5.6 UNITARY AND ORTHOGONAL MATRICES

The purpose of this section is to examine square matrices whose columns (or rows) are orthonormal. The standard inner product and the euclidean 2-norm are the only ones used in this section, so distinguishing subscripts are omitted.

Unitary and Orthogonal Matrices

- A ***unitary matrix*** is defined to be a *complex* matrix $\mathbf{U}_{n\times n}$ whose columns (or rows) constitute an orthonormal basis for $\mathcal{C}^n$.
- An ***orthogonal matrix*** is defined to be a *real* matrix $\mathbf{P}_{n\times n}$ whose columns (or rows) constitute an orthonormal basis for $\Re^n$.

Unitary and orthogonal matrices have some nice features, one of which is the fact that they are easy to invert. To see why, notice that the columns of $\mathbf{U}_{n\times n} = \big(\mathbf{u}_1 \,|\, \mathbf{u}_2 \,|\, \cdots \,|\mathbf{u}_n\big)$ are an orthonormal set if and only if

$$[\mathbf{U}^*\mathbf{U}]_{ij} = \mathbf{u}_i^*\mathbf{u}_j = \begin{cases} 1 & \text{when } i=j, \\ 0 & \text{when } i\neq j, \end{cases} \Longleftrightarrow \mathbf{U}^*\mathbf{U} = \mathbf{I} \Longleftrightarrow \mathbf{U}^{-1} = \mathbf{U}^*.$$

Notice that because $\mathbf{U}^*\mathbf{U} = \mathbf{I} \Longleftrightarrow \mathbf{U}\mathbf{U}^* = \mathbf{I}$, the columns of $\mathbf{U}$ are orthonormal if and only if the rows of $\mathbf{U}$ are orthonormal, and this is why the definitions of unitary and orthogonal matrices can be stated either in terms of orthonormal columns or orthonormal rows.

Another nice feature is that multiplication by a unitary matrix doesn't change the length of a vector. Only the direction can be altered because

$$\|\mathbf{U}\mathbf{x}\|^2 = \mathbf{x}^*\mathbf{U}^*\mathbf{U}\mathbf{x} = \mathbf{x}^*\mathbf{x} = \|\mathbf{x}\|^2 \quad \forall\ \mathbf{x}\in\mathcal{C}^n. \tag{5.6.1}$$

Conversely, if (5.6.1) holds, then $\mathbf{U}$ must be unitary. To see this, set $\mathbf{x} = \mathbf{e}_i$ in (5.6.1) to observe $\mathbf{u}_i^*\mathbf{u}_i = 1$ for each i, and then set $\mathbf{x} = \mathbf{e}_j + \mathbf{e}_k$ for $j \neq k$ to obtain $0 = \mathbf{u}_j^*\mathbf{u}_k + \mathbf{u}_k^*\mathbf{u}_j = 2\,\mathrm{Re}\left(\mathbf{u}_j^*\mathbf{u}_k\right)$. By setting $\mathbf{x} = \mathbf{e}_j + \mathrm{i}\mathbf{e}_k$ in (5.6.1) it also follows that $0 = 2\,\mathrm{Im}\left(\mathbf{u}_j^*\mathbf{u}_k\right)$, so $\mathbf{u}_j^*\mathbf{u}_k = 0$ for each $j \neq k$, and thus (5.6.1) guarantees that $\mathbf{U}$ is unitary.

In the case of orthogonal matrices, everything is real so that $(\star)^*$ can be replaced by $(\star)^T$. Below is a summary of these observations.