Quiz Section 4 - Best Approximation and Orthogonal Projection

Theorem:

Let V be an inner-product space, $W \subset V$ a subspace. Then $u^* \in W$ is the best approximation of $u \in V$ within W if and only if $(u - u^*) \perp W$, that is $\langle u - u^*, w \rangle = 0 \ \forall w \in W$.

Definition:

Let V be an inner-product space, $W \subset V$ a subspace, and $u \in V$. Then a vector $u^* \in W$ satisfying $(u - u^*) \perp W$ is called the *orthogonal projection* of u into W.

Theorem:

Let V be an inner-product space, and $W \subset V$ a subspace of dimension n, spanned by the orthonormal system $\{e_1, \ldots, e_n\}$. Then $\forall u \in V$ the orthogonal projection of u into W is given by:

$$u^* = \sum_{i=1}^{n} \langle u, e_i \rangle e_i$$

Proof:

We will show that u^* is indeed the orthogonal projection, that is, $\langle u - u^*, w \rangle = 0 \ \forall w \in W$.

Let $w \in W$, $W = $span$\{e_1, \ldots, e_n\}$. Then there exist $\alpha_1, \ldots \alpha_n \in \mathbb{C}$ such that $w = \sum_{i=1}^{n} \alpha_i e_i$.

$$\begin{aligned}
\langle u-u^*,w\rangle &= \left\langle u-\sum_{i=1}^{n}\langle u,e_i\rangle e_i, \sum_{j=1}^{n}\alpha_j e_j\right\rangle \\
&= \sum_{j=1}^{n}\overline{\alpha_j}\left\langle u-\sum_{i=1}^{n}\langle u,e_i\rangle e_i, e_j\right\rangle \\
&= \sum_{j=1}^{n}\overline{\alpha_j}\left(\langle u,e_j\rangle-\left\langle\sum_{i=1}^{n}\langle u,e_i\rangle e_i, e_j\right\rangle\right) \\
&= \sum_{j=1}^{n}\overline{\alpha_j}\left(\langle u,e_j\rangle-\sum_{i=1}^{n}\langle u,e_i\rangle\cdot\langle e_i,e_j\rangle\right) \\
&\underset{*}{=} \sum_{j=1}^{n}\overline{\alpha_j}\left(\langle u,e_j\rangle-\langle u,e_j\rangle\right) \\
&= \sum_{j=1}^{n}\overline{\alpha_j}\cdot 0 \\
&= 0
\end{aligned}$$

The transition marked with * is justified by the fact that $\{e_1,\dots e_n\}$ is an orthonormal system, namely:

$$\langle e_i,e_j\rangle=\begin{cases}1, & i=j.\\ 0, & i\neq j.\end{cases}$$

This proves the theorem.

Exercise 1 (From a quiz in 2008)

Let V be the set of continuous functions from $\mathbb{R}$ to $\mathbb{C}$ with the property:

$$\int_{\mathbb{R}}|f(x)|^2 e^{-x^2}dx<\infty$$

1. Prove that V is a vector space (with the standard operations of addition and multiplication by a scalar), and that V is an inner-product space with the inner product:

$$\langle f,g\rangle=\int_{\mathbb{R}}f(x)\overline{g(x)}e^{-x^2}dx$$

2. Prove that all polynomials belong to V.

3. Find the orthogonal projection of x^3 onto the subspace of polynomials of degree ≤ 2. (Hint: you may use the fact that $\int_{\mathbb{R}}e^{-x^2}=\sqrt{\pi}$)

Solution:

1. We will show the properties of closure under addition and multiplication by a scalar, other properties are trivial.

Multipilication by a scalar: Let $f \in V$, $\lambda \in \mathbb{C}$.

$$\int_{\mathbb{R}} |\lambda f(x)|^2 e^{-x^2} dx = |\lambda|^2 \int_{\mathbb{R}} |f(x)|^2 e^{-x^2} dx < \infty$$

Therefore V is closed under multiplication by a scalar.

Addition: let $f, g \in$ V

$$\begin{aligned} \int_{\mathbb{R}} |f(x) + g(x)|^2 e^{-x^2} dx &\leq \int_{\mathbb{R}} (|f(x)| + |g(x)|)^2 e^{-x^2} dx \\ &= \int_{\mathbb{R}} \left(|f(x)|^2 + |g(x)|^2 + |f(x) g(x)| \right) e^{-x^2} dx \end{aligned}$$

Let's look at the term $\int_{\mathbb{R}} |f(x) g(x)| e^{-x^2} dx$, $\forall x \in \mathbb{R}$ we have:

$$\begin{aligned} |f(x) g(x)| &\leq \max\left(|f(x)|^2, |g(x)|^2 \right) \\ &\leq |f(x)|^2 + |g(x)|^2 \end{aligned}$$

Therefore:

$$\int_{\mathbb{R}} |f(x) g(x)| e^{-x^2} dx \leq \int_{\mathbb{R}} |f(x)|^2 e^{-x^2} dx + \int_{\mathbb{R}} |g(x)|^2 e^{-x^2} dx$$

Back to proving closure under addition:

$$\begin{aligned} \int_{\mathbb{R}} |f(x) + g(x)|^2 e^{-x^2} dx &\leq \int_{\mathbb{R}} (|f(x)| + |g(x)|)^2 e^{-x^2} dx \\ &= \int_{\mathbb{R}} \left(|f(x)|^2 + |g(x)|^2 + |f(x) g(x)| \right) e^{-x^2} dx \\ &\leq \int_{\mathbb{R}} \left(|f(x)|^2 + |g(x)|^2 + |f(x)|^2 + |g(x)|^2 \right) e^{-x^2} dx \\ &\leq 2 \int_{\mathbb{R}} |f(x)|^2 e^{-x^2} dx + 2 \int_{\mathbb{R}} |g(x)|^2 e^{-x^2} dx \\ &< \infty \end{aligned}$$

Therefore V is closed under addition

We proved V is a vector space, it is left to prove that $\langle f, g \rangle = \int_{\mathbb{R}} f(x) \overline{g(x)} e^{-x^2} dx$ defines an inner-product on V.

First we will show that this is well defined, which means that $\forall f, g \in V$ $|\langle f, g \rangle| < \infty$

$$
\begin{aligned}
|\langle f,g\rangle| &= \left|\int_{\mathbb{R}} f(x)\,\overline{g(x)}e^{-x^2}dx\right| \\
&\leq \int_{\mathbb{R}} \left|f(x)\,\overline{g(x)}e^{-x^2}\right| dx \\
&= \int_{\mathbb{R}} |f(x)|\cdot|g(x)|\,e^{-x^2}dx \\
&\leq \int_{\mathbb{R}} \left(|f(x)|^2+|g(x)|^2\right)e^{-x^2}dx \\
&= \int_{\mathbb{R}} |f(x)|^2\,e^{-x^2}dx+\int_{\mathbb{R}} |g(x)|^2\,e^{-x^2}dx \\
&< \infty
\end{aligned}
$$

We will now show the simple properties of linearity and conjugate symmetry:

$$
\begin{aligned}
\langle f+g,h\rangle &= \int_{\mathbb{R}} (f(x)+g(x))\,\overline{h(x)}e^{-x^2}dx \\
&= \int_{\mathbb{R}} f(x)\,\overline{h(x)}e^{-x^2}dx+\int_{\mathbb{R}} g(x)\,\overline{h(x)}e^{-x^2}dx \\
&= \langle f,h\rangle+\langle g,h\rangle
\end{aligned}
$$

$$
\begin{aligned}
\langle \alpha f,g\rangle &= \int_{\mathbb{R}} \alpha f(x)\,\overline{g(x)}e^{-x^2}dx \\
&= \alpha\int_{\mathbb{R}} f(x)\,\overline{g(x)}e^{-x^2}dx \\
&= \alpha\,\langle f,g\rangle
\end{aligned}
$$

$$
\begin{aligned}
\overline{\langle f,g\rangle} &= \overline{\int_{\mathbb{R}} f(x)\,\overline{g(x)}e^{-x^2}dx} \\
&= \int_{\mathbb{R}} \overline{f(x)\,\overline{g(x)}e^{-x^2}}dx \\
&= \int_{\mathbb{R}} \overline{f(x)}\overline{\overline{g(x)}}\overline{e^{-x^2}}dx \\
&= \int_{\mathbb{R}} \overline{f(x)}g(x)\,e^{-x^2}dx \\
&= \int_{\mathbb{R}} g(x)\,\overline{f(x)}e^{-x^2}dx \\
&= \langle g,f\rangle
\end{aligned}
$$

The property $\langle f,f\rangle\geq 0$ follows from the fact that the integrand, $|f(x)|^2\,e^{-x^2}$ is non-negative.

It is left to show the property $\langle f, f\rangle = 0 \iff f = 0$.

The direction $f = 0 \Longrightarrow \langle f, f\rangle = 0$ is trivial.

To prove $\langle f, f\rangle = 0 \Longrightarrow f = 0$ we will show the equivalent $f \neq 0 \Longrightarrow \langle f, f\rangle \neq 0$.

$f \neq 0$ therefore $\exists x_0 \in \mathbb{R}$ such that $|f(x_0)| = M > 0$.

f is continuous, therefore $\exists \delta > 0$ such that $|f(x) - f(x_0)| < \frac{M}{2}$ for all x such that $|x - x_0| \leq \delta$.

Therefore for all $x \in [x_0 - \delta, x_0 + \delta]$ we have $|f(x)| > \frac{M}{2}$

For all $x \in [x_0 - \delta, x_0 + \delta]$ we have $e^{-x^2} \geq e^{-(|x_0|+\delta)^2} = C > 0$

Combining everything together:

$$\begin{aligned}
\langle f, f\rangle &= \int_{\mathbb{R}} |f(x)|^2 e^{-x^2} dx \\
&= \int_{[x_0-\delta, x_0+\delta]} |f(x)|^2 e^{-x^2} dx \\
&\geq \int_{[x_0-\delta, x_0+\delta]} \left(\frac{M}{2}\right)^2 C dx \\
&= \frac{M^2 C}{2}\delta \\
&> 0
\end{aligned}$$

This completes the proof that $\langle f, g\rangle$ is an inner-product on V.

2. We already showed that V is a vector space, therefore to show that every polynomial is in V it is enough to show that $x^n \in V \ \forall n \in \mathbb{N}$.

For any $n \in \mathbb{N}$ we have $\lim_{|x|\to\infty} x^n e^{-|x|} = 0$, therefore exists $M > 0$ such that $|x| > M$ implies $\left|x^n e^{-|x|}\right| < 1 \Rightarrow |x^n| < e^{|x|}$.

Therefore $\left|x^{2n}\right| < e^{2|x|}$ and $\left|x^{2n}\right| e^{-x^2} < e^{-x^2+2|x|}$ for any x such that $|x| > M$.

Note that for $|x| > 3$ we have $e^{-x^2+2|x|} < e^{-|x|}$ (because for any $|x| > 3$ we have $-x^2 + 2|x| < -|x|$).

Without loss of generality we assume $M > 3$

Therefore we have:

$$\begin{aligned}
\int_{\mathbb{R}} |x^n|^2 e^{-x^2} dx &= \int_{\mathbb{R}} \left|x^{2n}\right| e^{-x^2} dx \\
&= \int_{x \leq M} \left|x^{2n}\right| e^{-x^2} dx + \int_{|x|>M} \left|x^{2n}\right| e^{-x^2} dx
\end{aligned}$$

The first term is finite because the integrand and the interval of integration are both bounded.

It is left to show that the second term is finite.

$$
\begin{aligned}
\int_{|x|>M} \left|x^{2n}\right| e^{-x^2} dx &< \int_{|x|>M} e^{-x^2+2|x|} dx \\
&< \int_{|x|>M} e^{-|x|} dx \\
&= 2\int_{(M,\infty)} e^{-x} dx \\
&\leq 2\int_{(0,\infty)} e^{-x} dx \\
&= 2
\end{aligned}
$$

This completes the proof that $x^n \in V \ \forall n \in \mathbb{N}$.

Alternative proof:

The proof given above is the one in the lecturer's solutions, it is very short and efficient proof but perhaps hard to come up with. here is an alternative proof which is easier (but less compact):

We need to show that for any $n \in \mathbb{N} \cup \{0\}$ $x^n \in V$.

For this we need to show that for any $n \in \mathbb{N} \cup \{0\}$we have $\int_{\mathbb{R}} |x^n|^2 e^{-x^2} dx < \infty$. Since $x^n \in \mathbb{R}$ we have $|x^n|^2 = (x^n)^2 = x^{2n}$.

Therefore it is the same to show that $\int_{\mathbb{R}} x^{2n} e^{-x^2} dx < \infty$. We do this by induction. From the hint we know that this holds for $n = 0$ (because $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi} < \infty$).

We assume that the integral converges for $n-1$ and show that it converges for n.

$$
\begin{aligned}
\int_{\mathbb{R}} x^{2n} e^{-x^2} dx &= \int_{\mathbb{R}} x^{2n-1} x \left(-\frac{1}{2}\right) (-2)\, e^{-x^2} dx \\
&= -\frac{1}{2}\int_{\mathbb{R}} x^{2n-1} \left(-2x e^{-x^2}\right) dx \\
&= -\frac{1}{2}\int_{\mathbb{R}} x^{2n-1} d\left(e^{-x^2}\right) \\
&= -\frac{1}{2}\left(x^{2n-1} e^{-x^2}|_{-\infty}^{\infty} - \int_{\mathbb{R}} e^{-x^2} d\left(x^{2n-1}\right)\right) \\
&= -\frac{1}{2}\left(0 - 0 - (2n-1)\int_{\mathbb{R}} x^{2n-2} e^{-x^2} dx\right) \\
&= \frac{2n-1}{2}\int_{\mathbb{R}} x^{2(n-1)} e^{-x^2} dx \\
&< \infty
\end{aligned}
$$

Where the last transition uses the induction assumption.

Therefore every polynomial is in V.

3. We first find an orthonormal basis for W (the space of polynomials of degree≤ 2)

In order to do that we use the Gram-Schmidt process for the vectors $1, x, x^2$

First we will calculate some integrals which will be useful later:

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$$

$$\int_{\mathbb{R}} xe^{-x^2} dx = 0$$

$$\begin{aligned}
\int_{\mathbb{R}} x^2 e^{-x^2} dx &= -\frac{1}{2} \int_{\mathbb{R}} xd\left(e^{-x^2}\right) \\
&= -\frac{1}{2}\left(xe^{-x^2}|_{-\infty}^{\infty} - \int_{\mathbb{R}} e^{-x^2} dx\right) \\
&= -\frac{1}{2}\left(0 - 0 - \sqrt{\pi}\right) \\
&= \frac{\sqrt{\pi}}{2}
\end{aligned}$$

$$\int_{\mathbb{R}} x^3 e^{-x^2} dx = 0$$

$$\begin{aligned}
\int_{\mathbb{R}} x^4 e^{-x^2} dx &= -\frac{1}{2} \int_{\mathbb{R}} x^3 d\left(e^{-x^2}\right) \\
&= -\frac{1}{2}\left(x^3 e^{-x^2}|_{-\infty}^{\infty} - \int_{\mathbb{R}} e^{-x^2} d\left(x^3\right)\right) \\
&= -\frac{1}{2}\left(x^3 e^{-x^2}|_{-\infty}^{\infty} - 3\int_{\mathbb{R}} x^2 e^{-x^2} dx\right) \\
&= -\frac{1}{2}\left(0 - 0 - 3\frac{\sqrt{\pi}}{2}\right) \\
&= \frac{3\sqrt{\pi}}{4}
\end{aligned}$$

Where the integrals of xe^{-x^2} and $x^3e^{-x^2}$ are zero because these are odd functions.

$v_1 = 1$, $\|v_1\|^2 = \int_{\mathbb{R}} e^{-x^2} = \sqrt{\pi}$ hence $e_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt[4]{\pi}}$.

$v_2 = x$, therefore:

$$\tilde{e_2} = v_2 - \langle v_2, e_1 \rangle e_1 = x - \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} xe^{-x^2} dx = x$$

We now normalize $\tilde{e}_2$:

$$
\begin{aligned}
\|\tilde{e}_2\|^2 &= \int_{\mathbb{R}} x^2 e^{-x^2} dx \\
&= \frac{\sqrt{\pi}}{2}
\end{aligned}
$$

Therefore $e_2 = \frac{\tilde{e}_2}{\|\tilde{e}_2\|} = \frac{\sqrt{2}}{\sqrt[4]{\pi}} x$
$v_3 = x^2$, so we have:

$$\tilde{e}_3 = v_3 - \langle v_3, e_1 \rangle e_1 - \langle v_3, e_2 \rangle e_2$$

We calculate the inner-products:

$$
\begin{aligned}
\langle v_3, e_1 \rangle &= \int_{\mathbb{R}} x^2 \frac{1}{\sqrt[4]{\pi}} e^{-x^2} dx \\
&= \frac{1}{\sqrt[4]{\pi}} \int_{\mathbb{R}} x^2 e^{-x^2} dx \\
&= \frac{1}{\sqrt[4]{\pi}} \frac{\sqrt{\pi}}{2} \\
&= \frac{\sqrt[4]{\pi}}{2}
\end{aligned}
$$

$$
\begin{aligned}
\langle v_3, e_2 \rangle &= \int_{\mathbb{R}} x^2 \frac{\sqrt{2}}{\sqrt[4]{\pi}} x e^{-x^2} dx \\
&= \frac{\sqrt{2}}{\sqrt[4]{\pi}} \int_{\mathbb{R}} x^3 e^{-x^2} dx \\
&= 0
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
\tilde{e}_3 &= x^2 - \frac{\sqrt[4]{\pi}}{2} \frac{1}{\sqrt[4]{\pi}} \\
&= x^2 - \frac{1}{2}
\end{aligned}
$$

We now normalize $\tilde{e}_3$:

$$\begin{aligned}
\|\tilde{e}_3\|^2 &= \int_{\mathbb{R}} \left(x^2 - \frac{1}{2}\right)^2 e^{-x^2}\,dx \\
&= \int_{\mathbb{R}} \left(x^4 - x^2 + \frac{1}{4}\right) e^{-x^2}\,dx \\
&= \frac{3\sqrt{\pi}}{4} - \frac{\sqrt{\pi}}{2} + \frac{\sqrt{\pi}}{4} \\
&= \frac{\sqrt{\pi}}{2}
\end{aligned}$$

Hence:

$$e_3 = \frac{\tilde{e}_3}{\|\tilde{e}_3\|} = \frac{\sqrt{2}}{\sqrt[4]{\pi}} \left(x^2 - \frac{1}{2}\right)$$

And the desired orthogonal projection is $\sum_{i=1}^{3} \left\langle x^3, e_i \right\rangle e_i$

Note that $\left\langle x^3, e_1 \right\rangle = \left\langle x^3, e_3 \right\rangle = 0$ (Integrating an odd function on a symmetrical interval).

$$\begin{aligned}
\left\langle x^3, e_2 \right\rangle &= \int_{\mathbb{R}} x^3 \frac{\sqrt{2}}{\sqrt[4]{\pi}} x e^{-x^2}\,dx \\
&= \frac{\sqrt{2}}{\sqrt[4]{\pi}} \int_{\mathbb{R}} x^4 e^{-x^2}\,dx \\
&= \frac{\sqrt{2}}{\sqrt[4]{\pi}} \frac{3\sqrt{\pi}}{4} \\
&= \frac{3\sqrt[4]{\pi}}{2\sqrt{2}}
\end{aligned}$$

And therefore the orthogonal projection is $\left\langle x^3, e_2 \right\rangle e_2 = \frac{3\sqrt[4]{\pi}}{2\sqrt{2}} \frac{\sqrt{2}}{\sqrt[4]{\pi}} x = \frac{3}{2} x$.