The Gramm-Schmidt algorithm

Let V be a vector space over a field F. Recall that a bilinear form is a function B which takes as input two vectors $\vec{v}$ and $\vec{w}$ and is linear in each input, meaning

$$B(\vec{v}_1 + \vec{v}_2, \vec{w}) = B(\vec{v}_1, \vec{w}) + B(\vec{v}_2, \vec{w})$$

$$B(\vec{v}, \vec{w}_1 + \vec{w}_2) = B(\vec{v}, \vec{w}_1) + B(\vec{v}, \vec{w}_2)$$

$$B(c\vec{v}, \vec{w}) = B(\vec{v}, c\vec{w}) = cB(\vec{v}, \vec{w}).$$

The bilinear form B is called ***symmetric*** if $B(\vec{v}, \vec{w}) = B(\vec{w}, \vec{v})$.

A symmetric bilinear form over $\mathbb{R}$ is called ***positive definite*** if, for all nonzero vectors $\vec{v}$, we ave $B(\vec{v}, \vec{v}) > 0$.

A positive definite symmetric bilinear form is called an ***inner product***. We'll often denote an inner product as $(\vec{v}|\vec{w})$, $\langle \vec{v}, \vec{w} \rangle$ or $\vec{v} \cdot \vec{w}$.

If V is a vector space with an inner product, and $X \subseteq V$, then $X \cap X^\perp = \{0\}$ and it is often true that $V = X \oplus X^\perp$. If this holds, then we write p_X for the projection onto X whose kernel is $X^\perp$.

In particular, if X has a finite orthonormal basis $\vec{u}_1$, $\vec{u}_2$, ..., $\vec{u}_n$, then $V = X \oplus X^\perp$ and

$$p_X(\vec{v}) = \sum_{i=1}^{n} B(\vec{u}_i, \vec{v})\vec{u}_i.$$

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

So, take an $n-1$ dimensional subspace X of V. By induction, X has an orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$.

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

So, take an $n-1$ dimensional subspace X of V. By induction, X has an orthonormal basis $\vec{u}_1$, $\vec{u}_2$, $\dots$, $\vec{u}_{n-1}$.

Let $\vec{v}$ be any vector not in X. Put $\vec{u} = \vec{v} - p_X(\vec{v})$. Then $\vec{u}$ is orthogonal to X and (since $\vec{v} \notin X$), $\vec{u}$ is not 0. This means that we have

$$B(\vec{u}_1, \vec{u}) = B(\vec{u}_2, \vec{u}) = \cdots = B(\vec{u}_{n-1}, \vec{u}) = 0$$

$$B(\vec{u}, \vec{u}) > 0.$$

Finally, let us show that any finite dimensional vector space V, with an inner product B, has an orthonormal basis. Our proof is by induction on $\dim V$; the base case $\dim V = 0$ is trivial.

So, take an $n-1$ dimensional subspace X of V. By induction, X has an orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$.

Let $\vec{v}$ be any vector not in X. Put $\vec{u} = \vec{v} - p_X(\vec{v})$. Then $\vec{u}$ is orthogonal to X and (since $\vec{v} \notin X$), $\vec{u}$ is not 0. This means that we have

$$B(\vec{u}_1, \vec{u}) = B(\vec{u}_2, \vec{u}) = \cdots = B(\vec{u}_{n-1}, \vec{u}) = 0$$

$$B(\vec{u}, \vec{u}) > 0.$$

Rescaling $\vec{u}$ by $\sqrt{B(\vec{u}, \vec{u})}$, we get an orthonormal basis for X. $\square$

Remark: It is often convenient to work with orthogonal but not orthonormal bases in this algorithm. If $\vec{u}_1$, $\vec{u}_2$, $\ldots$, $\vec{u}_n$ is an orthogonal basis for X, then

$$p_X(\vec{v}) = \sum_{i=1}^{n} \frac{B(\vec{u}_i, \vec{v})}{B(\vec{u}_i, \vec{u}_i)} \vec{u}_i.$$

A boring example: Let

$$\vec{u}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} 3 & 1 & 2 & 0 & 0 \end{bmatrix} \qquad \vec{u}_3 = \begin{bmatrix} 6 & 0 & 3 & 2 & 3 \end{bmatrix}.$$

- Find a vector $\vec{v}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2)$ which is orthogonal to $\vec{u}_1$:

- Find a vector $\vec{w}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ which is orthogonal to $\vec{u}_1$ and $\vec{u}_2$:

A boring example: Let

$$\vec{u}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} 3 & 1 & 2 & 0 & 0 \end{bmatrix} \qquad \vec{u}_3 = \begin{bmatrix} 6 & 0 & 3 & 2 & 3 \end{bmatrix}.$$

- Find a vector $\vec{v}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2)$ which is orthogonal to $\vec{u}_1$:

  The orthogonal projection of $\vec{u}_2$ onto $\mathbb{R}\vec{u}_1$ is $\frac{\vec{u}_1 \cdot \vec{u}_2}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 = \frac{4}{2}\vec{u}_1 = 2\vec{u}_1$. So we should take $\vec{v} = \vec{u}_2 - 2\vec{u}_1 = \begin{bmatrix} 1 & -1 & 2 & 0 & 0 \end{bmatrix}$.

- Find a vector $\vec{w}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ which is orthogonal to $\vec{u}_1$ and $\vec{u}_2$:

A boring example: Let

$$\vec{u}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} 3 & 1 & 2 & 0 & 0 \end{bmatrix} \qquad \vec{u}_3 = \begin{bmatrix} 6 & 0 & 3 & 2 & 3 \end{bmatrix}.$$

- Find a vector $\vec{v}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2)$ which is orthogonal to $\vec{u}_1$:

  The orthogonal projection of $\vec{u}_2$ onto $\mathbb{R}\vec{u}_1$ is $\frac{\vec{u}_1 \cdot \vec{u}_2}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 = \frac{4}{2}\vec{u}_1 = 2\vec{u}_1$. So we should take $\vec{v} = \vec{u}_2 - 2\vec{u}_1 = \begin{bmatrix} 1 & -1 & 2 & 0 & 0 \end{bmatrix}$.

- Find a vector $\vec{w}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ which is orthogonal to $\vec{u}_1$ and $\vec{u}_2$:

  We compute the orthogonal projection onto $\mathrm{Span}(\vec{u}_1, \vec{u}_2)$ using the orthonormal basis $\vec{u}_1$, $\vec{v}$. The projection is

$$\frac{\vec{u}_1 \cdot \vec{u}_3}{\vec{u}_1 \cdot \vec{u}_1}\vec{u}_1 + \frac{\vec{v} \cdot \vec{u}_3}{\vec{v} \cdot \vec{v}}\vec{v} = \frac{6}{2}\vec{u}_1 + \frac{12}{6}\vec{v} = \begin{bmatrix} 5 & 1 & 4 & 0 & 0 \end{bmatrix}.$$

  So we should take $\vec{w} = \begin{bmatrix} 6 & 0 & 3 & 2 & 3 \end{bmatrix} - \begin{bmatrix} 5 & 1 & 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 2 & 3 \end{bmatrix}$.

A boring example: Let

$$\vec{u}_1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix} \qquad \vec{u}_2 = \begin{bmatrix} 3 & 1 & 2 & 0 & 0 \end{bmatrix} \qquad \vec{u}_3 = \begin{bmatrix} 6 & 0 & 3 & 2 & 3 \end{bmatrix}.$$

- Find a vector $\vec{v}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2)$ which is orthogonal to $\vec{u}_1$:

  The orthogonal projection of $\vec{u}_2$ onto $\mathbb{R}\vec{u}_1$ is $\frac{\vec{u}_1 \cdot \vec{u}_2}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 = \frac{4}{2}\vec{u}_1 = 2\vec{u}_1$. So we should take $\vec{v} = \vec{u}_2 - 2\vec{u}_1 = \begin{bmatrix} 1 & -1 & 2 & 0 & 0 \end{bmatrix}$.

- Find a vector $\vec{w}$ in $\mathrm{Span}(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ which is orthogonal to $\vec{u}_1$ and $\vec{u}_2$:

  We compute the orthogonal projection onto $\mathrm{Span}(\vec{u}_1, \vec{u}_2)$ using the orthonormal basis $\vec{u}_1$, $\vec{v}$. The projection is

$$\frac{\vec{u}_1 \cdot \vec{u}_3}{\vec{u}_1 \cdot \vec{u}_1} \vec{u}_1 + \frac{\vec{v} \cdot \vec{u}_3}{\vec{v} \cdot \vec{v}} \vec{v} = \frac{6}{2}\vec{u}_1 + \frac{12}{6}\vec{v} = \begin{bmatrix} 5 & 1 & 4 & 0 & 0 \end{bmatrix}.$$

  So we should take $\vec{w} = \begin{bmatrix} 6 & 0 & 3 & 2 & 3 \end{bmatrix} - \begin{bmatrix} 5 & 1 & 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1 & 2 & 3 \end{bmatrix}$.

  In short, our orthogonal basis is $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & -1 & 2 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & -1 & -1 & 2 & 3 \end{bmatrix}$.

Our orthogonal basis is $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & -1 & 2 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & -1 & -1 & 2 & 3 \end{bmatrix}$.

And the orthonormal version is $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \end{bmatrix}$, $\frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -1 & 2 & 0 & 0 \end{bmatrix}$, $\frac{1}{\sqrt{16}} \begin{bmatrix} 1 & -1 & -1 & 2 & 3 \end{bmatrix}$.

A more interesting example: Let V be the vector space of polynomials of degree $\leq n$ in $\mathbb{R}[x]$. Put

$$\langle f(x), g(x) \rangle = \int_0^1 f(x)g(x)dx.$$

Let's find an orthogonal basis. For $n = 0$, we just take $u_0(x) = 1$.

A more interesting example: Let V be the vector space of polynomials of degree $\leq n$ in $\mathbb{R}[x]$. Put

$$\langle f(x), g(x) \rangle = \int_0^1 f(x)g(x)dx.$$

Let's find an orthogonal basis. For $n = 0$, we just take $u_0(x) = 1$. For $n = 1$, we can take $u_1(x) = x - \frac{\int_0^1 1 \cdot x dx}{\int_0^1 1 \cdot 1 dx} = x - 1/2$.

A more interesting example: Let V be the vector space of polynomials of degree $\leq n$ in $\mathbb{R}[x]$. Put

$$\langle f(x), g(x) \rangle = \int_0^1 f(x)g(x)dx.$$

Let's find an orthogonal basis. For $n = 0$, we just take $u_0(x) = 1$. For $n = 1$, we can take $u_1(x) = x - \frac{\int_0^1 1 \cdot x dx}{\int_0^1 1 \cdot 1 dx} = x - 1/2$.

To find our next polynomial, project x^2 onto $\text{Span}(u_0, u_1)$:

$$\frac{\int_0^1 1 \cdot x^2 dx}{\int_0^1 1 \cdot 1 dx} 1 + \frac{\int_0^1 (x - 1/2) \cdot x^2 dx}{\int_0^1 (x - 1/2) \cdot (x - 1/2) dx} (x - 1/2) = x - 1/6$$

so $u_2 = x^2 - x + 1/6$.

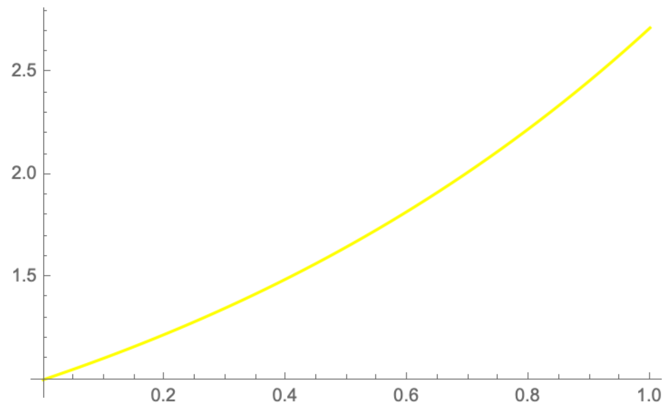
$$u_0 = 1 \quad u_1 = x - 1/2 \quad u_2 = x^2 - x + 1/6.$$

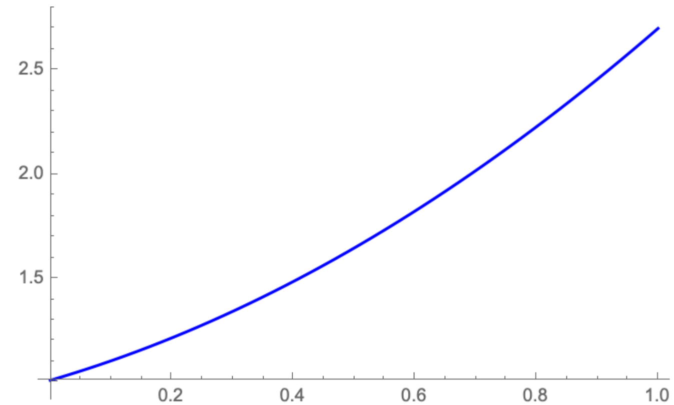
We can now take any function on $[0,1]$ and find the closest quadratic to it. For example, the orthogonal projection of e^x onto quadratics is

$$\frac{\int_0^1 e^x u_0(x)dx}{\int_0^1 u_0(x)^2 dx} u_0(x) + \frac{\int_0^1 e^x u_1(x)dx}{\int_0^1 u_1(x)^2 dx} u_1(x) + \frac{\int_0^1 e^x u_2(x)dx}{\int_0^1 u_2(x)^2 dx} u_2(x).$$

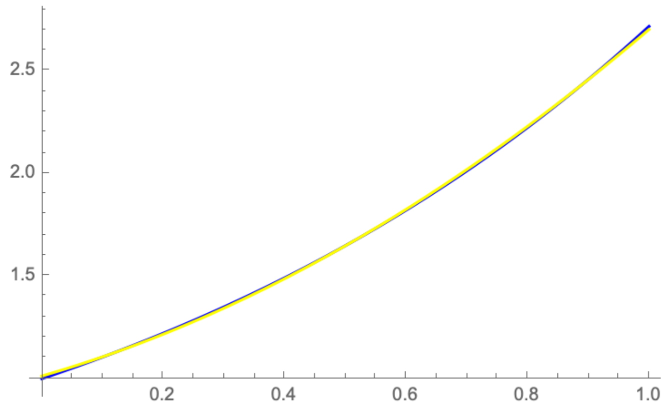
$$(e-1) * u_0(x) + (-6e+18) * u_1(x) + (210e - 570) * u_2(x) =: q(x).$$

How good is this? Here are plots of e^x and $q(x)$:

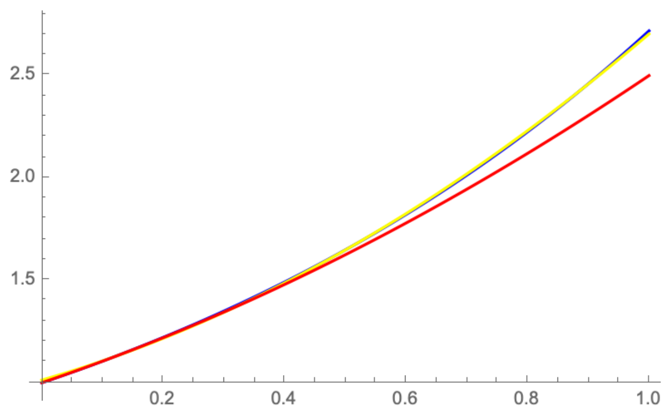




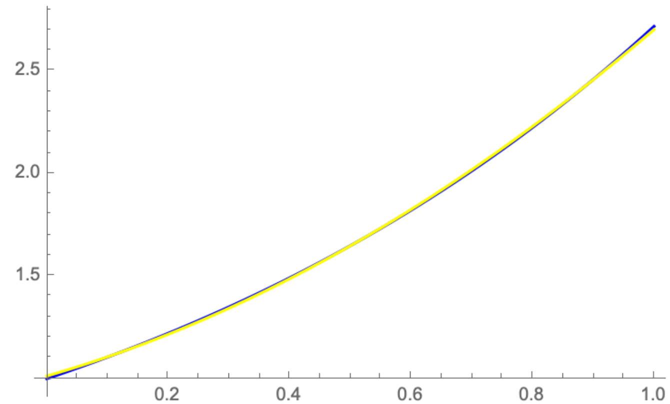
How good is this? Here they are together



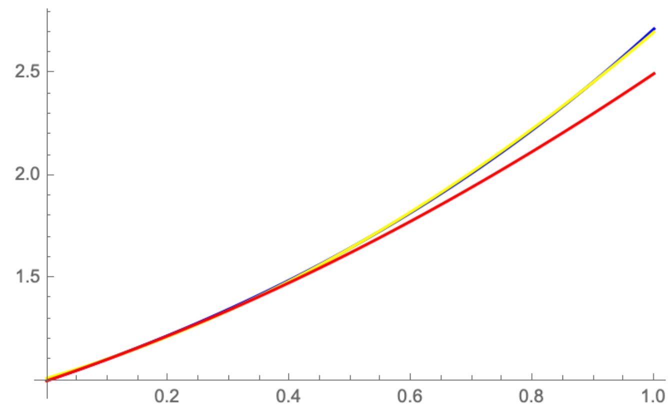
For contrast, I've added in the Taylor series $1 + x + x^2/2$:



How good is this? Here they are together



For contrast, I've added in the Taylor series $1 + x + x^2/2$:



Summary: Taking orthogonal projections onto a space of functions, even a low dimensional one like quadratics, can give great approximations!