THE AM≥GM THEOREM

The purpose of this supplement is to explain why the algebraic multiplicity of an eigenvalue is always greater than or equal to its geometric multiplicity. We will need two lemmas.

Lemma 1 (Linearly independent vectors extend to a basis)**.** *Let $v_1, \ldots, v_r$ be linearly independent vectors in $\mathbf{R}^n$. There exist vectors $v_{r+1}, \ldots, v_n$ such that $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ is a basis for $\mathbf{R}^n$.*

Proof. Let A be the $r \times n$ matrix with rows $v_1^T, \ldots, v_r^T$. Since $\{v_1, \ldots, v_r\}$ is linearly independent, it spans an r-dimensional subspace, so $\dim(\mathrm{Row}(A)) = r$. By rank-nullity, we have $\dim(\mathrm{Nul}(A)) = n - r$. Let $\{v_{r+1}, \ldots, v_n\}$ be any basis for $\mathrm{Nul}(A)$. Recall that $\mathrm{Nul}(A)$ is the orthogonal complement of $V = \mathrm{Row}(A)$. Let x be any vector in $\mathbf{R}^n$, and let $x = x_V + x_{V^\perp}$ be its orthogonal decomposition relative to $V = \mathrm{Row}(A)$. Since $x_V \in \mathrm{Row}(A) = \mathrm{Span}\{v_1, \ldots, v_r\}$ we can solve the equation

$$x_V = a_1 v_1 + \cdots + a_r v_r,$$

and since $x_{V^\perp} \in V^\perp = \mathrm{Nul}(A) = \mathrm{Span}\{v_{r+1}, \ldots, v_n\}$, we can solve the equation

$$x_{V^\perp} = a_{r+1} v_{r+1} + \cdots + a_n v_n.$$

Summing the previous two equations gives

$$x = x_V + x_{V^\perp} = a_1 v_1 + \cdots + a_r v_r + a_{r+1} v_{r+} + \cdots + a_n v_n.$$

This shows that any vector in $\mathbf{R}^n$ is in $\mathrm{Span}\{v_1, \ldots, v_n\}$. By the basis theorem, $\{v_1, \ldots, v_n\}$ is a basis for $\mathbf{R}^n$. □

The second lemma was an exercise in Homework 10.

Lemma 2 (Similar matrices have the same characteristic polynomial)**.** *Let C be an invertible $n \times n$ matrix and let D be any $n \times n$ matrix. Then D and CDC^{-1} have the same characteristic polynomial.*

Theorem 3 (AM≥GM)**.** *Let A be an $n \times n$ matrix. Suppose that λ_1 is an eigenvalue of A with geometric multiplicity r. Then the algebraic multiplicity of λ_1 is at least r.*

Proof. We proceed by a "partial diagonalization": we find an invertible matrix C such that the first r columns of $D = C^{-1}AC$ are $\lambda_1 e_1, \ldots, \lambda_1 e_r$. Let $\{w_1, \ldots, w_r\}$ be a basis for the λ_1-eigenspace. Extend this collection to a basis $\{w_1, \ldots, w_n\}$ for $\mathbf{R}^n$ using Lemma 1, and let C be the $n \times n$ matrix with columns $w_1, \ldots, w_n$. Note that C is invertible as it is a square matrix with linearly independent columns. Let $D = C^{-1}AC$, so that $A = CDC^{-1}$. For any $i = 1 \ldots, n$ we have $Ce_i = w_i$ (the ith column of C), so $e_i = C^{-1}w_i$. It follows that for $1 \le i \le r$ we have

$$De_i = (C^{-1}AC)e_i = C^{-1}A(Ce_i) = C^{-1}Aw_i = C^{-1}(Aw_i) = C^{-1}(\lambda_1 w_i) = \lambda_1 C^{-1} w_i = \lambda_1 e_i.$$

Therefore, the ith column of D is $\lambda_1 e_i$, as we wanted.

Since the first r columns of D are $\lambda_1 e_1, \ldots, \lambda_1 e_r$, the matrix $D - \lambda I_n$ has the form

$$D - \lambda I_n = \begin{pmatrix} \lambda_1 - \lambda & 0 & 0 & \cdots & 0 & * & \cdots & * \\ 0 & \lambda_1 - \lambda & 0 & \cdots & 0 & * & \cdots & * \\ 0 & 0 & \lambda_1 - \lambda & \cdots & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_1 - \lambda & * & \cdots & * \\ 0 & 0 & 0 & \cdots & 0 & * - \lambda & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & * & \cdots & * - \lambda \end{pmatrix}$$

(We do not know what the last $n-r$ columns of D contain, so they're denoted "$*$" above.) Expanding cofactors along the first column, then the second, and so on, we see that the characteristic polynomial of D has the form

$$p(\lambda) = \det(D - \lambda I_n) = (\lambda_1 - \lambda)^r \det \begin{pmatrix} * - \lambda & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * - \lambda \end{pmatrix}.$$

It follows that $(\lambda_1 - \lambda)^r$ divides $p(\lambda)$, so that the algebraic multiplicity of λ_1 as an eigenvalue of D is at least r. But D and $A = CDC^{-1}$ have the same characteristic polynomial by Lemma 2, so the same is true of A. □