{t_0}\{ e^{\langle t, Y_n \rangle} \}$$

$$= n^{-1} \log \mathbb{E}\{ e^{\langle t, Y_n \rangle} e^{\langle t_0, Y_n \rangle - n\Lambda_n(t_0)} \}$$

$$= \Lambda_n(t + t_0) - \Lambda_n(t_0), \quad t \in \mathbb{R}^d.$$

Lause 3.3. Olkoon $t_0 \in \mathring{D}(\Lambda)$ sellainen, että $\Lambda(t_0) \in \mathbb{R}$ ja $\nabla\Lambda(t_0)$ on olemassa. Oletetaan, että on olemassa raja-arvo

$$(3.15)\quad \Lambda(t_0) = \lim \Lambda_n(t_0).$$

Olkoon $G \subseteq \mathbb{R}^d$ avoin ja $\nabla\Lambda(t_0) \in G$. Silloin

$$\liminf n^{-1} \log \mathbb{P}(X_n \in G) \geq -\Lambda^*(\nabla\Lambda(t_0)).$$

Todistus. Valitaan $\varepsilon > 0$: $B(\nabla\Lambda(t_0), \varepsilon) \subseteq G$. Selvästi

$$\mathbb{P}(\bar{X}_n \in G) = \mathbb{P}(Y_n \in nG)$$

$$= \mathbb{E}_{t_0}\{ e^{-\langle t_0, Y_n\rangle + n\Lambda_n(t_0)} \mathbb{1}(Y_n \in nG)\}$$

$$\geq \mathbb{E}_{t_0}\{ e^{-\langle t_0, Y_n\rangle + n\Lambda_n(t_0)} \mathbb{1}(\tfrac{Y_n}{n} \in B(\nabla\Lambda(t_0), \varepsilon))\}$$

$$\geq e^{-\langle t_0, \nabla\Lambda(t_0)\rangle n - \varepsilon|t_0|n + n\Lambda_n(t_0)} \mathbb{P}_{t_0}(\bar{X}_n \in B(\nabla\Lambda(t_0), \varepsilon))$$

Osoitetaan, että

$$(3.16) \qquad \mathbb{P}_{t_0}(\bar{X}_n \in B(\nabla\Lambda(t_0), \varepsilon)) \to 1, \quad n \to \infty.$$

Sovelletaan lausetta 3.2. Raja-arvosta (3.15) ja yhteydestä (3.14) seuraa, että

$$\Lambda_{t_0}(t) = \Lambda(t + t_0) - \Lambda(t_0), \quad \forall t \in \mathbb{R}^d.$$

Nähdään, että $\nabla\Lambda_{t_0}(0) = \nabla\Lambda(t_0)$, joten (3.16) seuraa. Näin ollen

$$\liminf n^{-1} \log \mathbb{P}(\bar{X}_n \in G)$$

$$\geq -\langle t_0, \nabla\Lambda(t_0)\rangle + \Lambda(t_0) - \varepsilon|t_0|$$

$$= -\Lambda^*(\nabla\Lambda(t_0)) - \varepsilon|t_0|$$

(seuraus 3.1). $\square$

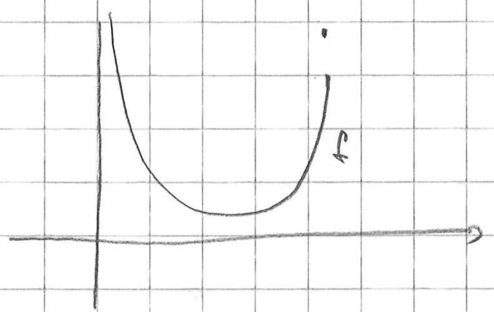
<u>Määritelmä 3.2.</u> Konveksi funktio $f: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ on <u>sileä</u> (essentially smooth), jos

(i) $\mathring{D}(f) \neq \emptyset$

(ii) $\nabla f(t)$ on olemassa $\forall t \in \mathring{D}(f)$

(iii) $\lim |\nabla f(t_n)| = +\infty$, aina kun $t_n \in \mathring{D}(f)$, $\forall n$, ja t_n suppenee kohti $D(f)$:n reunapistettä.

Erityisesti siis reaaliarvoinen kaikkialla differentioituva konveksi funktio on sileä.



Merkitään

$$\mathcal{H} = \{\nabla\Lambda(t) \mid t \in \mathring{D}(\Lambda),\ \nabla\Lambda(t)\ \exists,\ \lim \Lambda_n(t) = \Lambda(t)\}.$$

<u>Lause 3.4.</u> (Gärtner-Ellisin lause, alarajat). Olkoon $G \subseteq \mathbb{R}^d$ avoin. Silloin

(3.17) $$\liminf n^{-1} \log \mathbb{P}(Z_n \in G) \geq -\inf\{\Lambda^*(x) \mid x \in G \cap \mathcal{H}\}.$$

Jos Λ on sileä ja $\lim \Lambda_n(t) = \Lambda(t)$ kaikilla $t \in \mathring{D}(\Lambda)$, niin

(3.18) $$\liminf n^{-1} \log \mathbb{P}(Z_n \in G) \geq -\inf\{\Lambda^*(x) \mid x \in G\}.$$

Ensimmäinen tulos seuraa suoraan lauseesta 3.3 ottamalla supremum yli kysymykseen tulevien pisteiden $\in \mathcal{H}$.

Summien poikkeamien periaate saadaan sopivin oletuksin yhdistämällä lauseet 3.1 ja 3.4.

Lauseen todistus on pelkkää konveksia analyysiä.

Lemma 3.13. Olkoon $f: \mathbb{R}^d \to [0, +\infty]$ konveksi ja aps sekä

$$\inf\{f(t) \mid t \in \mathbb{R}^d\} = 0.$$

Jos $0 \in \operatorname{ri} D(f^*)$, niin $f(t_0) = 0$ eräälle $t_0 \in \mathbb{R}^d$.

Todistus. Ensinnäkin

$$f^*(0) = -\inf\{f(t) \mid t \in \mathbb{R}^d\} = 0.$$

Jos $y \in \mathbb{R}^d$ ja $0 < \delta_1 < \delta_2$, niin

$$\frac{f^*(\delta_1 y)}{\delta_1} = \frac{f^*\left(\left(1 - \frac{\delta_1}{\delta_2}\right)0 + \frac{\delta_1}{\delta_2}(\delta_2 y)\right)}{\delta_1}$$

$$\leq \frac{\frac{\delta_1}{\delta_2} f^*(\delta_2 y)}{\delta_1} = \frac{f^*(\delta_2 y)}{\delta_2}.$$

Voidaan siis määritellä funktio $g: \mathbb{R}^d \to \overline{\mathbb{R}}$ ehdosta

$$g(x) = \lim_{\delta \to 0+} \frac{f^*(\delta x)}{\delta}.$$

Selvästi

$$g(\alpha x) = \alpha g(x), \ \forall x \in \mathbb{R}^d, \ \alpha > 0.$$

Olkoon $x \neq 0$. Jos $f^*(\delta x) = \infty$, $\forall \delta > 0$, niin $g(x) = \infty$.
Jos taas $f^*(\delta_0 x) < \infty$ jollain $\delta_0 > 0$, niin

$$-\varepsilon x \in D(f^*) \text{ eräälle } \varepsilon > 0$$

(sillä $0 \in \text{ri}\, D(f^*)$, sovelletaan lemmaa 3.5, (ii)).

Olkoon $\lambda \in \mathbb{R}$:

$$0 = (1-\lambda)\delta + \lambda(-\varepsilon), \ \delta > 0,$$

ts. $\lambda = \frac{\delta}{\varepsilon + \delta} \in (0,1)$. Tällöin

$$0 = f^*(0) \leq (1-\lambda) f^*(\delta x) + \lambda f^*(-\varepsilon x)$$
$$= \frac{\varepsilon}{\varepsilon + \delta} f^*(\delta x) + \frac{\delta}{\varepsilon + \delta} f^*(-\varepsilon x),$$

joten

$$\frac{f^*(\delta x)}{\delta} + \frac{f^*(-\varepsilon x)}{\varepsilon} \geq 0.$$

Nähdään, että $g(x) > -\infty$. Koska

$$x \mapsto \frac{f^*(\delta x)}{\delta}$$

on konveksi ja $g(x) > -\infty$, $\forall x$, on myös g konveksi (Lemma 3.1). Lemman 3.4 nojalla myös $\underline{g}(x) > -\infty$, $\forall x$, joten $\underline{g}$ on konveksi. Lemman 3.9 nojalla voidaan määrätä affiini kuvaus h, jolle

$$h(x) = \langle b, x \rangle + a, \quad \underline{g}(x) \geq h(x), \ \forall x \in \mathbb{R}^d.$$

Siis myös $g(x) \geq h(x)$, $\forall x$.

Mielivaltaisella $\alpha > 0$, $x \in \mathbb{R}^d$,

$$\alpha g(x) = g(\alpha x) \geq \langle b, \alpha x \rangle + a$$

eli $g(x) \geq \langle b, x \rangle + \frac{a}{\alpha}$. Nähdään, että

$$g(x) \geq \langle b, x \rangle, \quad \forall x \in \mathbb{R}^d.$$

Edelleen

$$f^*(x) = \frac{f^*(x)}{1} \geq g(x) \geq \langle b, x \rangle, \quad \forall x \in \mathbb{R}^d.$$

Koska f on konveksi ja asj, niin lemman 3.10 nojalla

$$f(b) = \sup_{x \in \mathbb{R}^d} \{ \langle b, x \rangle - f^*(x) \} = 0$$

(koska $f(t) \geq 0$, $\forall t$, oletuksen nojalla). $\square$

Lemma 3.14. Olkoon $f: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ konveksi ja sileä. Jos $f(0) = 0$ ja $f^*(x) = 0$ jollain $x \in \mathbb{R}^d$, niin $0 \in \mathring{D}(f)$.

Todistus. Jos $\lambda \in (0,1)$, niin

$$f(\lambda t) \le (1-\lambda) f(0) + \lambda f(t) = \lambda f(t), \quad \forall t \in \mathbb{R}^d.$$

Koska $f^*(x) = 0$, niin

$$0 = f^*(x) \ge \langle \lambda t, x \rangle - f(\lambda t) \ge -\lambda |t|\,|x| - f(\lambda t)$$

joten

$$f(\lambda t) \ge -\lambda |t|\,|x|, \quad \forall \lambda \in (0,1).$$

Olkoon $t_0 \in \mathring{D}(f)$ ja $\delta > 0$: $\bar{B}(t_0, \delta) \subseteq \mathring{D}(f)$ ja

$$M = \max\{ f(t) \vee |t|\,|x| \mid t \in \bar{B}(t_0, \delta)\}.$$

Jos $\lambda \in (0,1)$, niin edellä esitetyn nojalla

$$\sup\{ |f(u)| \mid u \in B(\lambda t_0, \lambda\delta)\} \le \lambda M.$$

Olkoon $u \ne \lambda t_0$, $u \in B(\lambda t_0, \lambda \delta)$. Kirjoitetaan

$$\begin{cases} u = \mu v + (1-\mu)\lambda t_0, & \mu = \dfrac{|u - \lambda t_0|}{\lambda \delta} \in (0,1), \\ v = \lambda t_0 + \dfrac{\lambda\delta}{|u - \lambda t_0|}(u - \lambda t_0). \end{cases}$$

Siis

$$f(u) - f(\lambda t_0) \le \frac{|u-\lambda t_0|}{\lambda \delta} f(v) + \left(1 - \frac{|u-\lambda t_0|}{\lambda\delta}\right) f(\lambda t_0) - f(\lambda t_0)$$

$$= \frac{|u-\lambda t_0|}{\lambda\delta}\left(f(v) - f(\lambda t_0)\right) \le \frac{2M}{\delta} |u - \lambda t_0|,$$

sillä myös $v \in \bar{B}(\lambda t_0, \lambda \delta)$.

Samoin nähdään, että

$$f(\lambda t_0) - f(u) \le \frac{2M}{\delta} |u - \lambda t_0|.$$

Koska $\nabla f(\lambda t_0)$ on olemassa, ja edellä todetun nojalla

$$|\nabla f(\lambda t_0)| \le \frac{2M}{\delta},$$

ei 0 voi olla $D(f)$:n reunapiste ($\lambda t_0 \to 0$, kun $\lambda \to 0$, f on sileä). $\square$

Lemma 3.15. Olkoon $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ apj, konveksi ja sileä. Silloin

$$\mathrm{ri}\, \mathcal{D}(f^*) \subseteq \{ \nabla f(t) \mid t \in \mathring{\mathcal{D}}(f) \}.$$

Todistus. Olkoon $x \in \mathrm{ri}\, \mathcal{D}(f^*)$ ja

$$g(t) = f(t) - \langle t, x \rangle + f^*(x), \quad t \in \mathbb{R}^d.$$

Tällöin myös g on apj, konveksi ja sileä. Lisäksi

$$g(t) \geq 0, \ \forall t, \quad \text{ja} \quad \inf \{ g(t) \mid t \in \mathbb{R}^d \} = 0.$$

Jos $y \in \mathbb{R}^d$, niin

$$g^*(y) = \sup_{t \in \mathbb{R}^d} \{ \langle t, x+y \rangle - f(t) \} + f^*(x)$$

$$= f^*(x+y) - f^*(x).$$

Koska $x \in \mathrm{ri}\, \mathcal{D}(f^*)$, niin $0 \in \mathrm{ri}\, \mathcal{D}(g^*)$. Lemman 3.13 nojalla

$$g(t_0) = 0 \quad \text{eräälle} \quad t_0 \in \mathbb{R}^d.$$

Olkoon

$$h(t) = f(t+t_0) - f(t_0), \quad t \in \mathbb{R}^d.$$

Myös h on apj, konveksi ja sileä. Lisäksi $h(0) = 0$ ja

$$h^*(x) = \sup_{t \in \mathbb{R}^d} \{ \langle t + t_0, x \rangle - f(t+t_0) \} - \langle t_0, x \rangle + f(t_0)$$

$$= f^*(x) + g(t_0) - f^*(x) = 0.$$

Lemman 3.14 nojalla $0 \in \mathring{\mathcal{D}}(h)$. Nähdään, että

$$t_0 \in \mathring{\mathcal{D}}(f), \quad t_0 \in \mathring{\mathcal{D}}(g).$$

Nyt g on differentioituva $\mathring{\mathcal{D}}(g)$:ssä ja saavuttaa globaalin miniminsä pisteessä $t_0 \in \mathring{\mathcal{D}}(g)$. Välttämättä $\nabla g(t_0) = 0$, joten $x = \nabla f(t_0)$. $\square$

Lauseen 3.4 todistus. Alaraja (3.12) seuraa suoraan Lauseesta 3.3. Jälkimmäisen väitteen todistamiseksi todetaan ensin, että $\Lambda(t) > -\infty$, $\forall t \in \mathbb{R}^d$ Lemman 3.1 nojalla (differentioituvuus $\mathring{\mathcal{D}}(\Lambda)$:ssa merkitsee, että $\Lambda(t) \in \mathbb{R}$, $\forall t \in \mathring{\mathcal{D}}(\Lambda)$). Lemmojen 3.6 ja 3.8 nojalla $\underline{\Lambda}(t) > -\infty$, $\forall t$, ja $\underline{\Lambda}$ on konveksi ja apj. Lisäksi

$$\underline{\Lambda}(t) = \Lambda(t), \quad \forall t \in \mathring{\mathcal{D}}(\Lambda) \cup (\bar{\mathcal{D}}(\Lambda))^c$$

(lemma 3.8). Nähdään, että $\underline{\Lambda}$ on sileä. Lemman 3.15 nojalla

$$\begin{aligned} \operatorname{ri} \mathcal{D}(\underline{\Lambda}^*) &\subseteq \{ \nabla \underline{\Lambda}(t) \mid t \in \mathring{\mathcal{D}}(\underline{\Lambda}) \} \\ &= \{ \nabla \Lambda(t) \mid t \in \mathring{\mathcal{D}}(\Lambda) \}. \end{aligned}$$

Lemman 3.10 nojalla $\underline{\Lambda}^* = \Lambda^*$, joten

$$\operatorname{ri} \mathcal{D}(\Lambda^*) \subseteq \{ \nabla \Lambda(t) \mid t \in \mathring{\mathcal{D}}(\Lambda) \} = \mathcal{H}.$$

Alarajan (3.12) nojalla mielivaltaiselle avoimelle $G \subseteq \mathbb{R}^d$,

$$\begin{aligned} \liminf_n n^{-1} \log \mathbb{P}(\bar{Z}_n \in G) &\geq -\inf \{ \Lambda^*(x) \mid x \in G \cap \mathcal{H} \} \\ &\geq -\inf \{ \Lambda^*(x) \mid x \in G \cap \operatorname{ri} \mathcal{D}(\Lambda^*) \}. \end{aligned}$$

Lause on todistettu, jos voidaan näyttää, että

$$\inf\{\Lambda^*(x) \mid x \in G \cap \operatorname{ri} D(\Lambda^*)\}$$

$$\leq \inf\{\Lambda^*(x) \mid x \in G\}.$$

Olkoon $y \in G$ mielivaltainen. Riittää näyttää, että

(3.19) $$\inf\{\Lambda^*(x) \mid x \in G \cap \operatorname{ri} D(\Lambda^*)\} \leq \Lambda^*(y),$$

Voidaan olettaa, että $y \in D(\Lambda^*)$. Olkoon $x_0 \in \operatorname{ri} D(\Lambda^*)$ kiinteä. Tällöin

$$(1-\lambda)x_0 + \lambda y \in \operatorname{ri} D(\Lambda^*), \quad \forall \lambda \in [0,1),$$

Lemma 3.5, (iii). Kun λ on riittävän lähellä ykköstä, on

$$(1-\lambda)x_0 + \lambda y \in G \cap \operatorname{ri} D(\Lambda^*).$$

Saadaan

$$\inf\{\Lambda^*(x) \mid x \in G \cap \operatorname{ri} D(\Lambda^*)\}$$

$$\leq \liminf_{\lambda \to 1-} \Lambda^*((1-\lambda)x_0 + \lambda y)$$

$$\leq \liminf_{\lambda \to 1-} ((1-\lambda)\Lambda^*(x_0) + \lambda \Lambda^*(y)) = \Lambda^*(y). \quad \square$$