高等代数杂谈——关于特征值、可对角化矩阵、正交变换和二次型

作者：21 AI 钟韵骅

笔者非常喜欢高等代数，但高等代数不喜欢笔者，所以笔者学的很烂。在和bhgg的协♂商中，决定写此系列，谈一谈从矩阵到二次型中一条隐藏的支线剧情。

~~如果说高等代数的主线像旅行者和戴因斯雷布之间的故事，那这篇文章的目的就是带领大家探索层岩巨渊。~~

从一个问题谈起：如何快速计算矩阵的幂？

既然矩阵可以自乘，则这必然是一个方阵。不妨假设是n×n的方阵，我们知道矩阵相乘的规则：

$$a_{ij} = \sum_{k=1}^{n} a_{ik} \times a_{kj}$$

计算一个数的计算次数是n次，时间复杂度为O(n)，因此矩阵相乘的时间复杂度为O(n^3)，这在n很大的时候是一个不太可接受的复杂度。

算法方面，矩阵的幂可以用“矩阵快速幂”算法进行快速的计算，具体的原理是用了二分的乘积来加快计算。本篇文章中，笔者不打算阐述这个算法的详细细节，而是打算从数学方面快速的计算矩阵的幂。

（PS：或许会出算法番外篇，不过笔者也没学过这个算法，可能会拖一拖...）

考虑对角形矩阵Λ，Λ的幂相乘比较容易，手推一下递推式知：

$$\text{如果}\Lambda\text{的第}i\text{行第}j\text{列元素是}a_{ij}\text{，而}\Lambda\text{又是对角矩阵，故}i \neq j\text{时}a_{ij} = 0$$
$$\text{计算知：}\Lambda^n = \begin{cases} a_{ij}^n, & \text{i=j} \\ 0, & \text{otherwise} \end{cases}$$

我们就将矩阵的每个元素的积算出来，即可直接得到Λ的幂。

但更多的矩阵并不是对角型矩阵，**如何联系可对角化矩阵和任意的矩阵呢？**

考虑矩阵的相似，即满足如下条件的关系：

$$\text{若存在矩阵}P\text{，有}P^{-1}AP = B\text{，即}A\text{和}B\text{相似，记为}A \sim B$$

如果任意矩阵P和对角型矩阵Λ相似，则可以很容易的计算P的幂：

$$\text{假设}P \sim \Lambda\text{，即存在}Q\text{使得}Q^{-1}\Lambda Q = P\text{，则}P^n = (Q^{-1}\Lambda Q)^n$$
$$P^n = (Q^{-1}\Lambda Q)^n = Q^{-1}\Lambda Q Q^{-1}\Lambda Q \ldots Q^{-1}\Lambda Q$$
$$= Q^{-1}\Lambda^n Q$$
$$\text{若}R = Q^{-1}\text{，则}R\Lambda R^{-1} = P\text{，则有}\Lambda = R^{-1}PR\text{，因此相似关系具有对称性。}$$

那P和Λ是否相似呢？我们进行如下推理：

$$\text{如果存在}R\text{，满足：}R^{-1}PR = \Lambda \iff PR = R\Lambda$$
$$\text{其中}\Lambda\text{是对角形矩阵，因此假设}R = \{\beta_1, \beta_2, \ldots, \beta_n\}$$
$$\text{即}R\Lambda = \{a_{11}\beta_1, a_{22}\beta_2, \ldots, a_{nn}\beta_n\}$$
$$\text{而}PR = \{P\beta_1, P\beta_2, \ldots, P\beta_n\}$$

$$\text{以第一列为例，注意到：}P\beta_1 = a_{11}\beta_1$$
$$\text{这符合特征值的定义！}$$

因此不难知道：

$$a_{11}, a_{22} \ldots, a_{nn}\text{分别是}P\text{的特征值，而}\beta_1, \ldots, \beta_n\text{是特征值对应的特征向量。}$$

因此可知，矩阵P可对角化的充分必要条件是，P有n个（可以相同）的特征值，以及n个特征向量。R^{-1}的存在要求矩阵满秩，即这n个特征向量是线性无关的。可什么时候有n个线性无关的特征向量呢？特征值一定有n个，还是说和矩阵的秩有关？

回想起特征值的定义，假设矩阵为A，特征值为λ，对应的特征向量为α，则$A\alpha = \lambda\alpha$。

$$A\alpha = \lambda\alpha = \lambda I\alpha$$
$$\iff (A - \lambda I)\alpha = 0$$

满秩矩阵和非零向量相乘不可能为0向量，因此可知$(A - \lambda I)$非满秩，即$|A - \lambda I| = 0$。

令$f(\lambda) = |A - \lambda I|$，由行列式的值可知，$f(\lambda)$是个n次的多项式。原矩阵需要有n个特征值，$f(\lambda)$就要有n个根。

由代数基本定理知，n次多项式总存在n个复根，也就是说，原矩阵总存在n个复特征值，但或许并没有n个实特征值。也就是说，如果我们想从多项式方面入手，我们可以通过多项式的根来确定特征值。

不过，我们也可以从另一个方向入手，即线性无关的特征向量起手。对于每个特征值，满足此条件的x都是可能的特征向量。

$$(A - \lambda I)x = 0$$

这个方程其实是齐次线性方程组，x的维数即是对应线性方程组的解空间维度（我们定义这个空间叫特征子空间），即对于不同的特征子空间$U_1, U_2 \ldots U_k$，需要满足方程：

$$\sum_{i=1}^{k} Dim(U_i) = n$$

至此，我们大概明白了如何快速求矩阵的幂，以及可以快速求幂的条件，这无疑非常有用。**那这些知识还有什么用呢？**

让我们继续飞快的向后走，来到二次型的章节。可以看出，二次型的标准型即是一个对角矩阵，但二次型相关的关系是合同关系，并非相似关系：

$$\text{若存在矩阵}P\text{，有}P'AP = B\text{，即}A\text{和}B\text{合同，记为}A \simeq B$$

那如果可以联系相似关系和合同关系，那问题就迎刃而解了。有没有一种特殊的矩阵，它既可以作为合同的条件，又可以做为相似的条件呢？

$$\text{如果}Q^{-1} = Q', \text{即}QQ' = I\text{，那}Q\text{正好符合正交矩阵的定义!}$$

看来正交矩阵是研究这个问题的核心，也难怪正交变换会是二次型化简的一部分了。

原来正交矩阵位于合同关系和相似关系的相交点。因此作为特殊的相似关系，我们定义新的关系，正交相似关系：

$$\text{若存在正交矩阵}T\text{，有}T^{-1}AT = B\text{，即}A\text{和}B\text{正交相似。}$$

是不是正交变换的方法就是找到正交矩阵T，把二次型矩阵化成对角形矩阵呢?

且慢，先得保证前提是否成立，也就是说，**任意一个实对称矩阵能否化成实对角形矩阵？是否n个特征值都是实数?**

从上文我们知道，特征值即是多项式的根，又从多项式板块的知识中，我们知道高维多项式的解总是单个或成对出现。因此如果存在复数根，那么一定存在一组共轭复数解。

不妨反设，存在一组复数λ_0和$\overline{\lambda_0}$，使得它们均是对称矩阵P的特征值，因为P对称，$P = P'$。假设λ_0的一个特征向量是α，有:

$$A\alpha = \lambda_0\alpha$$

$$\lambda_0\text{是复数，对两边求转置知：}\alpha'A' = \lambda_0\alpha'(1)$$

$$\text{同时，因为}A\text{是实矩阵，对两边求共轭：}A\overline{\alpha} = \overline{\lambda_0}\overline{\alpha}(2)$$

$$(1)\text{知：}\alpha'A'\overline{\alpha} = \lambda_0\alpha'\overline{\alpha}$$

$$(2)\text{知：}\alpha'A'\overline{\alpha} = \overline{\lambda_0}\alpha'\overline{\alpha}$$

$$\text{两边相减知：}(\overline{\lambda_0} - \lambda_0)\alpha'\overline{\alpha} = 0\text{，但}\alpha'\overline{\alpha} \neq 0\text{，故}\overline{\lambda_0} - \lambda_0 = 0$$

$$\text{即}\lambda_0\text{是实数。}$$

因此，对称矩阵全是实特征值，因此答案是成立的。

于是来到了最后一个问题，即：任意实对称矩阵是否正交相似于对角矩阵？如果是，如何去求正交矩阵?

这个问题的回答十分精彩，其中一个回答是：https://www.zhihu.com/question/38801697/answer/869412501，出于篇幅原因笔者无法阐述，但这一部分**十分重要。**

终于，我们铺平了所有的道路，也间接的回答了这些知识点间的一些关系。希望这篇文章的阐述可以帮助读者串起这一条支线。