This form insures that $(\mathbf{T}-\lambda_1\mathbf{I})^{a_1}(\mathbf{T}-\lambda_2\mathbf{I})^{a_2}\cdots(\mathbf{T}-\lambda_k\mathbf{I})^{a_k}=\mathbf{0}$. The characteristic equation for $\mathbf{A}$ is $p(\lambda)=(\lambda-\lambda_1)^{a_1}(\lambda-\lambda_2)^{a_2}\cdots(\lambda-\lambda_k)^{a_k}=0$, so

$$\begin{aligned}\mathbf{U}^*p(\mathbf{A})\mathbf{U}&=\mathbf{U}^*(\mathbf{A}-\lambda_1\mathbf{I})^{a_1}(\mathbf{A}-\lambda_2\mathbf{I})^{a_2}\cdots(\mathbf{A}-\lambda_k\mathbf{I})^{a_k}\mathbf{U}\\&=(\mathbf{T}-\lambda_1\mathbf{I})^{a_1}(\mathbf{T}-\lambda_2\mathbf{I})^{a_2}\cdots(\mathbf{T}-\lambda_k\mathbf{I})^{a_k}=\mathbf{0},\end{aligned}$$

and thus $p(\mathbf{A})=\mathbf{0}$. **Note:** A completely different approach to the Cayley–Hamilton theorem is discussed on p. 532.

Schur's theorem is not the complete story on triangularizing by similarity. By allowing nonunitary similarity transformations, the structure of the upper-triangular matrix $\mathbf{T}$ can be simplified to contain zeros everywhere except on the diagonal and the superdiagonal (the diagonal immediately above the main diagonal). This is the Jordan form developed on p. 590, but some of the seeds are sown here.

Multiplicities

For $\lambda\in\sigma(\mathbf{A})=\{\lambda_1,\lambda_2,\ldots,\lambda_s\}$, we adopt the following definitions.

- The ***algebraic multiplicity*** of λ is the number of times it is repeated as a root of the characteristic polynomial. In other words, $alg\ mult_{\mathbf{A}}(\lambda_i)=a_i$ if and only if $(x-\lambda_1)^{a_1}\cdots(x-\lambda_s)^{a_s}=0$ is the characteristic equation for $\mathbf{A}$.
- When $alg\ mult_{\mathbf{A}}(\lambda)=1$, λ is called a ***simple eigenvalue.***
- The ***geometric multiplicity*** of λ is $\dim N(\mathbf{A}-\lambda\mathbf{I})$. In other words, $geo\ mult_{\mathbf{A}}(\lambda)$ is the maximal number of linearly independent eigenvectors associated with λ.
- Eigenvalues such that $alg\ mult_{\mathbf{A}}(\lambda)=geo\ mult_{\mathbf{A}}(\lambda)$ are called ***semisimple eigenvalues*** of $\mathbf{A}$. It follows from (7.2.2) on p. 511 that a simple eigenvalue is always semisimple, but not conversely.

Example 7.2.3

The algebraic and geometric multiplicity need not agree. For example, the nilpotent matrix $\mathbf{A}=\begin{pmatrix}0&1\\0&0\end{pmatrix}$ in (7.2.1) has only one distinct eigenvalue, $\lambda=0$, that is repeated twice, so $alg\ mult_{\mathbf{A}}(0)=2$. But

$$\dim N(\mathbf{A}-0\mathbf{I})=\dim N(\mathbf{A})=1 \implies geo\ mult_{\mathbf{A}}(0)=1.$$

In other words, there is only one linearly independent eigenvector associated with $\lambda=0$ even though $\lambda=0$ is repeated twice as an eigenvalue.

Example 7.2.3 shows that $geo\ mult_{\mathbf{A}}(\lambda)<alg\ mult_{\mathbf{A}}(\lambda)$ is possible. However, the inequality can never go in the reverse direction.