1 Cameron-Martin theory

We take a Gaussian measure in a separable Banach space X (or a separable Fréchet space). Recall that this means that for any linear functional $f \in X^*$, the evaluation $f(x)$ is a normal random variable on $\mathbb{R}$. We can define a mean and a covariance by setting

- $a_\gamma(f) = \int_X f(x)\gamma(\mathrm{d}x)$, so $a_\gamma : X^* \to \mathbb{R}$, and hence $a_\gamma \in (X^*)'$.
- $B_\gamma(f,g) = \int_X [f(x) - a_\gamma(f)] \cdot [g(x) - a_\gamma(g)]\gamma(\mathrm{d}x)$, so $B_\gamma \in (X^* \times X^*)'$.

So far we only know that a_γ is a linear map from X^* (the so-called topological dual) to $\mathbb{R}$, i.e. a_γ is an element of the "algebraical dual of the topological dual", because we don't know yet whether a_γ is continuous (i.e. bounded). The exponential Gaussian moments we obtain from Fernique's theorem prove exactly this,[insert Fernique] and thus actually $a_\gamma \in X^{**}$. This is still kind of impractical: We can interpret $a_\gamma(f)$ as the mean of the measure "in direction of f", so we would like to view a_γ as the "general mean" of the measure. But we think of the mean as an object in the probability space (think of sampling from a measure; the empirical mean will lie in the same space X as the samples), but here we can only state $a_\gamma \in X^{**}$. Because $X \subsetneq X^{**}$, we need to do some work in order to identify a_γ with an actual element in X. [Insert proof for $a_\gamma \in X$]

1.1 Creating a Hilbert space structure on linear functionals from scratch

We are working in a Banach space, so it is a small miracle that we are able to conjure a Hilbert space structure (but at first only in the realm of functionals X^*) out of thin air. Later we will construct an isometry from this Hilbert space into a strict subset of X. This Hilbert subspace is the Cameron-Martin space and has a set of very useful properties. But let's start at the beginning:

It is a fundamental fact that every bounded linear functional has second moments with respect to a Gaussian measure, i.e. $X^* \subset L^2(X,\gamma)$. Even more, the embedding

$$\begin{aligned} j : X^* &\hookrightarrow L^2(X,\gamma) \\ f &\mapsto f - a_\gamma(f) \end{aligned}$$

is continuous. The image of j is not closed in the L^2 topology, so we define:

$$X^*_\gamma = \overline{j(X^*)}^{L^2(X,\gamma)} \text{ is called the reproducing kernel Hilbert space,}$$

i.e. X^*_γ consists of all L^2-limits of sequences of the form $(f_n - a_\gamma(f_n))_n$, with $f_n \in X^*$. [Characterize X^*_γ further: All Gaussian random variables?]

This is a definition one easily glances over but there are a few subtleties here that should be pointed out:

- Although each $f \in X^*$ is also a $f \in L^2(X,\gamma)$ (and by extension, $f - a_\gamma(f) \in X^*_\gamma$), there is quite a difference in viewing f as a linear functional in X^* or a square-integrable random variable: The main point is that two "directions" f_1, f_2 that differ only by a component that is degenerate (i.e. the measure puts no mass along this axis) are *different* objects in X^*, but identical as elements in $L^2(X,\gamma)$. Take for example the degenerate Gaussian measure $\delta_0 \otimes N(0,1)$ in two dimensions: Here, $f_1(x) = x_1 + x_2$ and $f_2(x) = x_2$ are definitely not the same functional, but as random variables they are identical, because $x_1 = 0$ a.s.
- While $f \in X^*$ is a linear functional (i.e. specifically $f(0) = 0$), $j(f)$ isn't anymore, but is is rather an affine functional because its mean has been subtracted. On the other hand, $j(f)$ has zero mean, while this is generally not true for $f \in X^*$. This reasoning holds identically for elements $g \in X^*_\gamma$: By definition, $a_\gamma(g) = 0$, but $g(0) \neq 0$ in general. One has to be careful not to mix up those two ideas of "going through the origin".
- By using the covariance inner product $\langle f, g\rangle_{L^2(X,\gamma)} = B_\gamma(f,g) = \int_X [f(x) - a_\gamma(f)] \cdot [g(x) - a_\gamma(g)]\gamma(\mathrm{d}x)$, we obtain a Hilbert space structure in the image of j. This will be of great use later.

By L^2 continuity, a_γ, $\hat{\gamma}$ and B_γ can be extended with their domains changed: For any X^* we can also use elements from X^*_γ, while for $g \in X^*_\gamma$, $a_\gamma(g) = 0$, $\hat{\gamma}(f) = \exp\{-\frac{1}{2}\|f\|_{L^2(X,\gamma)}\}$ and $B_\gamma(g_1,g_2) = \int_X g_1(x)g_2(x)\gamma(\mathrm{d}x)$ for $g_1, g_2 \in X^*_\gamma$. Note that X^*_γ is not a superset of X^* because its elements all need to be centred (as random variables, i.e. $a_\gamma(f) = 0$ for $f \in X^*_\gamma$.

2 Transporting the Hilbert space structure into the probability space

So far, we only have a Hilbert space structure in the domain of linear functionals $f \in X^*$ (strictly speaking on X^*_γ).

3 Fréchet spaces

Let X be a topological vector space with topology $\mathcal{T}$. A few preliminary definitions:

Definition 1. A set C is called

- balanced if for all $x \in C$ and all $|\lambda| \leq 1$ also $\lambda x \in C$.
- convex if for all $x, y \in C$ and all $0 \leq t \leq 1$, also $tx + (1-t)y \in C$.
- absorbent if $\bigcup_{t>0} tC = X$.
- absolutely convex, if C is both convex and balanced. This is the same as saying that for every $x, y \in C$ and real numbers $|t| + |s| = 1$, $tx + sy \in C$. The best image for this is that C contains every parallelogramm with vertices x and y symmetric to the origin.

Definition 2.

- A collection of sets $\mathcal{B} \subset \mathcal{T}$ is called a base of the topology $\mathcal{T}$ if for every open set $U \in \mathcal{T}$ there is a $B \in \mathcal{B}$ with $B \in U$.
- A collection of sets $\mathcal{B}_x \subset \mathcal{T}$ is called a local base at x of the topology $\mathcal{T}$ if for every neighborhood $U \in \mathcal{T}$ of x (i.e. $x \in U$) there is a $B \in \mathcal{B}_x$ with $B \in U$.

We are going to compare two definitions of locally convex spaces.

Definition 3. A topological vector space X is called *locally convex* if the origin has a local base of absolutely convex absorbent sets.

Definition 4. A topological vector space X is called *locally convex* if there is a family of seminorms generating the topology.

Lemma 1. *Definitions* 3 *and* 4 *are equivalent.*

Proof. Proof of $3 \Rightarrow 4$. Define the Minkowski gauge of a set $C \in X$ as

$$\mu_C(x) = \inf\{\lambda > 0 : x \in \lambda C\}.$$

Then we can show that the origin's local base of absolutely convex absorbent sets $\mathcal{B}_0 = \{B_a, a \in I\}$ constitute a family of seminorms via their Minkowski gauges $\{\mu_{B_a}, a \in I\}$ and the latter generates the topology of X. XXX

Proof of $4 \Rightarrow 3$ Using the family of seminorms we can easily build a local base of absolutely convex absorbent sets for the origin.XXX □

Remark 1.

1. In Definition 4, note that the family of seminorms is not necessarily countable.
2. If the family of seminorms $\|\cdot\|_k$ generating the topology is *countable*, then this constitutes a pseudometric generating the topology. To see this, define

$$d(x, y) = \sum_{k=1}^{\infty} 2^{-k} \frac{\|x-y\|_k}{1 + \|x-y\|_k}. \tag{1}$$

   Note that this pseudo metric is neither unique (so the metrization is not unique) nor homogenous (so we can't define a suitable norm).
3. As translation is continuous in any t.v.s, we can use a base for the origin as a base for any point by translation.

Lemma 2. *Let X be a locally convex TVS with its family of seminorms $(\|\cdot\|_\alpha)_{\alpha\in I}$. Then X is Hausdorff if and only if*

$$\forall \alpha \in I : \ \|x\|_\alpha = 0 \quad \Leftrightarrow \quad x = 0 \tag{2}$$

Proof. Let X be Hausdorff. We show (2). If for every $\alpha \in I$, $\|x\|_\alpha = 0$ and we assume that $x \neq 0$, we know by the Hausdorff property that there exists an open neighborhood U_x of x and another open neighborhood of 0 such that they do not intersect. As the topology is generated by the seminorms, we can set $U_x = \{u : \ \|u - x\|_\alpha < \varepsilon \ \forall k \leq K\}$ for some $K \in \mathbb{N}, \varepsilon > 0$. Now obviously $0 \in U_x$ as $\|x\|_\alpha = 0$ for all α. Hence there cannot exist any neighborhood of 0 such that we can separate x and 0. The other direction of (2) is obvious by the definition of seminorms.

Now suppose (2) holds. If we set $x \neq y$, by (2) we know that there is at least one $\alpha \in I$ such that $\|x - y\|_\alpha = M \neq 0$.

By setting $U_x = \{u \in X : \ \|u - x\|_\alpha < M/2\}$ and $V_y = \{v \in X : \ \|v - y\|_\alpha < M/2\}$, we have $U_x \cap V_y = \emptyset$ by the triangle inequality (assume there is a w both in U_x and V_y, then one can show that $\|x-y\|_\alpha \leq \|x-w\|_\alpha + \|w-y\|_\alpha < M$). □

We are going to compare two definitions of Fréchet spaces.

Definition 5. A Fréchet space X is a topological vector space X such that

1. X is locally convex, i.e. $0 \in X$ has a local base of absorbent and absolutely convex sets.
2. The topology $\mathcal{T}$ can be induced by a translation invariant metric d, i.e. $d(x + a, y + a) = d(x, y)$.
3. (X, d) is a complete metric space.

Definition 6. A Fréchet space X is a topological vector space X such that

1. The topology can be induced by a countable family of seminorms $\|\cdot\|_k$, i.e. U open if and only if for all $u \in U$ there is an integer $K \geq 0$ and an $\varepsilon > 0$ such that
$$\{v : \|v - u\|_k < \varepsilon : \ k \leq K\} \subset U$$
2. X is Hausdorff.
3. X is complete w.r.t. the family of seminorms $\|\cdot\|_k$, i.e. if $(x_m)_m$ is a Cauchy sequence w.r.t. all seminorms, then there exists an $x \in X$ with $x_n \to x$ w.r.t. $\|\cdot\|_k$ (and by property 2 even $x_n \to x$).

Remark 2. 1. Note that the items 1-3 are not all "parallely equivalent", but as a set, they are equivalent:

- 5.1 and 5.2 $\Rightarrow$ 6.1 (the countability property of the family of seminorms is stronger than plain local convexity)
- 5.2 $\Rightarrow$ 6.2 (metrizability implies Hausdorff)
- 6.1 and 6.2 $\Rightarrow$ 5.2
- 6.1$\Rightarrow$ 5.1
- 5.3 $\Leftrightarrow$ 6.3

Lemma 3. *Definitions* 5 *and* 6 *are equivalent.*

Proof. **5 $\Rightarrow$ 6.**

ad 1.): From the equivalent definitions of locally convex vector spaces we know that 6.1 is *almost* local convexity. The trouble is that we claim the existence of a *countable* family of seminorms whereas local convexity just gives some family of seminorms. By combining 5.1 (local convexity in its set-theoretic version) and 5.2 (metrizability), we obtain 6.1.

Indeed, the metric d allows us to define small balls $B_n = \{x : \ d(x, 0) < \frac{1}{n}\}$ for each $n \in \mathbb{N}$. As the topology can be induced by d, we know that there is an open set inside each B_n and by local convexity we know of the existence of an absorbent and absolutely convex set $C_n \subset B_n$. Those sets in turn define via their Minkowski gauges μ_n a countable family of seminorms. Equivalence of the topology of X and the topology generated by the family μ_n follows from the fact that the C_n are a local base of the topology of X.

ad 2.): Every metric (or metrizable) space is Hausdorff.

ad 3.): This is a direct consequence of completeness of d.

6 $\Rightarrow$ 5 ad 1.): As noted above, 6.1 is slightly stronger than 5.1.

ad 2.): By 6.2 we know by lemma 2 that the countable family of seminorms in 6.1 fulfills

$$\forall n \in \mathbb{N} : \ \|x\|_k = 0 \quad \Leftrightarrow \quad x = 0.$$

This property makes the pseudometric in (1) (which is one possible pseudometrization) a proper metric. The fact that the topology of X is induced by d is equivalent to 6.1. □

4 The space $\mathbb{R}^\infty$

Lemma 4. *The topological vector space $W = \mathbb{R}^\infty = \{(x_n)_{n\in\mathbb{N}} : \; x_n \in \mathbb{R} : \; \forall n \in \mathbb{N}\}$ with the product topology (i.e. the topology generated by all sets of the form $U_1 \times \cdots \times U_m \times \mathbb{R}^\infty$ with all $U_i \subset \mathbb{R}$ open) is a Fréchet space. One possible metrization is*

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \cdot \frac{|x_n - y_n|}{1 + |x_n - y_n|} \tag{3}$$

Proof. We will prove this by showing that $\mathbb{R}^\infty$ fulfills every item in definition 6.

ad 1.): We define $\|x\|_k := |x_k|$, i.e. the absolute value of the k-th item. This is obviously a seminorm on $\mathbb{R}^\infty$. We show that the product topology an be generated by this family of seminorms.

Let U' be an open set in the product topology. Then U' is a union of sets of the form $U_1 \times \cdots \times U_m \times \mathbb{R}^\infty$. It suffices to consider only open sets out of this basis. Let $U = U_1 \times \cdots \times U_m \times \mathbb{R}^\infty$. Choose a point $u \in U$, i.e. $u = (u_1, \ldots, u_m, u_{m+1}, \ldots)$ where $u_i \in U_i$ for $i = 1, \ldots, m$. Define $\delta = \min_{i=1}^{M} \operatorname{dist}(u_i, \partial U_i)$, where dist is the distance function on $\mathbb{R}$. Then the set $\{v : \|v - u\|_k < \delta : k \leq M\}$ is a subset of U, which constitutes the first part of Definition 6.1. The other direction is obvious by definition of the product topology.

ad 2.): By lemma 2 we know that we only need to show that if and only if for given $u \in \mathbb{R}^\infty$ and all $n \in \mathbb{N}$ we have $\|u\|_k = |u_k| = 0$, then $u = 0 \in \mathbb{R}^\infty$. But this is obvious.

ad 3.): This is easily shown by using completeness of $\mathbb{R}$ in each dimension of $\mathbb{R}^\infty$. $\square$

Lemma 5. *The space $\mathbb{R}^\infty$ with the product topology is a polish space.*

Proof. $\mathbb{R}^\infty$ is separable as $\mathbb{Q}^\infty$ is countable and dense and it is completely metrizable by definition of Fréchet spaces. $\square$

Lemma 6. *The product topology of $\mathbb{R}^\infty$ cannot be generated by a norm.*

Proof. Open sets of $\mathbb{R}^\infty$ are necessarily unbounded but balls defined by any norm must be (by definition) bounded. $\square$

Lemma 7. *The Borel σ-algebra is the same as the product σ-algebra.*

Proof. The Borel-σ-algebra is the σ-algebra generated by open sets (i.e. sets of the form $U_1 \times \cdots \times U_m \times \mathbb{R}^\infty$), which is the same as the product *sigma*-algebra. $\square$

We write e_i for the element of W with $e_i(j) = \delta_{ij}$. Any element $x \in W$ can be written as $x = \sum_{i=1}^{\infty} x_i e_i$.

Lemma 8. *Every continuous linear functional $f \in W^*$ is of the form*

$$f(x) = \sum_{i=1}^{n} a_i x_i \tag{4}$$

for some $a_1, \ldots, a_n \in \mathbb{R}$ and with the notation $W \ni x = (x_1, x_2, \ldots)$. Thus W^* can be identified with c_{00}, the set of all real sequences which are eventually zero.

Proof. Let $f \in W^*$. Define $a_i = f(e_i)$. We show that only finitely many a_i are nonzero. Indeed, assume that for any $N \in \mathbb{N}$ there is an $i_N > N$ such that $a_{i_N} \neq 0$.

As we assumed f to be continuous, for any $\varepsilon > 0$ there is an open neighborhood U of 0 such that

$$\sup_{u \in U} |f(u)| < \varepsilon. \tag{5}$$

By the form of the topology, this U is of the form

$$U = \bigotimes_{i=1}^{n} U_i \times \mathbb{R}^\infty$$

for some $n \in \mathbb{N}$. Now we can define a sequence $(u_N)_N$ of elements u_N (which are sequences) in W as $u_N = \frac{1}{a_{i_N}} \cdot N \cdot e_{i_N}$, i.e. the N-th element is the sequence which has the value $\frac{N}{a_{i_N}}$ at the i_N-th position and 0 elsewhere. For M high enough, $(u_N)_{N \geq M}$ is a sequence in U. But

$$f(u_N) = \frac{N}{a_{i_N}} \cdot f(e_{i_N}) = N \xrightarrow{N \to \infty} \infty,$$

hence (5) is violated and f cannot be continuous.

This means that f acts on only a finite number of basis elements of W and as f is linear we can write it in the form (4). $\square$

We choose the measure μ to be an infinite product of Gaussian measures with variance 1 and mean 0. So the projections of elements $x \in W$ on the coordinates are i.i.d. Gaussians.

Lemma 9. *μ is a Gaussian measure. The covariance form of μ is given by*

$$q(f,g) = \sum_{i=1}^{\infty} f(e_i)g(e_i)$$

where the sum is actually finite.

Proof. We argue by using characteristic functions.

$$\int_W e^{if(x)}\mu(dx) = \int_{\mathbb{R}^\infty} e^{i\sum_{j=1}^n f(e_j)x_j} \bigotimes_{j=1}^{\infty} N(0,1)(dx_j) = \prod_{j=1}^{n} \int_{\mathbb{R}} e^{if(e_j)x_j} N(0,1)(dx_j)$$
$$= \prod_{j=1}^{n} e^{-\frac{f(e_j)^2}{2}} = e^{-\frac{1}{2}\sum_{j=1}^n f(e_j)^2}$$

and thus the covariance form is $q(f,f) = \sum_{j=1}^n f(e_j)^2$ and hence $q(f,g) = \sum_{j=1}^\infty f(e_j)g(e_j)$ where the sum runs until the largest non-zero entry of both f and g. $\square$

Lemma 10. *μ has full support, i.e. there is no open set other than the empty set having zero measure.*

Proof. Let U be open in W, i.e. $U = \bigotimes_{i=1}^n U_i \times \mathbb{R}^\infty$ with U_j open in $\mathbb{R}$. Then $\mu(U) = \prod_{j=1}^n N(0,1)(U_j) > 0$ as $N(0,1)$ has full support on $\mathbb{R}$. $\square$

Remark 3. q is actually positive definite: The only $f \in W^*$ with $q(f,f) = 0$ is $f = 0$. This means that $i : W^* \hookrightarrow L^2(W,\mu)$ with inner product q is an injection in the sense that $i(W^*) \subset L^2$ is isomorphic to W^*. This is actually cool because normally there may be $0 \neq f \in W^*$ with $0 = f \in L^2$ (for example in spaces with degenerate Gaussians like $W = \mathbb{R}^2$ with $N(0,1) \otimes \delta_0$. Here, choosing f as the projection on the second coordinate (written as $f = (0,1)^T$) is not 0 as a linear functional in $W^* = \mathbb{R}^2$ but $f(x) = 0$ μ-almost surely and thus $0 = f \in L^2$ as a square-integrable random variable. So, to reiterate, $W^* \hookrightarrow L^2(W,\mu)$ is an injection and we can view W^* as a **subspace** of L^2. This is because we don't lose "information" by viewing f as an element in L^2. In the example above with $\mathbb{R}^2$ we do lose information because after identifying W^* with a subset of L^2, we can't distinguish $(0,1)^T$ and $(0,0)^T$ as random variables in L^2, although they are quite distinct in $W^* = \mathbb{R}^2$.

Remark 4. As was pointed out, we can think of W^* as a subset of $L^2(W,\mu)$ with inner product q. But: W^* is not complete in the q inner product. Let's define $K = \overline{W^*}^{L^2(W,\mu)}$, the L^2 -closure of W^*.

Lemma 11. *K consists of all functions $f : W \to \mathbb{R}$ of the form*

$$f(x) = \sum_{j=1}^{\infty} a_j x_j \tag{6}$$

with $\sum_{j=1}^\infty |a_j|^2 < \infty$.

Proof. As W^* can be identified with c_{00} and $L^2(W,\mu$ with inner product q can be identified with l^2, the space of square summable sequences, this is equivalent to $\overline{c_{00}}^{l^2} = l^2$, which is for convenience proven in lemma 13. $\square$

Remark 5. Note one subtle point about (6): For an arbitrary $x \in W$, the sum may not converge. It does, however, converge for μ-a.e. $x \in W$. Indeed: Sums of independent random variables (and the $a_j \cdot x_j$ are i.i.d Gaussians) converge a.s. as soon as the sum of their variances converge. Thus, we only need to show

$$\sum_{j=1}^{\infty} a_j^2 \cdot \mathbb{E}x_j^2 < \infty$$

which follows immediately by $\mathbb{E}x_j^2 = 1$ and the condition on $(a_j)_j$.

By the same calculation as in lemma 9, we can show that f is a Gaussian random variable with covariance form $q(f,g) = \sum_{j=1}^\infty f(e_j)g(e_j)$ where now the sum may contain infinitely many terms (but still is of finite value by the Cauchy-Schwarz inequality in l^2).

5 Appendix

Lemma 12 (($\mathbb{R}^\infty, \|\cdot\|_\infty$) is complete)**.** *Let $(a^{(k)})_{k\in\mathbb{N}}$ be a Cauchy sequence (of sequences) in the space of sequences with the supremum norm, i.e.*

$$\text{for all } \varepsilon > 0 \text{ there is a } K_\varepsilon \text{ s. t. for all } k, l \geq K_\varepsilon : \|a^{(k)} - a^{(l)}\|_\infty = \sup_n |a_n^{(k)} - a_n^{(l)}| < \varepsilon.[1]$$

Then $a^{(k)} \to a$ in $(\mathbb{R}^\infty, \|\cdot\|_\infty)$, i.e. there is a sequence $a \in \mathbb{R}^\infty$ such that

$$\text{for all } \varepsilon > 0 \text{ there is a } M_\varepsilon \text{ s. t. for all } k \geq M_\varepsilon : \|a^{(k)} - a\|_\infty = \sup_n |a_n^{(k)} - a_n| < \varepsilon,$$

i.e. $a^{(k)}$ converges uniformly to a.

Proof. We use completeness of $\mathbb{R}$ when we realize that by assumption for every $\varepsilon > 0$ there is an index N_ε such that for every $n \in \mathbb{N}$ there is a limit a_n such that for all $l \geq N_\varepsilon$,

$$|a_n^{(l)} - a_n| < \varepsilon.$$

Now we fix $\varepsilon > 0$ and we want to bound $|a_n^{(k)} - a_n| < \varepsilon$ for all $n \in \mathbb{N}$ by choosing k large enough. For this, we write

$$|a_n^{(k)} - a_n| \leq |a_n^{(k)} - a_n^{(l)}| + |a_n^{(l)} - a_n|$$

Now first we choose $K = K_{\varepsilon/2}$, $N = N_{\varepsilon/2}$ and $l \geq \max\{K, N\}$. This makes the second part smaller than $\varepsilon/2$. As soon as we choose $k \geq K$, the first part is also smaller than $\varepsilon/2$. Hence we can choose $M_\varepsilon = K_{\varepsilon/2}$ and we are done. □

Lemma 13 ($\overline{c_{00}}^{l^2} = l^2$)**.** *Define $c_{00} = \{(a_n)_{n\in\mathbb{N}} : \exists m : \; \forall n \geq m : \; a_n = 0\}$ and choose a Cauchy sequence in c_{00} w.r.t. the l^2-norm, i.e.*

$$\text{for all } \varepsilon > 0 \text{ there is a } N_\varepsilon \text{ s. t. for all } k, l \geq N_\varepsilon : \sum_{n=1}^{\infty} |a_n^{(k)} - a_n^{(l)}|^2 < \varepsilon.$$

Then

$$a^{(k)} \to a \text{ as a sequence in } l^2$$

Proof. We need to show that there is a sequence a with

1. $\|a^{(k)} - a\|_2 \xrightarrow{k\to\infty} 0$ and
2. $a \in l^2$.

The existence of such an a follows immediately by dropping the summation symbol and using completeness of $\mathbb{R}$ to obtain a limit element a_n for every $n \in \mathbb{N}$, which constitutes a sequence $a = (a_n)_{n\in\mathbb{N}}$.

Now for 1.) we see that for any $R \in \mathbb{N}$

$$\sum_{n=1}^{R} |a_n^{(k)} - a_n|^2 \leq \sum_{n=1}^{R} 2 \cdot |a_n^{(k)} - a_n^{(l)}|^2 + 2 \cdot |a_n^{(l)} - a_n|^2.$$

If we choose $l \geq M := M_{\sqrt{\frac{\varepsilon}{4\cdot R}}}$ (the index from Lemma 12), the second sum is bounded by $\varepsilon/2$. After setting $N := N_{\sqrt{\frac{\varepsilon}{4}}}$ and claiming additionally $k, l \geq N$ (thus $l \geq \max\{M, N\}$), we see that uniformly in R we just need to set $k \geq N$ in order to bound

$$\sum_{n=1}^{R} |a_n^{(k)} - a_n|^2 < \varepsilon$$

and thus the bound also holds for the infinite sum, which proves 1.).

For 2.) we see that

$$\sum_{n=1}^{R} |a_n|^2 \leq \sum_{n=1}^{R} 2 \cdot |a_n - a_n^{(k)}|^2 + 2 \cdot |a_n^{(k)}|^2$$

which is finitely bounded uniformly in R as we can set k such that the first sum is arbitrarily small and the second sum is some finite value (the l^2-norm of $a^{(k)}$. □

---

[1] We can think of $(a^{(k)})_k$ being a "uniformly Cauchy" sequence.