MATH 470 Independent Study in Matrix Theory: The Kronecker Product

Thomas Zhang

Preview: Why Might We Care

There are certain kinds of matrix equations of particular interest; let's start with some linear ones

1. $AX = B$
2. $AX + XB = C$ (Lyapunov's equation when C is Hermitian; commute if $A = B, C = 0$)
3. $AXB = C$
4. $A_1XB_1 + A_2XB_2 + \cdots + A_kXB_k = C$ (Bapat and Sunder's equation of interest is a special case where $B_i = A_i^*$ and $C = I$)
5. $AX + YB = C$.

Some non-linear matrix equations are also of interest, in particular quadratic equations

1 Basic Properties

Definition 1.1 $\otimes : M_{m,n}(\mathbb{F}) \times M_{p,q}(\mathbb{F}) \to M_{mp,nq}(\mathbb{F})$,

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

Let's start with the straightforward properties of the Kronecker product

1. $(cA) \otimes B = A \otimes (cB)$
2. $(A \otimes B)^\top = A^\top \otimes B^\top$
3. $(A \otimes B)^* = A^* \otimes B^*$
4. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$ (associativity)
5. $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$ (distributivity pt 1)

6. $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$ (distributivity pt 2)

7. If A, B are Hermitian, then $A \otimes B$ is also Hermitian.

Next, we look at a few important properties that require a bit of proof.

Lemma 1.2 *(Mixed-product property) Let $A \in M_{m,n}, B \in M_{p,q}, C \in M_{n,k}, D \in M_{q,r}$ such that $A \otimes B \in M_{mp,nq}$ and $C \otimes D \in M_{nq,kr}$ where matrix multiplication makes sense. Then*

1. $(A \otimes B)(C \otimes D) = AC \otimes BD$

2. $\implies (A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ *(if A, B are nonsingular)*

Proof: 1: This follows if we're careful about indexing the blocks of our matrices.
2: Apply 1 and observe $(A \otimes B)(A^{-1} \otimes B^{-1}) = AA^{-1} \otimes BB^{-1} = I_m \otimes I_p$ but at the same time by the uniqueness of inverse this implies $(A \otimes B)^{-1} = (A^{-1} \otimes B^{-1})$.

Another magical thing about Kronecker products is how they preserve a lot of eigenvalue/singular value properties in some form. First we introduce the following notation:

Definition 1.3 *We can "spaghettify" a matrix $A \in M_{m,n}$ by turning it into a vector $vec(A) \in \mathbb{F}^{mn}$ by defining the following operation:*

$$vec(A) := [a_{11}, \ldots, a_{m1}, a_{12}, \ldots, a_{m2}, \ldots, a_{1n}, \ldots, a_{nm}]^\top,$$

basically putting the columns of the matrix end to end. A pretty cool corollary is that you can verify the Frobenius inner product $tr(A^*B)$ is in fact an inner product by verifying it in spaghettified form.

Now for some cool eigenvalue FACTS

Theorem 1.4 *Let $A \in M_n$, $B \in M_m$, with $\lambda \in \sigma(A), x \in \mathbb{C}^n$ and $\mu \in \sigma(B), y \in \mathbb{C}^m$ as accompanying eigenpairs, then*

$$\lambda\mu \in \sigma(A \otimes B), \textit{with corresponding eigenvector } x \otimes y \in \mathbb{C}^{mn}.$$

This implies that every eigenvalue of $A \otimes B$ is a product of eigenvalues of A and B, counting multiplicities. Therefore, $\sigma(A \otimes B) = \sigma(B \otimes A)$.

Proof: We apply the mixed-product property on $(A \otimes B)(x \otimes y)$:

$$(A \otimes B)(x \otimes y) = (Ax) \otimes (By) = (\lambda x) \otimes (\mu x) = \lambda\mu(x \otimes y).$$

We have just shown that $x \otimes y$ is indeed an eigenvector of $A \otimes B$ associated with eigenvalue $\lambda\mu$. To see that the spectrum $\sigma(A \otimes B)$ is precisely the product of all the eigenvalues of A and B accounting for multiplicity, we appeal to Schur's Decomposition Theorem, where we can find a unitary transformation of a matrix into an upper triangular matrix where the diagonal entries are precisely the eigenvalues. Say we have $A = SU_AS^*$, $B = TU_BT^*$. By Lemma 1.2.2, we see that the matrix $S \otimes T$ is also unitary. We then observe that

$$(S \otimes T)^*(A \otimes B)(S \otimes T) = (S^* \otimes T^*)(AS \otimes BT) = (S^*AS \otimes T^*BT) = U_A \otimes U_B.$$

We see that any product of eigenvalues from and A and B necessarily must be on the diagonal somewhere of $U_A \otimes U_B$, and $U_A \otimes U_B$ is precisely the Schur-triangularized matrix. Therefore, any eigenvalue of $A \otimes B$ is some product of eigenvalues of A and B.

Corollary 1.5 *An immediate result from the previous theorem is that the Kronecker product of two positive (negative) semi-definite matrices is positive semi-definite, since the pointwise products of eigenvalues continue to be positive.*

Theorem 1.6 *(Singular Value Decomposition) Let $A \in M_{m,n}$, $B \in M_{p,q}$ have singular value decompositions: $A = V_A \Sigma_A W_A^*$, $B = V_B \Sigma_B W_B^*$ and rank$(A) = r_1$ and rank$(B) = r_2$. The non-zero singular values of $A \otimes B$ are the $r_1 r_2$ positive numbers $\{\sigma_i(A)\sigma_j(B)\}$. Furthermore, the singular values of $A \otimes B$ are the same as $B \otimes A$, which implies rank$(A \otimes B)$ = rank$(B \otimes A) = r_1 r_2$.*

So far, we have observed a general phenomenon where things are the same between $A \otimes B$ and $B \otimes A$. The problems and their solutions will establish a few more. All of these together are ultimately explained by the fact that $B \otimes A$ is permutation equivalent to $A \otimes B$, i.e. there are permutation matrices P, Q such that $B \otimes A = P(A \otimes B)Q$.

1.1 Exercises

1. Show that if $A \in M_n$ and $B \in M_m$ are square, then $\det(A \otimes B) = \det(A)^m \det(B)^n$, and $\det(A \otimes B) = \det(B \otimes A)$.

2. Show that if $A \in M_n$ and $B \in M_m$ are square, then $\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\mathrm{tr}(B) = \mathrm{tr}(B \otimes A)$.

3. Show that $A \otimes B$ is normal if and only if $B \otimes A$ is normal.

2 Linear Matrix Equations and the Kronecker Product

Equipped with the basic properties of the Kronecker Product, we can go back and re-write the matrix equations in the Preview section. We first write them out then show a proof for one (I might come back and prove the rest when I'm done proving the cooler stuff):

1. $AX = B \implies (I \otimes A)\mathrm{vec}(X) = \mathrm{vec}(B)$

2. $AX + XB = C \implies [(I \otimes A) + (B^\top \otimes I)]\mathrm{vec}(X) = \mathrm{vec}(C)$

3. $AXB = C \implies (B^\top \otimes A)\mathrm{vec}(X) = \mathrm{vec}(C)$

4. $A_1XB_1 + A_2XB_2 + \cdots + A_kXB_k = C \implies [B_1^\top \otimes A_1 + \cdots + B_k^\top \otimes A_k]\mathrm{vec}(X) = \mathrm{vec}(C)$

5. $AX + YB = C \implies (I \otimes A)\mathrm{vec}(X) + (B^\top \otimes I)\mathrm{vec}(Y) = \mathrm{vec}(C)$

Proof of 3: Let $A \in M_{m,n}$, $B \in M_{p,q}$, and let A_k denote the k-th column of A. Observe that by matrix multiplication $(AXB)_k = A(XB)_k = AX(B_k)$. Using some more matrix multiplication facts, we get

$$\begin{aligned} AX(B_k) &= A\Big[\sum_{i=1}^{p} b_{ik} X_i\Big] \\ &= [b_{1k}A \quad b_{2k}A \ \ldots \ b_{pk}A]\mathrm{vec}(X) \\ &= (B_k^\top \otimes A)\mathrm{vec}(X). \end{aligned}$$

Therefore, stacking the q columns of B together we get

$$\mathrm{vec}(C) = \mathrm{vec}(AXB) = \begin{bmatrix} B_1^\top \otimes A \\ \vdots \\ B_q^\top \otimes A \end{bmatrix} \mathrm{vec}(X) = (B^\top \otimes A)\mathrm{vec}(X) \quad \blacksquare$$

A convenient, related fact is the following:

Lemma 2.1 *For $A, B \in M_n$, vec$(AB) = (I_n \otimes A)$vec(B).*

By spaghettifying things through the Kronecker product, it is often heuristically important to pre-process the matrix equation. An example of pre-processing Horn and Johnson give is: let's say we're dealing with equation 2 and we want to deal with an equation where different similar transformations have been applied on A and B. Then we can translate the equation in the following way:

$$\begin{aligned} AX + XB &= SAX + SXB \\ &= SAXT + SXBT \\ &= (SAS^{-1})SXT + SXT(T^{-1}BT) = SCT. \end{aligned}$$

If we define $A' = SAS^{-1}$, $B' = SBS^{-1}$, $X' = SXT$, $C' = SCT$, we now have a hopefully nicer system to work with.

A natural thought to have at this point is: well what happens to linear transformations on the original matrices after you've taken Kronecker products? Borrowing some basic abstract algebra, we observe that the mapping vec : $M_{m,n} \to \mathbb{C}^{mn}$ is clearly an isomorphism, since any matrix and be strung out, and any long vector can be sliced up and arranged to form a matrix. The next fact follows:

Observation 2.2 *$T : M_{m,n} \to M_{p,q}$ is a linear transformation. Then there exists a unique matrix $K(T) \in M_{mp,nq}$ such that if $T(X) = Y$, then vec$(T(X)) = K(T)$vec$(X) =$ vec(Y).*

Apparently, the next step is to classify all "linear derivations", where derivations are defined as all linear transformations where additionally $T(XY) = T(X)Y + XT(Y)$ for all $X, Y \in M_n$. Kronecker products give us a rather surprising and convenient characterization of linear derivations:

Theorem 2.3 *$T : M_n \to M_n$ is a linear derivation if and only if there is some $C \in M_n$ such that $T(X) = CX - XC$ for all $X \in M_n$.*

Proof: It is not difficult to verify that $T(X) = CX - XC$ is a linear derivation. For the forward direction, we use the previous observations and lemmas to expand the following:

$$\begin{aligned} \mathrm{vec}(T(XY)) &= \mathrm{vec}(T(X)Y) + \mathrm{vec}(XT(Y)) \\ K(T)\mathrm{vec}(XY) &= (I \otimes T(X))\mathrm{vec}(Y) + (I \otimes X)\mathrm{vec}(T(Y)) \\ K(T)(I \otimes X)\mathrm{vec}(Y) &= (I \otimes T(X))\mathrm{vec}(Y) + (I \otimes X)K(T)\mathrm{vec}(Y) \\ \iff (I \otimes T(X))\mathrm{vec}(Y) &= K(T)(I \otimes X)\mathrm{vec}(Y) - (I \otimes X)K(T)\mathrm{vec}(Y) \\ I \otimes T(X) &= K(T)(I \otimes X) - (I \otimes X)K(T). \end{aligned}$$

Observe that $I \otimes X$ is 0 apart from $[I \otimes X]_{ii}$, referring to the ii-blocks. Therefore, $[K(T)(I \otimes X)]_{ii} = K_{ii}X$.

$$\implies [I \otimes T(X)]_{ii} = T(X) = K_{ii}X - XK_{ii} \text{ for each } i = 1, \dots, n.$$

We can pick any K_{ii} to be the matrix C to finish the proof. ■

We mentioned in the first section that $A \otimes B$ is permutation-equivalent to $B \otimes A$. To prove this, we do the brunt of the work in proving that $\text{vec}(X)$ is permutation equivalent to $\text{vec}(X)^\top$.

Theorem 2.4 *For arbitrary $X \in M_{m,n}$, there is a unique matrix $P(m,n) \in M_{mn}$ such that*

$$P(m,n) = \sum_{i=1}^{m} \sum_{j=1}^{n} E_{ij} \otimes E_{ij}^\top,$$

where each $E_{ij} \in M_{m,n}$ has entry 1 in position i, j and 0 everywhere else. It turns out that $P(m,n)$ is a permutation matrix (such that $P(m,n) = P(n,m)^\top = P(n,m)^{-1}$).

Proof: Observe that we can write $x_{ij}E_{ij}^\top = E_{ij}^\top X E_{ij}^\top$ Therefore,

$$X^\top = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}E_{ij}^\top = \sum_{i=1}^{m} \sum_{j=1}^{n} E_{ij}^\top X E_{ij}^\top.$$

Now we can write out $\text{vec}(X^\top)$:

$$\begin{aligned} \text{vec}(X^\top) &= \sum_{i=1}^{m} \sum_{j=1}^{n} \text{vec}(E_{ij}^\top X E_{ij}^\top) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} (E_{ij} \otimes E_{ij}^\top)\text{vec}(X). \end{aligned}$$

Now we have to verify that $P(m,n) = \sum_{i=1}^{m} \sum_{j=1}^{n} (E_{ij} \otimes E_{ij}^\top)$ is indeed a permutation matrix. Let E'_{ij} be the unit matrices of the transposed matrix space M_{nm} such that $E'_{ij} = E_{ji}^\top$. Observe that

$$\begin{aligned} P(n,m) &= \sum_{i=1}^{m} \sum_{j=1}^{n} (E'_{ij} \otimes E_{ij}'^\top) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} (E_{ji}^\top \otimes E_{ji}) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} (E_{ij} \otimes E_{ij})^\top = P(m,n)^\top. \end{aligned}$$

To see that $P(m,n) = P(n,m)^{-1}$, observe that $X = (X^\top)^\top$, so $\text{vec}(X) = P(n,m)\text{vec}(X^\top) = P(n,m)P(m,n)\text{vec}(X) \implies P(m,n) = P(n,m)^{-1}$. This completes the proof.

Corollary 2.5 *For any* $A \in M_{m,n}$, $B \in M_{p,q}$, *then*

$$B \otimes A = P(m,p)^\top (A \otimes B) P(n,q).$$

In fact,

$$\sum_{i=1}^{k} B_i \otimes A_i = P(m,p)^\top \Big(\sum_{i=1}^{k} A_i \otimes B_i \Big) P(n,q)$$

Corollary 2.6 *Let* $A_1, \ldots, A-r$ *and* $B_1, \ldots, B_s$ *be given square complex matrices. Then*

$$\Big(\bigoplus_{i=1}^{r} A_i \Big) \otimes \Big(\bigoplus_{j=1}^{s} B_j \Big) = P(n,p)^\top \Big(\bigoplus_{i,j=1} A_i \otimes B_j \Big) P(n,p).$$

2.1 Exercises

1. Describe how to determine $K(T)$ from T.

3 Kronecker Sums and the Equation $AX + XB = C$

Earlier, we brought up the equation $AX + XB = C$, which we see is a generalization of the form in which we see important equations like Lyapunov's equation. In the previous section we were able to re-write equations of that form:

$$AX + XB = C \implies [(I_m \otimes A) + (B^] top \otimes I_n)]\mathrm{vec}(X) = \mathrm{vec}(C).$$

The "Kronecker form" of the above equation are the family of equations written in the form $(I_m \otimes A) + (B \otimes I_n)$. The Kronecker form actually tells us the following surprising fact

Theorem 3.1 *Let* $A \in M_n$ *and* $B \in M_m$ *be given. If* (λ, x) *is an eigenpair of* A *and* (μ, y) *is an eigenpair of* B, *then* $(\lambda+\mu, y \otimes x)$ *is an eigenpair of the Kronecker form* $(I_m \otimes A) + (B \otimes I_n)$. *Similar to Theorem 1.4, it turns out every eigenvalue of* $(I_m \otimes A) + (B \otimes I_n)$ *is a sum of an eigenvalue from* A *and* B: $\sigma[(I_m \otimes A) + (B \otimes I_n)] = \{\lambda_i + \mu_j\}$. *Furthermore,* $(I \otimes A)$ *commutes with* $(B^\top \otimes I)$, *and* $\sigma[(I_m \otimes A) + (B \otimes I_n)] = \sigma[(I_n \otimes B) + (A \otimes I_m)]$

Proof: to show commutativity, we use the mixed-product property:

$$(I_m \otimes A)(B \otimes I_n) = B \otimes A = (B \otimes I_n)(I_m \otimes A).$$

As for the claim about the eigenvalues, we do a similar trick we pulled in Theorem 1.4 and appeal to Schur's Decomposition Theorem such that $S^* A S = U_A$, $T^* B T = U_B$ are upper triangular and have the eigenvalues of A and B as their diagonals.

$$(S \otimes T)^*(I_m \otimes A)(S \otimes T) = I_m \otimes U_A = \begin{bmatrix} U_A & & & 0 \\ & U_A & & \\ & & \ddots & \\ 0 & & & U_A \end{bmatrix}$$
$$(S \otimes T)^*(B \otimes I_n)(S \otimes T) = U_B \otimes I_n = \begin{bmatrix} \mu_1 I_n & & & 0 \\ & \mu_2 I_n & & \\ & & \ddots & \\ 0 & & & \mu_m I_n \end{bmatrix}$$

both of which are upper triangular matrices that contain the eigenvalues of $I_m \otimes A$ and $B \otimes I_n$ on the diagonal. Therefore, $(S \otimes T)^*[(I_m \otimes A) + (B \otimes I_n)](S \otimes T)$ will be upper triangular and the eigenvalues of the Kronecker form are clearly sums of eigenvalues of A and B accounting for multiplicity. ■

We can use this theorem to prove some solvability facts about the matrix equation $AX + XB = C$ and its relatives.

Theorem 3.2 *Going back to the matrix equation $AX + XB = C$, the equation has a unique solution $X \in M_{n,m}$ for each $C \in M_{n,m}$ if and only if $\sigma(A) \cap \sigma(-B) = \emptyset$.*

Proof: We observe that if the Kronecker form $(I_m \otimes A) + (B^\top \otimes I_n)$ is invertible, then we have a solution for the equation given any C, otherwise, since there will be a non-trivial null space, the solutions to the equation will not be unique (special solution + particular solution). Since the eigenvalues of B and $B^\top$ are the same, we can consider the Kronecker form $(I_m \otimes A) + (B \otimes I_n)$ from the previous theorem. Observe that the triangularized will have a 0 (eigenvalue 0 a.k.a. nontrivial nullspace) if some eigenvalue of A cancels out with an eigenvalue of B. Therefore, the Kronecker form is invertible, hence the equation is uniquely solvable for any C if $\sigma(A) \cap \sigma(-B) = \emptyset$, i.e. no eigenvalue of A cancels out with an eigenvalue of B.

Corollary 3.3 *Let $A \in M_n$ and $B \in M_m$. $AX - XB = 0$ has a nonzero solution $X \in M_{n,m}$ if and only if $\sigma(A) \cap \sigma(-B) \neq \emptyset$ (i.e. Kronecker form has nontrivial nullspace).*

Corollary 3.4 *Let $A \in M_n$. The equation $XA + A^*X = C$ has a unique solution $X \in M_n$ for each $C \in M_n$ if and only if $\sigma(A) \cap \sigma(-A) = \emptyset$. Lookie here, smells like Lyapunov. We can show that $(I \otimes A^*) + (A^\top \otimes I)$ is the Kronecker form of the equation, and if we shift around the equation $AX + XA^* = C$, we get the Kronecker form $(I \otimes A) + (\overline{A} \otimes I)$*

Corollary 3.5 *Suppose that $XA + A^*X$ has a unique solution for every C. The solution X is Hermitian if and only if C is Hermitian. Furthermore, if $A \in M_n$ is positive stable (all eigenvalues have positive real part), then the equation has a unique solution X for each $C \in M_n$.*

Horn and Johnson jump off the deep end in terms of symbols and technical proofs involving Jordan forms after this point, so I'll stick to stating the cool results:

Definition 3.6 *Any polynomial that a matrix A satisfies must divide the characteristic polynomial $ch(A)$. The* ***minimal polynomial*** *of matrix A is the lowest degree monic polynomial that A satisfies. A matrix is* ***non-derogatory*** *if the smallest degree polynomial that A satisfies (minimal polynomial) is precisely the characteristic polynomial of A (up to a factor of ± 1).*

Theorem 3.7 *Given $A \in M_n$, the set of matrices that commute with A is a subspace of M_n of at least n. The dimension is equal to n if and only if A is non-derogatory.*

Definition 3.8 *Let $A \in M_n$ be a given matrix. The* ***centralizer*** *of A is the set $C(A) := \{B \in M_n : AB = BA\}$, i.e. the set of all B that commute with A. The set of all polynomials of A can be denoted $P(A) := \{p(A) : p(t) \text{ is a polynomial}\}$. It is immediately clear that $P(A) \subseteq C(A)$. However, their relationship does not stop there.*

Theorem 3.9 *Let $A \in M_n$ be a given matrix and let $q_A(t)$ be the minimal polynomial of A. Then:*

1. *$P(A)$ and $C(A)$ are subspaces of M_n.*
2. *degree of $q_A(t) =$ dim $P(A) \leq n$.*
3. *dim $C(A) \geq n$ with equality if and only if A is nonderogatory.*

Corollary 3.10 *A matrix $A \in M_n$ is non-derogatory if and only if every matrix that commutes with A is a polynomial in A.*

Corollary 3.11 *Given $A \in M_n$, $B \in M_n$ is a polynomial in A if and only if B commutes with every matrix that commutes with A.*

Theorem 3.12 *Let $A \in M_n$, $B \in M_n$ and $C \in M_{m,n}$ be given. Then there is some $X \in M_{m,n}$ such that $AX - XB = C$ if and only if*

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \text{ is similar to } \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

Theorem 3.13 *Let $A \in M_{m,n}$, $B \in M_{p,q}$ and $C \in M_{m,q}$ be given. There are matrices $X \in M_{m,q}$ and $Y \in M_{n,p}$ such that $AX - YB = C$ if and only if*

$$\text{rank} \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} = \text{rank} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$