The Hadamard Product

Elizabeth Million

April 12, 2007

1 Introduction and Basic Results

As inexperienced mathematicians we may have once thought that the natural definition for matrix multiplication would be entrywise multiplication, much in the same way that a young child might say, "I writed my name." The mistake is understandable, but it still makes us cringe. Unlike poor grammar, however, entrywise matrix multiplication has reason to be studied; it has nice properties in matrix analysis and has applications in both statistics (301 [4], 140 [5]) and physics (93, 149 [5]). Here we will only expore the properties of the Hadamard product in matrix analysis.

Definition 1.1. *Let A and B be $m \times n$ matrices with entries in* $\mathbf{C}$. *The* Hadamard product *of A and B is defined by $[A \circ B]_{ij} = [A]_{ij}[B]_{ij}$ for all $1 \leq i \leq m$, $1 \leq j \leq n$.*

As we can see, the Hadamard product is simply entrywise multiplication. Because of this, the Hadamard product inherits the same benefits (and restrictions) of multiplication in $\mathbf{C}$. Note also that both A and B need to be the same size, but not necessarily square. To avoid confusion, juxtaposition of matrices will imply the "usual" matrix multiplication, and we will always use "$\circ$" for the Hadamard product.

Now we can explore some basics properties of the Hadamard Product.

Theorem 1.2. *Let A and B be $m \times n$ matrices with entries in* $\mathbf{C}$. *Then $A \circ B = B \circ A$.*

Proof. The proof follows directly from the fact that multiplication in $\mathbf{C}$ is commutative. Let A and B be $m \times n$ matrices with entries in $\mathbf{C}$. Then $[A \circ B]_{ij} = [A]_{ij}[B]_{ij} = [B]_{ij}[A]_{ij} = [B \circ A]_{ij}$ and therefore $A \circ B = B \circ A$. $\square$

Theorem 1.3. *The identity matrix under the Hadamard product is the $m \times n$ matrix with all entries equal to 1, denoted J_{mn}. That is, $[J_{mn}]_{ij} = 1$ for all $1 \leq i \leq m$, $1 \leq j \leq n$.*

Note: we have denoted the Hadamard identity as J_{mn} as to avoid confusion with the "usual" identity matrix, I_n.

Proof. Take any $m \times n$ matrix A with entries in $\mathbf{C}$. Then $[J_{mn} \circ A]_{ij} = (1)([A]_{ij}) = [A]_{ij}$ and so $J_{mn} \circ A = A$. Since the Hadamard Product is commutative (1.2), we know $J_{mn} \circ A = A \circ J_{mn} = A$. Therefore, J_{mn} as defined above is indeed the identity matrix under the Hadamard product. $\square$

Theorem 1.4. *Let A be an $m \times n$ matrix. Then A has a Hadamard inverse, denoted $\hat{A}$, if and only if $[A]_{ij} \neq 0$ for all $1 \leq i \leq m$, $1 \leq j \leq n$. Furthermore, $[\hat{A}]_{ij} = ([A]_{ij})^{-1}$.*

Proof. ($\Rightarrow$) Let A be an $m \times n$ matrix with Hadamard inverse $\hat{A}$. Then we know $A \circ \hat{A} = J_{mn}$. That is, $[A \circ \hat{A}]_{ij} = [A]_{ij}[\hat{A}]_{ij} = 1$. Multiplying by inverses in $\mathbf{C}$ we know that $[\hat{A}]_{ij} = (1)(([A]_{ij})^{-1}) = ([A]_{ij})^{-1}$, which is only possible when all entries of A are invertible (in $\mathbf{C}$). In other words, $[A]_{ij} \neq 0$ for all $1 \leq i \leq m$, $1 \leq j \leq n$.

($\Leftarrow$) Take any $m \times n$ matrix A with entries in $\mathbf{C}$ such that $[A]_{ij} \neq 0$ for all $1 \leq i \leq m$, $1 \leq j \leq n$. Then there exists $([A]_{ij})^{-1}$ for all i,j. This implies $[A]_{ij}([A]_{ij})^{-1} = ([A]_{ij})^{-1}[A]_{ij} = 1$, and so A has an inverse $\hat{A}$ defined by $[\hat{A}]_{ij} = ([A]_{ij})^{-1}$ for all i,j. □

Note: again we have denoted the Hadamard inverse as $\hat{A}$ as to avoid confusion with the "usual" matrix inverse, A^{-1}.

The Hadamard identity matrix and the Hadamard inverse are both more limiting than helpful, so we will not explore their use further. One last fun fact: the set of $m \times n$ matrices with nonzero entries form an abelian (commutative) group under the Hadamard product (Prove this!).

Theorem 1.5. ***The Hadamard Product is Linear.*** *Suppose $\alpha \in \mathbf{C}$, and A, B and C are $m \times n$ matrices. Then $C \circ (A + B) = C \circ A + C \circ B$. Furthermore, $\alpha(A \circ B) = (\alpha A) \circ B = A \circ (\alpha B)$.*

Proof. Part 1.

$$\begin{aligned} [C \circ (A + B)]_{ij} &= [C]_{ij}[A + B]_{ij} \\ &= [C]_{ij}([A]_{ij} + [B]_{ij}) \\ &= [C]_{ij}[A]_{ij} + [C]_{ij}[B]_{ij} \\ &= [C \circ A]_{ij} + [C \circ B]_{ij} \\ &= [C \circ A + C \circ B]_{ij} \end{aligned}$$

Part 2.

$$\begin{aligned} [\alpha(A \circ B)]_{ij} &= \alpha[A \circ B]_{ij} \\ &= \alpha[A]_{ij}[B]_{ij} \\ &= [\alpha A]_{ij}[B]_{ij} \\ &= [\alpha A \circ B]_{ij} && \text{First Equality} \\ &= \alpha[A]_{ij}[B]_{ij} \\ &= [A]_{ij}\alpha[B]_{ij} \\ &= [A]_{ij}[\alpha B]_{ij} \\ &= [A \circ \alpha B]_{ij} && \text{Second Equality} \end{aligned}$$

□

2 Diagonal Matrices and the Hadamard Product

We can relate the Hadamard product with matrix multiplication via considering diagonal matrices, since $A \circ B = AB$ if and only if both A and B are diagonal. For example, a simple calculation reveals that the Hadamard product relates the diagonal values of a diagonalizable matrix A with its eigenvalues:

Theorem 2.1. *Let A be a diagonalizable matrix of size n with eigenvalues λ_1, λ_2, $\ldots$, λ_n and diagonalization $A = SDS^{-1}$ where D is a diagonal matrix such that $[D]_{ii} = \lambda_i$ for all $1 \leq i \leq n$. Also, let $[A]_{ij} = a_{ij}$ for all i,j. Then*

$$\begin{bmatrix} a_{11} \\ a_{22} \\ \vdots \\ a_{nn} \end{bmatrix} = \left[S \circ (S^{-1})^T\right] \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \quad \textit{(304, [4])}.$$

Proof. Let $\mathbf{d} \in \mathbf{C}^n$ such that $[\mathbf{d}]_i = \lambda_i$ for all $1 \le i \le n$. We will show $a_{ii} = [S \circ (S^{-1})^T \mathbf{d}]_i$.

$$\begin{aligned}
a_{ii} &= [A]_{ii} \\
&= [SDS^{-1}]_{ii} \\
&= \sum_{j=1}^{n} \sum_{k=1}^{n} [S]_{ik} [D]_{kj} [S^{-1}]_{ji} \\
&= \sum_{k=1}^{n} [S]_{ik} [D]_{kk} [S^{-1}]_{ki} \\
&= \sum_{k=1}^{n} [S]_{ik} \lambda_k [S^{-1}]_{ki} \\
&= \sum_{k=1}^{n} [S]_{ik} [S^{-1}]_{ki} \lambda_k \\
&= \sum_{k=1}^{n} [S]_{ik} [(S^{-1})^T]_{ik} \lambda_k \\
&= \sum_{k=1}^{n} [S \circ (S^{-1})^T]_{ik} \lambda_k \\
&= \sum_{k=1}^{n} [S \circ (S^{-1})^T]_{ik} [\mathbf{d}]_k \\
&= [S \circ (S^{-1})^T \mathbf{d}]_i
\end{aligned}$$

□

We obtain a similar result when we look at the singular value decomposition of a square matrices:

Theorem 2.2. *Let A be a matrix of size n with singular values σ_1, σ_2, …, σ_n and singular value decomposition $A = UDV^*$, where D is a diagonal matrix with $[D]_{ii} = \sigma_i$ for all $1 \le i \le n$. Then,*

$$\begin{bmatrix} a_{11} \\ a_{22} \\ \vdots \\ a_{nn} \end{bmatrix} = \left[U \circ \bar{V} \right] \begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_n \end{bmatrix}$$

Note that this relation only makes sense with square matrices, since otherwise we cannot compute $U \circ \bar{V}$.

Proof. Note that $(V^*)^T = \bar{V}$. The proof is similar to our proof of (2.1). □

We have started to see that the Hadamard product behaves nicely with respect to diagonal matrices and normal matrix multiplication. The following are some more general properties that expand on this idea. The proofs all involve straightforward sum manipulation and are left as exercises for the reader.

Theorem 2.3. *Suppose A, B are $m \times n$ matrices, and D and E are diagonal matrices of size m and n, respectively. Then,*

$$\begin{aligned}
D(A \circ B)E &= (DAE) \circ B = (DA) \circ (BE) \\
&= (AE) \circ (DB) = A \circ (DBE) \quad \textit{(304, [4])}
\end{aligned}$$

We will need a quick definition before our next theorem.

Definition 2.4. *Define the diagonal matrix, $D_{\mathbf{x}}$, of size n with entries from a vector $\mathbf{x} \in \mathbf{C}^n$ by*

$$[D_{\mathbf{x}}]_{ij} = \begin{cases} [\mathbf{x}]_i & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 2.5. *Suppose A, B are $m \times n$ matrices and $\mathbf{x} \in \mathbf{C}^n$. Then the ith diagonal entry of the matrix $AD_{\mathbf{x}}B^T$ coincides with the ith entry of the vector $(A \circ B)\mathbf{x}$ for all $i = 1, 2, \ldots, m$ (305, [4]). That is,*

$$[AD_{\mathbf{x}}B^T]_{ii} = [(A \circ B)\mathbf{x}]_i \quad \text{for all} \quad 1 \le i \le m.$$

Theorem 2.6. *Let A, B and C be $m \times n$ matrices. Then the ith diagonal entry of the matrix $(A \circ B)C^T$ coincides with the ith diagonal entry of the matrix $(A \circ C)B^T$. That is, $[(A \circ B)C^T]_{ii} = [(A \circ C)B^T]_{ii}$ for all $1 \le i \le m$ (305, [4]).*

3 Schur Product Theorem

The purpose of this section is to set up and prove the Schur Product Theorem. This theorem relates positive definite matrices to the Hadamard product, which is important when analyzing the determinants of matrices since we want real, nonnegative numbers to compare. Note that if A is positive semidefinite of size n then clearly $|A| = \prod_{i=1}^{n} \lambda_i \ge 0$ since $\lambda_i \ge 0$ for all $1 \le i \le n$.

Lemma 3.1. *Any rank one positive semidefinite matrix A of size n can be written as the product $A = \mathbf{x}\mathbf{x}^T$ where $\mathbf{x}$ is some vector in $\mathbf{C}^n$.*

Proof. This proof follows the constructive proof of Theorem ROD [3].

Let A be a rank one positive semidefinite matrix of size n. Since A is positive semidefinite, we know it is Hermitian and therefore normal. This allows us to write the orthonormal diagonalization $A = UDU^*$ where D is the matrix of eigenvalues of A and U is a unitary matrix made up of orthonormalized eigenvectors of A (Note that U has real entries, so $U^* = U^T$). Also since A is positive semidefinite, we know all the diagonal values of D are nonnegative numbers.

In the proof of Theorem ROD [3] we now let $X^* = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be the columns of U and let $Y^* = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be the rows of U^* (which are the rows of U^T) converted to column vectors (which are the columns of U). Next we define $A_k = \lambda_k \mathbf{u}_k \mathbf{u}_k^T$ and we see that $A = A_1 + \ldots + A_n$. But A is rank one, so (ordered properly) $A = A_1 + \mathcal{O} + \ldots + \mathcal{O} = A_1 = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T$. We also know $\lambda_1 > 0$, so we can define $\mathbf{x} = \sqrt{\lambda_1}\mathbf{u}_1$. So now we have $\mathbf{x}\mathbf{x}^T = (\sqrt{\lambda_1}\mathbf{u}_1)(\sqrt{\lambda_1}\mathbf{u}_1)^T = \sqrt{\lambda_1}\mathbf{u}_1\mathbf{u}_1^T\sqrt{\lambda_1} = \sqrt{\lambda_1}\sqrt{\lambda_1}\mathbf{u}_1\mathbf{u}_1^T = \lambda_1\mathbf{u}_1\mathbf{u}_1^T = A$.

Therefore any positive semidefinite matrix A of rank one and size n can be written $\mathbf{x}\mathbf{x}^T$, where $\mathbf{x} \in \mathbf{C}^n$. ☐

Lemma 3.2. *Let B be a positive semidefinite matrix of size n and rank r. Then the matrices B_1, $\ldots$, B_r from the rank one decomposition of B are all also positive semidefinite.*

Proof. Let B be a positive semidefinite matrix of size n and rank r with rank one decomposition $B = B_1 + \ldots + B_r$. Then from the constructive proof of Theorem ROD [3] together with the orthonormal diagonalization of B we know that $B_k = UD_kU^*$ where U is a unitary matrix and D_k is the diagonal matrix with $[D_k]_{kk} = \lambda_k$ and $[D_k]_{ij} = 0$ for all other entries. Since B is positive semidefinite, then we know that all eigenvalues are positive (as ordered for the decomposition). That is, $\lambda_k > 0$ for all $1 \leq k \leq r$. Clearly then D_k is positive semidefinite for all $1 \leq k \leq r$.

Now consider $< B_k\mathbf{x}, \mathbf{x} >$ for any $\mathbf{x} \in \mathbf{C}^n$. Then

$$< B_k\mathbf{x}, \mathbf{x} >=< UD_kU^*\mathbf{x}, \mathbf{x} >=< D_kU^*\mathbf{x}, U^*\mathbf{x} >\geq 0.$$

This is true for all $1 \leq k \leq r$ and so B_k is positive semidefinite for all k. □

Lemma 3.3. *Let B be a matrix of size n and rank r. Furthermore, suppose the matrices B_1, ..., B_r from the rank one decomposition of B are all positive semidefinite. Then B itself is positive semidefinite.*

Proof. Let $B = B_1 + \ldots + B_r$ as described above and let $\mathbf{x} \in \mathbf{C}^n$. Then

$$\begin{aligned} < B\mathbf{x}, \mathbf{x} > &= < (B_1 + \ldots + B_r)\mathbf{x}, \mathbf{x} > \\ &= < B_1\mathbf{x} + \ldots + B_r\mathbf{x}, \mathbf{x} > \\ &= < B_1\mathbf{x}, \mathbf{x} > + \ldots + < B_r\mathbf{x}, \mathbf{x} > \\ &\geq 0. \end{aligned}$$

Therefore B is positive semidefinite. □

Theorem 3.4. ***Schur Product Theorem.*** *Suppose A and B are positive semidefinite matrices of size n. Then $A \circ B$ is also positive semidefinite.*

Proof. (141, [2]). Let A and B be positive semidefinite matrices of size n. Suppose B is of rank zero. This is true if and only if $B = \mathcal{O}$, and therefore $A \circ B = \mathcal{O}$, which is clearly positive semidefinite.

Now suppose B is of rank one. Then we can write $B = \mathbf{x}\mathbf{x}^T$ for some vector $\mathbf{x} \in \mathbf{C}^n$ (Lemma 3.1). Then $[A \circ B]_{ij} = [A]_{ij}[B]_{ij} = [A]_{ij}[\mathbf{x}]_i[\mathbf{x}^T]_j = [D_\mathbf{x}AD_\mathbf{x}]_{ij}$. Now, since A is positive semidefinite, then so is $D_\mathbf{x}AD_\mathbf{x}$. Take any vector $\mathbf{v} \in \mathbf{C}^n$. Then

$$\begin{aligned} < D_\mathbf{x}AD_\mathbf{x}\mathbf{v}, \mathbf{v} > &= < AD_\mathbf{x}\mathbf{v}, (D_\mathbf{x})^*\mathbf{v} > \\ &= < AD_\mathbf{x}\mathbf{v}, D_\mathbf{x}\mathbf{v} > \\ &= < A(D_\mathbf{x}\mathbf{v}), (D_\mathbf{x}\mathbf{v}) > \\ &\geq 0. \end{aligned}$$

Now suppose B is of rank r, $1 < r \leq n$. Then we can decompose B into a sum of rank one matrices B_1, ..., B_r (Theorem ROD, [3]) where each matrix B_i is also positive semidefinite (Lemma 3.2). Then $A \circ B = A \circ (B_1 + \ldots + B_r) = A \circ B_1 + \ldots + A \circ B_r$ (Theorem 1.5). We know that each $A \circ B_i$ is positive semidefinite for each i, and so $A \circ B$ is positive semidefinite (Lemma 3.3).

Therefore for any two positive semidefinite matrices A and B, $A \circ B$ is also positive semidefinite. □

4 Some Inequalities

Now for some matrix analysis. In mathematics, the term "analysis" means there are tons of inequalities (I have no proof for this). Recall the importance of the Schur Product Theorem is that if two matrices A and B are positive semidefinite, then so is $A \circ B$. This allows us to compare determinants of these matrices, since they are always nonnegative, real numbers.

Theorem 4.1. ***Oppenheim's Inequality.*** *Let A and B be positive semidefinite matrices of size n. Then $|A \circ B| \geq [A]_{11} \cdots [A]_{nn}|B|$.*

I am not including a proof of Oppenheim's Inequality because it is somewhat long and requires a bit of setup. However, it is not hard to follow and so instead I will simply refer you to page 144 of Bapat [2].

Theorem 4.2. ***Hadamard's Inequality.*** *Suppose A is positive semidefinite of size n. Then* $|A| \leq [A]_{11} \cdots [A]_{nn}$.

Proof. Let A be any positive semidefinite matrix of size n. Note that I_n is a postive semidefinite matrix of size n. Now we have the following:

$$\begin{aligned} |A| &= [I_n]_{11} \cdots [I_n]_{nn}|A| \\ &\leq |I_n \circ A| && \text{(Oppenheim's Inequality)} \\ &= [A]_{11} \cdots [A]_{nn} \end{aligned}$$

□

Corollary 4.3. *Let A and B be positive semidefinite matrices of size n. Then $|A \circ B| \geq |AB|$.*

Proof.

$$\begin{aligned} |A \circ B| &\geq [A]_{11} \cdots [A]_{nn}|B| && \text{(Oppenheim's Inequality)} \\ &\geq |A||B| && \text{(Hadamard's Inequality)} \\ &= |AB| \end{aligned}$$

□

Theorem 4.4. *Let A and B be positive semidefinite matrices of size n. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A \circ B$ and let $\widehat{\lambda_1}, \ldots, \widehat{\lambda_n}$ be the eigenvalues of AB. Then*

$$\prod_{i=k}^{n} \lambda_i \geq \prod_{i=k}^{n} \widehat{\lambda_i}, \quad k = 1, 2, \ldots, n. \quad \textit{(144, [2])}$$

Note that if $k = 1$, then this is true by Corollary 4.3.

This theorem dates back to Oppenheim in 1930, but has only recently been proven by Ando [1] and Visick [6] in 1995 (144, [2]).

This last inequality does not require A and B to be positive semidefinite, although the statement can be made stronger if they are (107 [5]).

Theorem 4.5. *Let A and B be square matrices of size n. Then rank $(A \circ B) \leq$ (rank A)(rank B).*

Proof. We will write the rank one decompositions of A and B. Suppose A has rank ρ_1 with eigenvalues λ_k, $1 \leq k \leq n$ and B has rank ρ_2 with eigenvalues $\widehat{\lambda_l}$, $1 \leq l \leq n$. Then by the constructive proof of Theorem ROD [3]

$$A = \sum_{k=1}^{\rho_1} \lambda_k \mathbf{x}_k \mathbf{y}_k^T, \quad \text{and} \quad B = \sum_{l=1}^{\rho_2} \widehat{\lambda_l} \mathbf{v}_l \mathbf{w}_l^T.$$

Then

$$\begin{aligned}
[A \circ B]_{ij} &= [A]_{ij}[B]_{ij} \\
&= \left[\sum_{k=1}^{\rho_1} \lambda_k \mathbf{x}_k \mathbf{y}_k^T\right]_{ij} \left[\sum_{l=1}^{\rho_2} \widehat{\lambda}_l \mathbf{v}_l \mathbf{w}_l^T\right]_{ij} \\
&= \sum_{k=1}^{\rho_1} \sum_{l=1}^{\rho_2} [\lambda_k \mathbf{x}_k \mathbf{y}_k^T]_{ij} [\widehat{\lambda}_l \mathbf{v}_l \mathbf{w}_l^T]_{ij} \\
&= \sum_{k=1}^{\rho_1} \sum_{l=1}^{\rho_2} [((\lambda_k \mathbf{x}_k) \circ (\widehat{\lambda}_l \mathbf{v}_l))(\mathbf{y}_k \circ \mathbf{w}_l)^T]_{ij} \\
&= \left[\sum_{k=1}^{\rho_1} \sum_{l=1}^{\rho_2} ((\lambda_k \mathbf{x}_k) \circ (\widehat{\lambda}_l \mathbf{v}_l))(\mathbf{y}_k \circ \mathbf{w}_l)^T\right]_{ij}
\end{aligned}$$

So $A \circ B$ has at most rank $\rho_1 \rho_2 = (\text{rank } A)(\text{rank } B)$. □

5 Conclusion

As we have seen, the Hadamard product has nice properties when we work with diagonal matrices and regular matrix multiplication, and when we work with positive semidefinite matrices. These properties (and numerous others that have been discovered; see references) can help us analyze matrices and understand the freedoms and limitations of the rank, eigenvalues and diagonal values of related matrices.

6 Addendum: License

References

[1] T. Ando. Majorization relations for hadamard products. *Linear Algebra and its Applications*, 223 - 224:57 – 64, July 1995.

[2] R. Bapat. *Nonnegative Matrices and Applications*. Cambridge University Press, 1997.

[3] Robert A. Beezer. *A First Course in Linear Algebra*. Robert A. Beezer, 1.08 edition, 2007.

[4] Roger Horn. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

[5] Charles Johnson. *Matrix Theory and Applications*. American Mathematical Society, 1990.

[6] George Visick. A weak majorization involving the matrices $a \circ b$ and ab. *Linear Algebra and its Applications*, 223 - 224:731 – 744, July 1995.