LECTURE 7

Gram-Schmidt orthogonalisation

7.1 Orthogonalisation

It is clear from previous lectures that the availability of an orthogonal basis allows for convenient evaluations or derivations of results. However, it is not immediately clear how one can come across such a useful set: often, standard bases fail to be orthogonal with respect to a given inner product.

Example 7.1 Let $V = \mathcal{P}_n([-1,1])$ be the space of polynomials defined on the interval $[-1,1]$ and consider the following inner product:

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\mathrm{d}x.$$

Then the elements of the power basis $B = \{1, x, \ldots, x^n\}$ satisfy

$$\left\langle x^i, x^j\right\rangle = \int_{-1}^{1} x^{i+j}\mathrm{d}x = \left[\frac{x^{i+j+1}}{i+j+1}\right]_{-1}^{1} = \frac{1-(-1)^{i+j+1}}{i+j+1} = 0, \qquad \text{provided } i+j \text{ is odd.}$$

To devise an orthogonal basis for a finite dimensional inner product space V, we can employ the concept of orthogonal decompositions. To see how this works, Let U be a subspace of an inner product space V spanned by two non-orthogonal vectors $\mathbf{u}, \mathbf{v}$. Consider now the orthogonal decomposition

$$\mathbf{v} = \mathbf{v}_{\mathbf{u}}^{\|} + \mathbf{v}_{\mathbf{u}}^{\perp}.$$

Then $\left\{\mathbf{u}, \mathbf{v}_{\mathbf{u}}^{\perp}\right\}$ is an orthogonal set. Moreover, $\operatorname{span}\{\mathbf{u}, \mathbf{v}\} = \operatorname{span}\left\{\mathbf{u}, \mathbf{v}_{\mathbf{u}}^{\perp}\right\}$. As an orthogonal spanning set, $\left\{\mathbf{u}, \mathbf{v}_{\mathbf{u}}^{\perp}\right\}$ is a basis for U. In other words, we generated an orthogonal basis for U, starting from a generic basis. Can we generalise this approach? Consider the subspace spanned by the non-orthogonal vectors $\mathbf{u}, \mathbf{v}, \mathbf{w}$. Let $W = \operatorname{span}\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$. By the previous discussion, we also have $W = \operatorname{span}\left\{\mathbf{u}, \mathbf{v}^{\perp}, \mathbf{w}\right\}$. Moreover, U is a subspace of W so that we can use Theorem 6.7 to write down the orthogonal decomposition

$$\mathbf{w} = \mathbf{w}_U^{\|} + \mathbf{w}_U^{\perp},$$

where, by the proof of the theorem,

$$\mathbf{w}_U^{\perp} = \mathbf{w} - \frac{\langle \mathbf{w}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} - \frac{\left\langle \mathbf{w}, \mathbf{v}^{\perp}\right\rangle}{\left\langle \mathbf{v}^{\perp}, \mathbf{v}^{\perp}\right\rangle}\mathbf{v}^{\perp}.$$

As before, we note that $\{\mathbf{u}, \mathbf{v}^\perp, \mathbf{w}^\perp\}$ is a spanning orthogonal set, hence a basis for W.

Assume now that U is a proper subspace of V and (somehow) we already have an orthogonal basis B for U. Since the orthogonal complement U^T contains vectors orthogonal to U, these will also be orthogonal to the elements in B. We can then select one such vector and add it to the set B. This will result in an orthogonal basis for a subspace of increased dimension (by one). Continuing in this way, we eventually generate an orthogonal set of dimesion n, which will then have to be a basis for V. This procedure is guaranteed to work due to the existence and uniqueness result of Theorem 6.7. An iterative procedure for the above approach is included below.

Algorithm: generic orthogonalisation: compute orthogonal basis B_n

choose any basis $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of V
define $\mathbf{u}_1 := \mathbf{v}_1; B_1 = \{\mathbf{u}_1\}$
for $k = 1, 2, \dots, n-1$
  define $U_k = \text{span} B_k$
  generate $\mathbf{u}_{k+1} \in U_k^\perp$ using $\mathbf{v}_{k+1}$
  define $B_{k+1} = B_k \cup \{\mathbf{u}_{k+1}\}$
end
return orthogonal basis B_n for V

Note that the output B_n of the above algorithm is an orthogonal set with n elements, as we append to the initial (single-vector) set B_1 another $n-1$ vectors. Hence B_n is a basis.

R If we run the above algorithm for $k = m < n-1$ steps, we will construct an orthogonal basis B_m for a subspace of U_m of V.

7.2 Gram-Schmidt procedure

The only instruction that needs to be described is that referring to the construction of the vectors $\mathbf{u}_k$. By Theorem 6.7, in order to construct a vector orthogonal to U we need to

- choose a vector $\mathbf{v}$;
- subtract from it the orthogonal projections onto each of the basis elements of U.

In the case of the Gram-Schmidt procedure for orthogonalisation, the vector $\mathbf{v}$ is chosen to be $\mathbf{v}_k$, while the basis elements of U are chosen to be the previously computed orthogonal vectors. The resulting procedure is known as the Gram-Schmidt procedure (or process).

Algorithm: Gram-Schmidt orthogonalisation: compute orthogonal basis B_n

choose any basis $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of V
define $\mathbf{u}_1 := \mathbf{v}_1; B_1 = \{\mathbf{u}_1\}$
for $k = 1, 2, \dots, n-1$

$$\mathbf{u}_{k+1} = \mathbf{v}_{k+1} - \sum_{i=1}^{k} \frac{\langle \mathbf{v}_{k+1}, \mathbf{u}_i \rangle}{\langle \mathbf{u}_i, \mathbf{u}_i \rangle} \mathbf{u}_i$$

  define $B_{k+1} = B_k \cup \{\mathbf{u}_{k+1}\}$
end
return orthogonal basis B_n for V

We can re-write the above algorithm to include normalisation of the new basis vectors. This results in a somewhat simplified algorithm, due to the fact that the previously generated basis elements are unit vectors, so that $\langle \mathbf{u}_i, \mathbf{u}_i \rangle = 1$.

Algorithm: Gram-Schmidt orthonormalisation: compute orthonormal basis B_n

choose any basis $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of V
define $\mathbf{u}_1 := \mathbf{v}_1; B_1 = \{\mathbf{u}_1\}$
for $k = 1, 2, \dots, n-1$

$$\mathbf{u}_{k+1} = \mathbf{v}_{k+1} - \sum_{i=1}^{k} \langle \mathbf{v}_{k+1}, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i$$

$\widehat{\mathbf{u}}_{k+1} = \mathbf{u}_{k+1} / \|\mathbf{u}_{k+1}\|$
define $B_{k+1} = B_k \cup \{\widehat{\mathbf{u}}_{k+1}\}$
end
return orthonormal basis B_n for V

Let us write the algorithm explicitly, indicating the first few steps and also the generic step from the above algorithm.

$$\begin{cases} \mathbf{u}_1 = \mathbf{v}_1 & \widehat{\mathbf{u}}_1 := \mathbf{u}_1 / \|\mathbf{u}_1\|, \\ \mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 & \widehat{\mathbf{u}}_2 := \mathbf{u}_2 / \|\mathbf{u}_2\|, \\ \mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_2 \rangle \widehat{\mathbf{u}}_2 & \widehat{\mathbf{u}}_3 := \mathbf{u}_3 / \|\mathbf{u}_2\|, \\ \qquad \cdots\cdots\cdots\cdots\cdots\cdots & \\ \mathbf{u}_k = \mathbf{v}_k - \sum_{i=1}^{k-1} \langle \mathbf{v}_k, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i & \widehat{\mathbf{u}}_k := \mathbf{u}_k / \|\mathbf{u}_k\|, \\ \qquad \cdots\cdots\cdots\cdots\cdots\cdots & \\ \mathbf{u}_n = \mathbf{v}_n - \sum_{i=1}^{n-1} \langle \mathbf{v}_n, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i & \widehat{\mathbf{u}}_n := \mathbf{u}_n / \|\mathbf{u}_n\|. \end{cases}$$

We note that at each step we need to use quantities (unit basis elements) computed in previous steps. Finally, we note that

$$\mathbf{v}_k = \mathbf{u}_k + \sum_{i=1}^{k-1} \langle \mathbf{v}_k, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i = \|\mathbf{u}_k\| \widehat{\mathbf{u}}_k + \sum_{i=1}^{k-1} \langle \mathbf{v}_k, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i = \sum_{i=1}^{k} \langle \mathbf{v}_k, \widehat{\mathbf{u}}_i \rangle \widehat{\mathbf{u}}_i,$$

since $\|\mathbf{u}_k\| = \langle \mathbf{u}_k, \widehat{\mathbf{u}}_k \rangle = \langle \mathbf{v}_k, \widehat{\mathbf{u}}_k \rangle$. Hence, every $\mathbf{v}_k$ can be written as a linear combination of only the first k vectors in the newly computed orthonormal basis.

We end this lecture with two examples involving standard inner product spaces.

7.3 Examples

Example 7.2 Consider the following non-orthogonal basis for $\mathbb{R}^3$

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Let $\langle \cdot, \cdot \rangle$ denote the Euclidean inner-product. Using the Gram-Schmidt iterative procedure we find

$$\mathbf{u}_1 = \mathbf{v}_1, \quad \widehat{\mathbf{u}}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix};$$

$$\mathbf{u}_2 = \mathbf{v}_2 - \langle \mathbf{v}_2, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \widehat{\mathbf{u}}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix};$$

$$\mathbf{u}_3 = \mathbf{v}_3 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_1 \rangle \widehat{\mathbf{u}}_1 - \langle \mathbf{v}_3, \widehat{\mathbf{u}}_2 \rangle \widehat{\mathbf{u}}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \widehat{\mathbf{u}}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

The above example confirms that orthonormal bases are not unique: the canonical basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is also orthonormal with respect to the Euclidean inner product.

Example 7.3 Consider the power basis $\{1, x, x^2\} =: \{p_1, p_2, p_3\}$ for $\mathcal{P}_2([-1, 1])$. Let us apply the Gram-Schmidt procedure without normalisation using the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(x) q(x) \mathrm{d}x.$$

We will denote the new basis by q_1, q_2, q_3. We find

$$q_1 = p_1 = 1;$$

$$q_2 = p_2 - \frac{\langle p_2, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 = x$$

since

$$\langle p_2, q_1 \rangle = \int_{-1}^{1} x \mathrm{d}x = 0$$

and finally

$$q_3 = p_3 - \frac{\langle p_3, q_2 \rangle}{\langle q_2, q_2 \rangle} q_2 - \frac{\langle p_3, q_1 \rangle}{\langle q_1, q_1 \rangle} q_1 = x^2 - \frac{1}{3}$$

since

$$\langle p_3, q_2 \rangle = \int_{-1}^{1} x^3 \mathrm{d}x = 0, \quad \langle p_3, q_1 \rangle = \int_{-1}^{1} x^2 = \frac{2}{3}, \quad \langle q_1, q_1 \rangle = \int_{-1}^{1} 1 \mathrm{d}x = 2.$$

Therefore, the resulting orthogonal basis is $\{1, x, x^2 - \frac{1}{3}\}$.