High-Dimensional Measures and Geometry

Lecture Notes from Feb 11, 2010

taken by Nick Maxwell

We will rexamine the application of the martingale technique, with $\{f_0, f_1, ..., f_n\}$.

Let f_k depend only on the first k coordinates, so

$$f_k(x_1, x_2, ..., x_k, x_{k+1}, ..., x_n) = \int f_k(x_1, x_2, ..., x_k, x'_1, x'_2, ..., x'_{n-k}) d\mu_{n-k}(x') =$$

$$\int f_k(x_1, x_2, ..., x_k, x'_1, x'_2, ..., x'_{n-k}) d\mu_{n-k}(x') = \sum_{i=1}^{k} x_i + (n-k)p$$

Computing $E_{n-k}[e^{\lambda g_k}]$, $g_k = f_k - f_{k-1}$, instead of estimating, gives, by

$$g_k(x_1, ..., x_n) = f_k(x_1, ..., x_n) - f_{k-1}(x_1, ..., x_n) = x_k - p.$$

that

$$e^{\lambda g_k} = \begin{cases} e^{\lambda(1-p)}, & x_k = 1 \\ e^{-\lambda p}, & x_k = 0 \end{cases}$$

which gives,

$$E_{k-1}[e^{\lambda g_k}] = pe^{\lambda(1-p)} + (1-p)e^{-\lambda p}.$$

Iterating as before, n times,

$$E[e^{\lambda(f-a)}] = \left(pe^{\lambda(1-p)} + (1-p)e^{-\lambda p}\right)^n.$$

Now using the Laplace transform method,

$$\mu_n(\{x \in I_n; f(x) - np \geq t\}) \leq e^{-\lambda t}\left(pe^{\lambda(1-p)} + (1-p)e^{-\lambda p}\right)^n.$$

Also, switching $f \to -f$, $np \to -np$ gives

$$\mu_n(\{x \in I_n; f(x) - np \leq t\}) \leq e^{-\lambda t}\left(pe^{-\lambda(1-p)} + (1-p)e^{\lambda p}\right)^n.$$

Choosing least λ gives,

$$\frac{t}{np(1-p)} = 1 - e^{-\lambda} \Rightarrow \lambda = -\log\left(1 - \frac{t}{np(1-p)}\right).$$

If we now fix $np = \beta^2$, $p \to 0$, and as $t = \alpha\sqrt{np(1-p)} \to \alpha\beta$ and $\lambda \to -\log(a - \frac{\alpha}{\beta})$, inserting this RHS estimate gives that

$$\mu_n(\{x \in I_n; f(x) - np \geq t\}) \leq e^{-\lambda t}\left((1-p)e^{\lambda p} + pe^{-\lambda(1-p)}\right)^n.$$

Denote the RHS of this inequality as " $e^{-\lambda t} *$ ".

Consider $*$ as $p \to 0$, $n \to \infty$, $p = \beta^2/n$

$$\left(pe^{\lambda(1-p)} + (1-p)e^{-\lambda p}\right)^n = \left(pe^{-\log(1-\alpha/\beta)} + 1 - p + \log(1-\alpha/\beta)p + C_n p^2\right)^n,$$

where C_n stays constnt in n. So,

$$* \to e^{\beta^2}(e^{-\log(1-\alpha/\beta)} + \log(1-\alpha/\beta))$$

now, for small α_k and large n,

$$* \approx e^{\beta^2(\log(1-\alpha/\beta))^2/2} \approx e^{\alpha^2/2}$$

together with $e^{\alpha\beta\lim(1-\alpha/\beta)} \approx e^{-\alpha^2}$. So, RHS is $\approx e^{-\alpha^2/2}$.

5.4 General result in product spaces

5.4.1 Definition. *If $g : \mathbb{R} \to \mathbb{R}$ is convex, then define its Legendre transform, g^*, by $g^*(x) = \sup_{\lambda\in\mathbb{R}}\{t\lambda - g(\lambda)\}$. Note, $(t\lambda - g(\lambda))$ is concave. If $g \in C^2(\mathbb{R})$ and g is strictly convex, then this supremum is attained, and λ^* solves $g'(\lambda^*) = t$ uniquely. So, λt and $g(\lambda)$ have same slope at λ^*.*

ADD diagram.

Examples: $g(\lambda) = \lambda^2$, $g^*(t) = t^2/4$

We will apply the Legendre transform to the function $L_f(\lambda) = \log\int_X e^{\lambda f}d\mu$. And by Jensen's inequality, and by convexity of $\exp$, $\int_X e^{\lambda f}d\mu \geq \exp\left(\lambda\int_X f d\mu\right)$. So, $\log\int_X e^{\lambda f} \geq \log\exp\lambda\int f d\mu = \lambda E[f] \in \mathbb{R}$, which is bounded below in the vicinity of $\lambda = 0$.

Also, assuming existence,

$$L_f'(\lambda) = \frac{E[fe^{\lambda f})}{E(e^{\lambda f}]}$$

and,

$$L_f''(\lambda) = \frac{E[f^2e^{\lambda f}]E[e^{\lambda f}] - (E[fe^{\lambda f}])^2}{(E[e^{\lambda f}])^2},$$

using Cauchy-Schwartz,

$$(E[fe^{\lambda f}])^2 \leq E[f^2e^{\lambda f}]E[e^{\lambda f}],$$

so $L''_f(\lambda) \geq 0$, meaning that L'_f is convex, so it is in the domain of the Legendre transformation.

5.4.2 Theorem (Cramér). *Let (X, μ) be a probability space, $f : \mathbb{R} \to \mathbb{R}$ and $t \in \mathbb{R}$ $s.t.$*

1) $E[f] = a_f$

2) $L_f(\lambda) = \log E[e^{\lambda f}]$ is finite near $\lambda = 0$

3) $t > a_f$ and $\mu(\{x; f(c) > t\}) > 0$

Let $X_n = \prod_{k=1}^{n} X$, $\mu_n = \mu \times \mu \times ... \times \mu$, and let $h : X_n \to \mathbb{R}$, $h(x) = f(x_1) + f(x_2) + ... + f(x_n)$. Then,

$$\lim_{n\to\infty} \frac{1}{n} \log\left(\mu\left(\{x \in X_n; h(x) > nt\}\right)\right) = -L_f^*(t),$$

where $L_f(\lambda) = \log \int_X e^{\lambda f} d\mu$. Moreover, for all $n \in \mathbb{N}$,

$$\mu\left(\{x \in X_n; h(x) > nt\}\right) \leq e^{-nL_f^*(t)}.$$

Proof. Only inequality part.

$$\mu_n \{x \in X_n; h(x) > nt\}) \leq e^{-nt\lambda} E[e^{\lambda h}] = \left(e^{-t\lambda} E_{X_1}[e^{\lambda f}]\right)^n = \left(e^{-t\lambda} e^{\log E[e^{\lambda f}]}\right)^n = e^{-n(t\lambda - L_f(\lambda))}$$

Optimizing this with respect to λ gives

$$\mu_n \{x \in X_n; h(x) > nt\}) \leq e^{-nL_f^*(t)}$$

□