Definition. A sequence \( \{ \theta_n \} \subset \Theta \) converges to \( \Theta_0 \)

\[
\theta_n \rightarrow \Theta_0
\]

if for every subsequence \( \{ \theta_{n_k} \} \) we have \( \theta_{n_k} \rightarrow \Theta^* \) or \( \| \theta_{n_k} \| \to \infty \) as \( k \to \infty \)

Note. It appears we could just define \( \theta_n \to \Theta_0 \) to mean that either \( \theta_n \to \Theta^* \) or \( \| \theta_n \| \to \infty \) as \( n \to \infty \).

We won't say BAD talk about subsequences.

Example. \( N(\mu, \sigma^2) \) distribution \( \Theta = \{ (\mu, \sigma^2) : \mu \in \mathbb{R}, \sigma^2 > 0 \} \)

Lemma. If \( l : \Theta \to \mathbb{R} \) is continuous, \( \Theta \subset \mathbb{R}^d \) open

\[
\lim_{n \to \infty} l(\theta_n) = -\infty
\]

then there exists \( \theta^* \in \Theta \) (the interior of \( \Theta \)) where \( l \)

is maximized.

Proof. Fix \( \theta^* \in \Theta \) and set \( c_0 = l(\theta^*) \)

Define \( B(K) = \sup_{\theta \in \Theta : \| \theta \| \geq K} l(\theta) \)

\[ K = 1, 2, 3, \ldots \]

Note that \( B(1) \geq B(2) \geq \ldots \)

so either

(A) \( B(K) \to -\infty \) as \( K \to \infty \)

or

(B) there exists \( \Theta^* \) s.t. \( B(K) = \Theta^* \) for all \( K \).

But if \( B \) holds, there exist \( \theta, \theta_2, \ldots \) such that

\[
l(\theta_n) > \Theta^* \quad \text{and} \quad \| \theta_n \| \geq K = 1, 2, \ldots
\]

we would have \( \{ \theta_n \} \to \Theta^* \) but \( l(\theta_n) \not\to -\infty \), a contradiction.
Thus (A) holds. Now pick $K^*$ so that $D(K^*) < c_0$.

Then $\sup_{\theta \in \Theta} l(\theta) = D(K^*) < c_0$.

We can decompose $\Theta$ into two pieces:

$\Theta_0 = \Theta \cap \left\{ \theta : \|\theta\| \leq K^* \right\}$

$\Theta_1 = \Theta \cap \left\{ \theta : \|\theta\| > K^* \right\}$

with $l(\theta) < c_0$ for $\theta \in \Theta_1$,

and $l(\theta) = c_0$ for some $\theta \in \Theta_0$, so $\Theta \subseteq \Theta_0$.

The set $\Theta_0$ is closed and bounded and is hence compact.

If $l$ is not maximized in $\Theta_0$ there must exist an infinite sequence of points in $\Theta_0$ say $\theta_1, \theta_2, \theta_3, \ldots$ with

$l(\theta_k) \rightarrow \sup_{\theta \in \Theta_0} l(\theta)$

Since $\Theta_0$ is compact, there exists a subsequence $\theta_{k_1}, k_1, k_2, \ldots$ with $\theta_{k_n} \rightarrow \theta^* \in \Theta_0$. It cannot be the case that $\theta^* \in \Theta_0$ since we would then have $l(\theta_1) \rightarrow -\infty$. Thus, $\theta^* \notin \Theta_0$,

and we get the desired conclusion. □

Proposition.

Suppose $X \sim P \in \left\{ P_\theta, \theta \in \Theta \right\}$, $\Theta \subseteq \mathbb{R}^d$ open

where $P_\theta$ has density $p_{\theta}(x)$. If $\log p_{\theta}(x)$ is strictly concave in $\Theta$ and $l(\theta) \rightarrow -\infty$ as $\theta \rightarrow \partial \Theta$.

The MLE exists and is unique.

Proof.

Concavity in an open set implies continuity.

So by the previous Lemma, $\hat{\theta}_{MLE}$ exists.

For uniqueness, $\hat{\theta}_1, \hat{\theta}_2$ are MLE for $i = 1, 2$.

We have

$l(\hat{\theta}_1) = l(\hat{\theta}_2)$

Let $\tilde{\theta}_t = (1-t)\hat{\theta}_1 + t\hat{\theta}_2$. Since $\Theta$ is open there exists $\lambda \in (0, 1)$

such that $\tilde{\theta}_t \in \Theta$. By strict concavity we have

$l(\tilde{\theta}_t) > \lambda l(\hat{\theta}_1) + (1-\lambda) l(\hat{\theta}_2) = l(\hat{\theta}_1)$

a contradiction. □

Theorem.

Let $\Theta$ be the canonical exponential family defined

by $\Theta^n \left( T = (T_1, \ldots, T_n)^T \right)$ and suppose that

(i) $\mathcal{E}_k = \left\{ \mu \in \mathbb{R}^k : \int e^{y^T \mu} h(x) \, dx < +\infty \right\}$ is open.
(ii) the family has rank $k$.
Suppose we observe $X = x_0$ and set $t_0 = T(x_0)$.
(a) If
\[ p[c^T T(x) > c^T t_0] > 0 \quad \text{for all } c \in \mathbb{R}^k, c \neq 0 \]
then the MLE exists and is unique and satisfies
\[ \hat{A}(\eta) = \text{Eq } T(x) = t_0. \]
(b) If (a) fails then the MLE does not exist and
(a) has no solution.

Proof:
Fix $\eta_0$ in $\mathcal{F}$.
We can assume without loss of generality that $h(\eta_0) = p(x_0, \eta_0)$.
To see this, write
\[
    p(\eta|y) = \exp \left\{ \eta^T T(x) - A(\eta) \right\} \cdot h(\eta)
    = \exp \left\{ \eta^T T(x) - A(\eta) \right\} \frac{p(x_0, \eta_0)}{\exp \left\{ \eta^T T(x) - A(\eta_0) \right\}}
    = \exp \left\{ (\eta - \eta_0)^T T(x) - (A(\eta) - A(\eta_0)) \right\} \cdot h(\eta_0)
    = \exp \left\{ \eta^T t_0 - A(\eta) \right\} \cdot h(\eta_0).
\]

In addition, we can replace $T(x)$ by $T(x - t_0)$ and
without loss of generality set $t_0 = 0$.

Note that we then have
\[
    l(x|\eta) = \log p(x|\eta) = \log \left[ \exp \left\{ \eta^T t_0 - A(\eta) \right\} \cdot h(\eta_0) \right]
    = \eta^T t_0 - A(\eta) + \log h(\eta_0).
\]

Now suppose $\{\eta_n\} \subseteq \mathcal{F}$ and $\eta_n \rightarrow \partial \mathcal{F}$.
We proceed to prove that $l(x|\eta_n) \rightarrow -\infty$.  