Higher Algebra and Semi-Definite Programming
5/11/2020

Anthony Della Pella

These notes and results are largely based on selected chapters from the book [Bôcher & Duval, 1907], and most notation and terminology is that of the author of the book. Frequent reference will also be made to the lectures [Sturmfels, 2014]. The proofs are my own, and I refer to results and proofs from the book or lecture notes when appropriate.

1. Polynomials

Definition 1.1 (Homogeneous)**.** A polynomial all of whose terms are of the same degree is *homogeneous*. We often refer to such polynomials as forms although this convention varies among other authors.

Definition 1.2 (Homogeneous Coordinates)**.** When describing points in n-dimensional Euclidean space it is sometimes helpful to use an $n+1^{\text{th}}$ quantity. For example, in the Cartesian plane we have points specified by (x, y). Using a third quantity t we have:

$$X = \frac{x}{t} \qquad Y = \frac{y}{t}. \tag{1}$$

Here (X, Y) will be the Cartesian coordinates of a point.

A couple remarks on the above definition:

- We exclude the point $(0, 0, \ldots, 0)$ from consideration.
- The point $(x_1, \ldots, x_n, 0)$ is "the point at ∞". This follows logically from a limiting argument noting that as the last (t) coordinate becomes smaller and smaller the point described by the homogeneous coordinates moves farther and farther away from the origin.

As an example we can consider a degree 2 polynomial in the Cartesian plane.

Example 1.3. Let $AX^2 + BXY + CY^2 + DX + EY + F = 0$ be a polynomial in the 2 Cartesian coordinates X and Y. Writing this in homogeneous coordinates, it becomes:

$$A\frac{x^2}{t^2} + B\frac{xy}{t^2} + C\frac{y^2}{t^2} + D\frac{x}{t} + E\frac{y}{t} + F = 0. \tag{2}$$

Rewriting Eq. (2) we have:

$$Ax^2 + Bxy + Cy^2 + Dxt + Eyt + Ft^2 = 0. \tag{3}$$

This formulation of our original polynomial equation has two advantages:

1. We have obtained a *homogeneous* polynomial.
2. This provides a dictionary between homogeneous polynomials in 3 dimensions of this form, and 2 dimensional polynomials.

Item 2 is worth noting because in this special case we can view an inherently 3 dimensional object by drawing a 2 dimensional picture.

2. Matrices And Quadratic Forms

Fix $V \cong \mathbb{R}^n$. Quadratic forms on V are represented by symmetric $n \times n$ matrices. When X is a symmetric $n \times n$ matrix, and ϕ is the quadratic form corresponding to X, we have:

$$\phi : V \to \mathbb{R} \tag{4}$$
$$\vec{u} \mapsto \vec{u}^T X \vec{u}. \tag{5}$$

Where there is no possibility of ambiguity, we will often drop the vector notation instead representing vectors by lower case math script letter.

We now recall that for the quadratic form ϕ, the associated symmetric bi linear form $B(x, y) = \frac{1}{2}(\phi(x + y) - \phi(x) - \phi(x))$ is equivalent to $x^T Xy$. We also recall the spectral theorem in it's general form:

Fact 2.1. We can diagonal ice the matrix X over $\mathbb{R}$, that is there exists Λ such that

$$\Lambda^T X \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n). \tag{6}$$

The entries λ_i are the (real) eigenvalues of X, and Λ is an orthogonal matrix whose columns are the corresponding age vectors of X.

We now remark that, in addition to the "inner product" representation for ϕ, we also have a sum of squares representation for ϕ:

Theorem 2.2. *For any quadratic form* $q(x) = q(x_1, \ldots, x_n) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j$ *we have a representation of* $q(u)$ *as a sum of squares in the entries of* u *whenever there is at least one coefficient* a_{ii} *that is nonzero.*

Proof. We provide an algorithm to obtain the desired sum of squares representation. We begin by assuming without loss of generality, that $a_{11} \neq 0$. Our goal will be to successively remove the non squared terms from the sum. We start by getting rid of those terms involving x_1. Consider

$$p(x_2, x_3, \ldots, x_n) = q(x) - \frac{1}{a_{11}} \left(a_{11}x_1 + \cdots + a_{1n}x_n\right)^2. \tag{7}$$

Rearranging, $q(x) = \frac{1}{a_{11}} \left(a_{11}x_1 + \cdots + a_{1n}x_n\right)^2 + p(x_2, x_3, \ldots, x_n)$.

Changing coordinates so that $u_1 = a_{11}x_1 + \cdots + a_{1n}x_n$, with $u_2 = x_2, \ldots, u_3 = x_3, \ldots, u_n = x_n$ remaining the same, we can write:

$$\begin{aligned} q(u) &= \frac{1}{a_{11}} u_1^2 + p(x_2, x_3, \ldots, x_n) &(8)\\ &= \frac{1}{a_{11}} u_1^2 + p(u_2, u_3, \ldots, u_n). &(9)\end{aligned}$$

Iterating, we obtain a representation for $q = q(u)$ as a sum of squares in the terms u_i. $\square$

Theorem 2.3. *We have the following explicit representation for* $\phi(u)$*:*

$$\phi(u) = \sum_{j=1}^{n} \lambda_j \left(\sum_{i=1}^{n} \Lambda_{ij} u_i \right)^2. \tag{10}$$

Proof. We let $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ be the diagonal matrix of eigenvalues. Then, $X = \Lambda D \Lambda^T$. Thus, $\phi(u) = u^T \Lambda D \Lambda^T u$.

$$\left(\Lambda^T u\right)_i = \sum_{j=1}^{n} \Lambda_{ji} u_j. \tag{11}$$

Using this, we have:

$$\left(D\Lambda^T u\right)_i = \sum_{j=1}^{n} \lambda_i \Lambda_{ji} u_j. \tag{12}$$

Finally,

$$\left(u^T\Lambda\right)_i = \sum_{j=1}^{n} u_j \Lambda_{ji}. \tag{13}$$

Combining Eq. (13) and Eq. (12) we have

$$\phi(u) = \sum_{i=1}^{n} \lambda_i \left(\sum_{j=1}^{n} u_j \Lambda_{ji} \right)^2. \tag{14}$$

Using our algorithm from Theorem 2.2 we can obtain:

$$\phi(u) = \sum_{j=1}^{n} \lambda_j \ell_j u_j^2. \tag{15}$$

□

We now shift our focus to determinants.

3. Determinants

3.1. Complementary Determinants

Definition 3.1 (Rank)**.** A matrix is *rank* r if it contains at least one r-rowed determinant which is not zero, while all determinants of order higher than r are 0.

We will often speak about the determinant of rank r by which we mean the rank of the matrix of the determinant being r.

Definition 3.2 (Complement)**.** As usual, to each element of a matrix (or determinant) we can find a corresponding first minor obtained by striking out the row and column of the determinant in which the given element lies. Since the elements of a determinant of n^{the} order can be regarded as the $(n-1)^{\text{the}}$ minors of that determinant, we have a natural pairing between 1 rowed minors (elements of the matrix) and $(n-1)$-rowed minors. We can do this in a similar fashion for 2 rowed minors in which we will pair these with $(n-2)$-rowed minors. This procedure can be carried out for any r-rowed minor (and its corresponding $(n-r)$-rowed minor). In general, the two minors we pair up will be called *complementary*.

For example, we consider the case of a 5×5 determinant below.

Example 3.3.
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{vmatrix}$$
where the two minors:
$$\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} a_{12} & a_{14} & a_{15} \\ a_{42} & a_{44} & a_{45} \\ a_{52} & a_{54} & a_{55} \end{vmatrix}$$
are complementary.

We now define a slightly more nuanced version of complements that was stated above.

Definition 3.4 (Algebraic Complement, Principal Minor)**.** If M is the m-rowed minor of D in which the rows $k_1, \dots, k_m$ and the columns $l_1, \dots, l_m$ are represented, then the *algebraic complement* of M is defined by the equation

$$(-1)^{k_1+\cdots+k_m+l_1+\cdots+l_m} \cdot [\text{complement of } M]. \tag{16}$$

We also remark now the well known convention that the minor of a determinant obtained by striking out the same rows as columns is referred to as the *principal minor*.

Theorem 3.5. *The algebraic complement of any principal minor is equal to its plain complement.*

Proof. This follows from noting that $k_1 = l_1$ in the notation used in Definition 3.4, as well as $k_2 = l_2, \dots, k_m = l_m$. The sign term $(-1)^{k_1+\cdots+k_m+l_1+\cdots+l_m}$ is always positive and the result follows. □

Theorem 3.6. *If M and N are complementary minors either M and N are the algebraic complements of one another, or $-N$ is the algebraic complement of M and $-M$ is the algebraic complement of N.*

Proof. We begin by (utilizing the notation of Definition 3.4) considering M as the minor obtained when crossing out rows $k_1, \dots, k_m$, and columns $l_1, \dots, l_m$. Then, since it is complementary to M, N is obtained by crossing out the $(n-m)$ rows $k_{m+1},\dots,k_n$ and the $(n-m)$ columns $l_{m+1},\dots,l_n$.

There are n total rows and n total columns, thus

$$\sum_{i=1}^{n}(k_i+l_i)=n(n+1) \tag{17}$$

which is even.

Finally, we utilize the fact that if we partition an even number into a sum of two numbers those two numbers must have the same sign. Realizing that $k_1+\cdots+k_m+l_1+\cdots+l_m$ provides such a partition of $n(n+1)$ we conclude the result. □

It is also useful in this section to remark some determinant identities that will be useful later:

3.2. Determinant Identities

For the following identities we assume that all dimensions line up, and in particular det is always called on a square matrix.

1. (Algebraic Definition) For an $n\times n$ matrix A we can write the determinant as a polynomial in the entries (a_{ij}) of A:

$$\sum_{\sigma\in S_n}\prod_{i=1}^{n}\operatorname{sgn}(\sigma)a_{i,\sigma(i)} \tag{18}$$

   where S_n is the symmetric group on n elements.

2. The homomorphism identity: $\det(AB)=\det(A)\det(B)$.

3. (Schur Decomposition Identity) If A is invertible, then

$$\det\begin{pmatrix}A & B\\ C & D\end{pmatrix}=\det(A)\det\left(D-CA^{-1}B\right). \tag{19}$$

4. (Dodgson Condensation Identity) Let $n\geq 3$ and M an $n\times n$ matrix with M_j^i denoting the $(n-1)\times(n-1)$ submatrix with the i^{th} row and j^{th} column removed. If $M_{j,k}^{h,i}$ denotes the matrix with the i^{th} and h^{th} row and j^{th} and k^{th} column removed, then

$$\det(M)\det\left(M_{1,n}^{1,n}\right)=\det\left(M_1^1\right)\det(M_n^n)-\det\left(M_n^1\right)\det(M_1^n). \tag{20}$$

5. (Sylvester's Identity) In [Tao, 2017] A more general form of Item 4 is given as Sylvester's Identity. Here, $1 \leq k < n$ and S, S' are k-element subsets of $[n]$. $M^S_{S'}$ is the matrix formed from M by removing the rows associated to S and the columns associated to S'. We have:

$$\det(M)\det\left(M^S_S\right)^{k-1} = \det\left(\det\left(M^{S\setminus\{i\}}_{S\setminus\{j\}}\right)\right)_{i,j\in S} \tag{21}$$

6. Another identity for block matrices of determinants is given below. Here X, Y, Z, W are column vectors, and A is an $n \times (n-2)$ matrix.

$$\det\begin{pmatrix}\det(X,Y,A) & \det(X,W,A)\\ \det(Z,Y,A) & \det(Z,W,A)\end{pmatrix} = \det(X,Z,A)\det(Y,W,A). \tag{22}$$

4. The Spectral Theorem

We start this section (which will be useful later) by stating the spectral theorem for Hermitian matrices. We recall the definition of Hermitian below:

Definition 4.1 (Hermitian)**.** An $n \times n$ matrix $A = (a_{ij})^n_{i,j=1}$ is *Hermitian* or *Self-Adjoint* if it is equal to its own conjugate transpose. That is, if $a_{ij} = \overline{a_{ji}}$.

A few useful remarks about Hermitian matrices are:

- A is Hermitian if and only if it is equal to its adjoins, that is it satisfies

$$\begin{aligned} w(Av)^T &= \langle w, Av\rangle \\ &= \langle Aw, v\rangle \\ &= (Aw)b^T \end{aligned}$$

- A is Hermitian if and only if it has real quadratic form, that is:

$$\langle v, Av\rangle \in \mathbb{R} \tag{23}$$

  where Eq. (23) holds for all vectors v.

We can now state the spectral theorem for Hermitian matrices:

Theorem 4.2. *If A is Hermitian, there exists an orthonormal basis of V consisting of eigenvectors of A. Moreover, each eigenvalue is real.*

The spectral theorem holds for the more broader class of matrices known as Normal matrices.

Definition 4.3. An $n \times n$ matrix A is *Normal* if

$$A^* A = A A^* \tag{24}$$

where A^* represents the conjugate transpose of the matrix, *i.e.* $(a_{ij})^* = (\overline{a_{ji}})$.

References

Bôcher, Maxime, & Duval, E. P. R. 1907. *Introduction to higher algebra.* Dover.

Sturmfels, Bernd. 2014. *Introduction to non-linear algebra.* KMRS & Mathnet Korea.

Tao, Terence. 2017 (Sep). *large-dimensional gelfand-tsetlin patterns.*