# Further Mathematical Methods (Linear Algebra) 2002

# Lecture 11: 'Special' Real and Complex Matrices

### Notation

We denote the complex conjugate of a matrix $\mathsf{A}$ by $\mathsf{A}^*$ and the complex conjugate transpose of a matrix $\mathsf{A}$ by $\mathsf{A}^\dagger$. Consequently, we can see that $\mathsf{A}^\dagger = (\mathsf{A}^*)^t = (\mathsf{A}^t)^*$, and indeed, that $(\mathsf{A}^\dagger)^\dagger = \mathsf{A}$. Notice that this is *not* the notation used in the book by Ostaszewski where the complex conjugate of a matrix $\mathsf{A}$ is denoted by $\overline{\mathsf{A}}$ and the complex conjugate transpose of a matrix is denoted by $\mathsf{A}^*$.

### Convention

Throughout this part of the course, we will assume that we are working with the real Euclidean inner product when in $\mathbb{R}^n$ (i.e. if $\mathbf{x} = [x_1, \ldots, x_n]^t$ and $\mathbf{y} = [y_1, \ldots, y_n]^t$ are vectors in $\mathbb{R}^n$, then $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n$) and the complex Euclidean inner product when in $\mathbb{C}^n$ (i.e. if $\mathbf{x} = [x_1, \ldots, x_n]^t$ and $\mathbf{y} = [y_1, \ldots, y_n]^t$ are vectors in $\mathbb{C}^n$, then $\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1^* + \cdots + x_n y_n^*$). Further, note that for such vectors in $\mathbb{C}^n$, $\mathbf{x}^\dagger \mathbf{y} = x_1^* y_1 + \cdots + x_n^* y_n = (x_1 y_1^* + \cdots + x_n y_n^*)^* = \langle \mathbf{x}, \mathbf{y} \rangle^*$.

### Definitions

Here is a summary of the special types of complex matrix (and their real analogues) that we shall meet in the lecture. If we have an $n \times n$ matrix $\mathsf{A}$ with complex entries, i.e. the vectors that form the column space of the matrix are in $\mathbb{C}^n$, we say that

- $\mathsf{A}$ is *Hermitian* if $\mathsf{A}^\dagger = \mathsf{A}$.
- $\mathsf{A}$ is *unitary* if $\mathsf{A}\mathsf{A}^\dagger = \mathsf{I}$.

However, if we have an $n \times n$ matrix $\mathsf{A}$ with real entries, i.e. the vectors that form the the column space of the matrix are in $\mathbb{R}^n$, we say that

- $\mathsf{A}$ is *symmetric* if $\mathsf{A}^t = \mathsf{A}$.
- $\mathsf{A}$ is *orthogonal* if $\mathsf{A}\mathsf{A}^t = \mathsf{I}$.

As $\mathbb{R} \subseteq \mathbb{C}$ (and, $\mathbb{R}^n \subseteq \mathbb{C}^n$), a matrix with real entries is a special case of a matrix with complex entries (as if a matrix $\mathsf{A}$ has real entries, $\mathsf{A}^* = \mathsf{A}$). It should be clear that all symmetric matrices are Hermitian (as for a symmetric matrix $\mathsf{A}$, $\mathsf{A}^\dagger = (\mathsf{A}^*)^t = \mathsf{A}^t = \mathsf{A}$) and all orthogonal matrices are unitary (as for an orthogonal matrix $\mathsf{A}$, $\mathsf{A}\mathsf{A}^\dagger = \mathsf{A}(\mathsf{A}^*)^t = \mathsf{A}\mathsf{A}^t = \mathsf{I}$). Consequently, most of our theorems about Hermitian and unitary matrices can be translated into theorems about symmetric and orthogonal matrices [respectively] by using the following rules:

| | | |
|---|---|---|
| the complex matrix $\mathsf{A}$ (i.e. $CS(\mathsf{A}) \subseteq \mathbb{C}^n$) | becomes | the real matrix $\mathsf{A}$ (i.e. $CS(\mathsf{A}) \subseteq \mathbb{R}^n$) |
| the complex conjugate transpose of $\mathsf{A}$ (i.e. $\mathsf{A}^\dagger$) | becomes | the transpose of $\mathsf{A}$ (i.e. $\mathsf{A}^t$) |
| the unitary matrix $\mathsf{A}$ (i.e. $\mathsf{A}\mathsf{A}^\dagger = \mathsf{I}$) | becomes | the orthogonal matrix $\mathsf{A}$ (i.e. $\mathsf{A}\mathsf{A}^t = \mathsf{I}$) |
| the Hermitian matrix $\mathsf{A}$ (i.e. $\mathsf{A}^\dagger = \mathsf{A}$) | becomes | the symmetric matrix $\mathsf{A}$ (i.e. $\mathsf{A}^t = \mathsf{A}$) |

For example: Notice that if $\mathsf{A}$ and $\mathsf{B}$ are two complex matrices, then $(\mathsf{A}\mathsf{B})^\dagger = \mathsf{B}^\dagger \mathsf{A}^\dagger$. But, if $\mathsf{A}$ and $\mathsf{B}$ are two real matrices, then $(\mathsf{A}\mathsf{B})^t = \mathsf{B}^t \mathsf{A}^t$. We will assume this result (but, see Problem Sheet 6, Question 5).

Lastly, as unitary matrices have the property that $\mathsf{A}\mathsf{A}^\dagger = \mathsf{I}$, it should be clear that if $\mathsf{A}$ is a unitary matrix, then $\mathsf{A}^\dagger = \mathsf{A}^{-1}$. Thus, for a unitary matrix we have $\mathsf{A}^\dagger \mathsf{A} = \mathsf{I}$ as well. Similarly, if $\mathsf{A}$ is an orthogonal matrix, then $\mathsf{A}^t = \mathsf{A}^{-1}$ and $\mathsf{A}^t \mathsf{A} = \mathsf{I}$ too.

**Unitary Diagonalisation**

As we are often keen to diagonalise matrices, it is useful to note that a particularly nice form of diagonalisation is available in some cases. To see this, observe that:

**Theorem:** A [square] complex matrix $\mathsf{A}$ has an orthonormal set of eigenvectors iff there exists a unitary matrix $\mathsf{P}$ such that the matrix $\mathsf{P}^\dagger\mathsf{A}\mathsf{P}$ is diagonal. If such a $\mathsf{P}$ exists, then $\mathsf{A}$ is said to be *unitarily diagonalisable.*

**Proof:** It should be clear that a matrix $\mathsf{P}$ is unitary iff the column vectors of $\mathsf{P}$ form an orthonormal set. To see this, observe that if the column vectors of a $3 \times 3$ matrix $\mathsf{P}$ are the vectors $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$, then the product $\mathsf{P}^\dagger\mathsf{P}$ is given by:

$$\begin{bmatrix} — & \mathbf{v}_1^\dagger & — \\ — & \mathbf{v}_2^\dagger & — \\ — & \mathbf{v}_3^\dagger & — \end{bmatrix} \begin{bmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^\dagger\mathbf{v}_1 & \mathbf{v}_1^\dagger\mathbf{v}_2 & \mathbf{v}_1^\dagger\mathbf{v}_3 \\ \mathbf{v}_2^\dagger\mathbf{v}_1 & \mathbf{v}_2^\dagger\mathbf{v}_2 & \mathbf{v}_2^\dagger\mathbf{v}_3 \\ \mathbf{v}_3^\dagger\mathbf{v}_1 & \mathbf{v}_3^\dagger\mathbf{v}_2 & \mathbf{v}_3^\dagger\mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} \langle\mathbf{v}_1,\mathbf{v}_1\rangle^* & \langle\mathbf{v}_1,\mathbf{v}_2\rangle^* & \langle\mathbf{v}_1,\mathbf{v}_3\rangle^* \\ \langle\mathbf{v}_2,\mathbf{v}_1\rangle^* & \langle\mathbf{v}_2,\mathbf{v}_2\rangle^* & \langle\mathbf{v}_2,\mathbf{v}_3\rangle^* \\ \langle\mathbf{v}_3,\mathbf{v}_1\rangle^* & \langle\mathbf{v}_3,\mathbf{v}_2\rangle^* & \langle\mathbf{v}_3,\mathbf{v}_3\rangle^* \end{bmatrix}$$

where we have used the fact mentioned earlier when we set up our convention. So, clearly, $\mathsf{P}^\dagger\mathsf{P} = \mathsf{I}$ iff the column vectors of $\mathsf{P}$ form an orthonormal set. Now, to prove the theorem, we need to show that it holds in both 'directions':
**LTR:** Assume that $\mathsf{A}$ is a [square] complex matrix with an orthonormal set of eigenvectors. This means that $\mathsf{A}$ is diagonalisable (as orthonormal vectors are linearly independent) and so there exists a $\mathsf{P}$ such that the matrix $\mathsf{P}^{-1}\mathsf{A}\mathsf{P}$ is diagonal. This matrix $\mathsf{P}$ is the matrix whose columns are the eigenvectors of $\mathsf{A}$, and as these are orthonormal, we can see from the above observation that $\mathsf{P}$ will be unitary. For unitary matrices, $\mathsf{P}^{-1} = \mathsf{P}^\dagger$, and therefore there exists a $\mathsf{P}$ such that $\mathsf{P}^\dagger\mathsf{A}\mathsf{P}$ is diagonal (as required).
**RTL:** Assume that there exists a unitary matrix $\mathsf{P}$ such that $\mathsf{P}^\dagger\mathsf{A}\mathsf{P}$ is diagonal. Now, as $\mathsf{P}$ is unitary, from the observation above, its column vectors must be orthonormal. Further, as $\mathsf{P}^\dagger\mathsf{A}\mathsf{P}$ is diagonal and unitary matrices are such that $\mathsf{P}^{-1} = \mathsf{P}^\dagger$, we have a matrix $\mathsf{P}^{-1}\mathsf{A}\mathsf{P}$ that is diagonal. Now, this is the result of our standard diagonalisation procedure, and so $\mathsf{P}$ is a matrix whose columns are the eigenvectors of the [in general, complex and square] matrix $\mathsf{A}$. But we have seen that these column vectors are orthonormal, and so the eigenvectors of $\mathsf{A}$ must form an orthonormal set (as required).

It is now convenient to define another special kind of complex matrix, namely:

**Definition:** A [square] complex matrix $\mathsf{A}$ is *normal* iff $\mathsf{A}^\dagger\mathsf{A} = \mathsf{A}\mathsf{A}^\dagger$.

and this leads to the following useful theorem:

**Theorem:** A [square] complex matrix is normal iff $\mathsf{A}$ is unitarily diagonalisable.

**Proof:** We have to establish that this theorem holds by proving it in both 'directions'.
**RTL:** We assume that $\mathsf{A}$ is unitarily diagonalisable, that is, there exists a unitary matrix $\mathsf{P}$ such that the matrix $\mathsf{P}^\dagger\mathsf{A}\mathsf{P} = \mathsf{D}$ is diagonal. Now $\mathsf{P}$ is unitary, and so $\mathsf{P}^\dagger = \mathsf{P}^{-1}$, which means that $\mathsf{A} = \mathsf{P}\mathsf{D}\mathsf{P}^\dagger$, and taking the complex conjugate transpose of this we get $\mathsf{A}^\dagger = (\mathsf{P}\mathsf{D}\mathsf{P}^\dagger)^\dagger = \mathsf{P}\mathsf{D}^\dagger\mathsf{P}^\dagger$. Further, using these, we can see that

$$\mathsf{A}\mathsf{A}^\dagger = (\mathsf{P}\mathsf{D}\mathsf{P}^\dagger)(\mathsf{P}\mathsf{D}^\dagger\mathsf{P}^\dagger) = \mathsf{P}\mathsf{D}\mathsf{D}^\dagger\mathsf{P}^\dagger \quad \text{and} \quad \mathsf{A}^\dagger\mathsf{A} = (\mathsf{P}\mathsf{D}^\dagger\mathsf{P}^\dagger)(\mathsf{P}\mathsf{D}\mathsf{P}^\dagger) = \mathsf{P}\mathsf{D}^\dagger\mathsf{D}\mathsf{P}^\dagger$$

where, again, we have used the fact that $\mathsf{P}^\dagger = \mathsf{P}^{-1}$. But, for the diagonal matrix $\mathsf{D} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ it should be clear that $\mathsf{D}\mathsf{D}^\dagger = \mathsf{D}^\dagger\mathsf{D}$ as

$$\mathsf{D}\mathsf{D}^\dagger = \text{diag}(|\lambda_1|^2, |\lambda_2|^2, \ldots, |\lambda_n|^2) = \mathsf{D}^\dagger\mathsf{D}$$

Consequently, $\mathsf{A}\mathsf{A}^\dagger = \mathsf{A}^\dagger\mathsf{A}$, and so the matrix $\mathsf{A}$ is normal (as required).
**LTR:** This is quite hard, and so due to considerations of space, we shall omit it here.