Review of Matrix Theorems PhL 12.18.04

********** **M&M Chapter 10 notes** **********

All lemmas, theorems, facts given here are proven in the notes, so only results are quoted here.

10.2 Determinants **********

$$\det a = \varepsilon_{\ell mn....q}\, a_{1\ell}\, a_{2m}\, a_{3n}.....a_{Nq} = = \varepsilon_{\ell mn....q}\, a_{\ell 1}\, a_{m2}\, a_{n3}.....a_{qN}$$

Theorem 2Z. det(a) = det(a^T), so determinant is the same if you swap rows and columns.

Theorem 2A: If a row = 0, then det = 0. // same for col

Lemma 2B for next Theorem: $\varepsilon_{\ell mn....q} M_{\ell mn....q} = 0$ if M is symmetric in *any* pair of indices

Theorem 2C: If one row is a multiple of another, then det = 0. // same for col

Theorem 2D: Add a multiple of one row to another row, det does not change.

Lemma 2E for next Theorem: If you swap two indices on M, sign of $\varepsilon_{\ell mn....q} M_{\ell mn....q}$ changes.

Theorem 2F: Swap two columns and sign of det changes.

Theorem 2G: Multiply a column by k, you then multiply det by k.

Theorem 2H. The determinant of an **upper triangular matrix** is the product of the diagonal elements.

10.3 Minors and Cofactors **********

Define: **minor(a_{pq})** = det of the N-1 x N-1 matrix you get by crossing out row p and column q.

Define: **cof(a_{pq})** = $(-1)^{p+q}$ minor(a_{pq}).

Theorem 3A: $\det(A) = \Sigma_q\, a_{pq}\ \mathrm{cof}(a_{pq})$ [across row p] = $\Sigma_p\, a_{pq}\ \mathrm{cof}(a_{pq})$ [down column q] (10-2)

Theorem 3B: $\Sigma_q\, a_{pq}\ \mathrm{cof}(a_{p'q}) = 0$ when $p \neq p'$. Also $\Sigma_p\, a_{pq}\ \mathrm{cof}(a_{pq'}) = 0$ when $q \neq q'$

Definition: The **adjoint matrix** of A, written $A^{\wedge}$, is the transpose of the cofactor matrix, $A^{\wedge} \equiv \mathrm{cof}(A)^T$. Note that the word "adjoint" means different things to different authors, so beware. This is what it means to M&M.

Theorem 3C: For any square matrix A, $A^{-1} = (1/\det(A))\ \mathrm{cof}(A)^T = (1/\det(A))A^{\wedge}$, provided $\det(A) \neq 0$, which is to say, provided matrix A is not singular.

10.4 Multiplication and Differentiation of Determinants **********

Theorem 4A: $\det(AB) = \det(A)\det(B)$

Theorem 4B: $\partial |A|/\partial a_{pq} = \text{cof}(a_{pq})$

10.5 Preliminary Remarks on Matrices **********

Def: A square matrix A is **singular** if det(A) = 0. All rectangular matrices are defined as singular, where rectangular means non-square.

Def: **Rank of A**: If $\det(A) \neq 0$, A has full rank N. If det(A) = 0 but if ALL det(minors) $\neq 0$, then rank is N-1. And so on. Rank applies only to a square matrix.

Therefore, if A is non-singular, rank = N, but if A is singular, rank < N.

Theorem 5A: If rank A = r, then we know two things:

(1) *all* RxR sub-determinants of A for R > r vanish.

(2) *at least one* rxr sub-determinant does *not* vanish

10.6 Combinations of Matrices **********

10.7 Special Matrices **********

Def: **trace** is the sum of the diagonal elements.

Theorem 7B: Tr(AB) = Tr(BA), whether or not the matrices commute (Proof: $A_{ij}B_{ji}$ = reverse)

Theorem 7C: $\text{Tr}(A\otimes B) = \text{Tr}(A)\text{Tr}(B)$. (Proof: from definition of the cross product.)

def**: Transpose** Matrix: $\tilde{A}_{k\ell} = A_{\ell k}$ (I usually say A^T for transpose, they use twiddle over the A)

def**: Adjoint** Matrix: $\hat{A}_{pq} = \text{cof}(a_{pq})^T$ (I don't ever use this matrix, beware the name, see above.)

Theorem 7D: $\hat{A}A = E * \det(A) = A\hat{A}$

Theorem 7E: If A is non-singular, then $A^{-1} = \hat{A}/|A|$.

Fact: If you transpose or invert a product of matrices, you have to change the order.

def: **Complex Conjugate Matrix:** $A^*_{ij} = a_{ij}^*$

def: **Associate Matrix**: $A^\dagger = A^{*\sim} = A^{*T}$

Warning: In Stakgold, the matrix $A^\dagger$ is written A*, and is called the adjoint of A. This is a double confusion because (1) adjoint means cof^T to M&M, and (2) * means C.C. to M&M. Stakgold then must use the overbar for C.C. In Quantum Mechanics, $A^\dagger$ is usually called the "Hermitian conjugate" of A.

10.8 Real Linear Vector Space **********

Theorem 8A: If basis vectors are linearly dependent, the **Gram Determinant** shown on page 312 vanishes. Conversely, if it does not vanish, then you have a basis by using the rows or columns.

10.9 Linear Equations **********

Def: A set of vectors $\mathbf{v_i}$ is **linearly independent** in E^n if *none* of them can be written as a lincom of some other ones.

Def: A set of vectors $\mathbf{v_i}$ is **linearly dependent** if the set is *not* linearly independent. In this case, at least one of the vectors must be writeable as a lincom of some others. If none were so writeable, then you would have linear independence. So it is an either/or situation with a set of vectors.

Phil Theorem 9.1: (General E^n) In E^n if the vectors $\mathbf{v_i}$ are linearly independent, then $\Sigma\, x_i\, \mathbf{v_i} = 0 \Rightarrow x_i = 0$. The converse is also true: $\Sigma\, x_i\, \mathbf{v_i} = 0 \Rightarrow x_i = 0 \Rightarrow$ the vectors $\mathbf{v_i}$ are linearly independent.

Phil Theorem 9.2. If the N columns of a matrix A (possibly not square) are linearly independent, then $Ax=0 \Rightarrow x=0$. The converse is also true.

Phil Theorem 9.3: If the rows *or* cols of square matrix A are linearly dependent, then det(A) = 0.
Contrapositive: If $\det(A) \neq 0$, then the rows *and* cols of A are linearly independent.

Phil Theorem 9.4. If $\det(A) \neq 0$, then [$Ax=0 \Rightarrow x=0$].

Phil Theorem 9.5. If the columns or rows of A are linearly independent, then $\det(A) \neq 0$.
Contrapositive: if detA = 0, then columns and rows are linearly dependent.

Phil Theorem 9.6. Linear dependence of the columns $\Leftrightarrow$ Linear dependence of the rows.
Contrapositive: Linear independence of the columns $\Leftrightarrow$ Linear independence of the rows.

Phil Summary Theorem 9.7: By combinations of the above theorems, we can show that:

The following items all mean the same thing: (any one implies all the others)

(1) detA = 0, A is singular and rank(A) < N
(2) rows of A are linearly dependent
(3) columns of A are linearly dependent
(4) Ax = 0 does not imply that x=0, there can be non-zero solutions
(5) A^{-1} does not exist

The following items all mean the same thing: (any one implies all the others)

(1) $\det A \neq 0$, A is non-singular and rank(A) = N
(2) rows of A are linearly independent
(3) columns of A are linearly independent
(4) $Ax = 0 \Rightarrow x=0$

(5) A^{-1} exists

Define: **Nullspace of A**: the set of all vectors x that solve Ax=0 span a space of some dimension $k \leq N$ that is called the nullspace of A. Stakgold shows that this space is in fact a subspace of E^N. At this time, we have no theorems about nullspaces.

Phil Theorem 9.8 (Cramer's Rule) : If Ax=y and detA $\neq 0$, then get $x_q = | A(q \text{ col} \leftarrow y) | / | A |$.

10.11 Equivalent Matrices **********

Definition: B = PAQ where both P and Q are non-singular, then A and B are **equivalent**.

(a) $B = Q^{-1}AQ$ means A and B are equivalent and this is a **similarity** (= **collineatory**) transformation.

(b) $B = Q^TAQ$ means A and B are equivalent and A and B are **congruence** transformation.

(d) If it happens that $Q^T = Q^{-1}$ and Q has real elements (Q is **real orthogonal**), then $Q^{-1}AQ$ is a **real orthogonal** transformation. Since we have $Q^{-1}AQ = Q^TAQ$, such a transformation is both a similarity and a congruence at the same time.

(c) $B = Q^\dagger AQ$ means A and B are equivalent and A and B are called **conjunctive** transformation.

(e) If it happens that $Q^\dagger = Q^{-1}$ (Q is **unitary**) then $Q^{-1}AQ$ is a **unitary** transformation. Since we have $Q^{-1}AQ = Q^\dagger AQ$, such a transformation is both a similarity and a conjunct at the same time.

10.12 Bilinear and Quadratic Forms **********

Def: $A(x,y) = y^TAx$ is called **bilinear form**.

Def: $A(x,x) = x^TAx$ is called a **quadratic form**.

10.13. Similarity Transformations **********

10.14 The Characteristic Equation of a Matrix **********

Def: $K(\lambda) = + \det(\lambda E - A)$ is the **characteristic equation** of A. // notice choice of sign inside det
In other books, it is called the **secular equation**.

Lemma 14A: For a diagonal matrix, $\det(D) = \Pi(D_{ii})$.

Corollary: det(E) = 1.

Lemma 14B: $\det(Q) = 1/\det(Q^{-1})$.

Lemma 14C: If $B = Q^{-1}AQ$, then det(B) = det(A).

Theorem 14D: If A and B are related by a similarity, the two matrices have the same eigenvalues.

Theorem 14E: If A and B are related by a similarity, the two matrix have the same trace.

10.15 Reduction of a Matrix to Diagonal Form **********

Theorem 15A: The eigenvalues of a diagonal matrix are the diagonal values.

Theorem 15B: *If* we can find a similarity such that $Q^{-1}AQ = \Lambda$ where Λ is a diagonal matrix, then the eigenvalues of A are the diagonal elements of Λ.

Theorem 15C: In the case that all the eigenvalues are different, the eigenvectors are linearly independent.

Theorem 15D: If you have an eigenvalue problem $A\mathbf{x_i} = \lambda_i\mathbf{x}_i$, if you create a matrix X whose columns are the x_i, then you can write $AX = X\Lambda$ where Λ = diag (λ_1, λ_2.....).

Theorem 15E: Let X be the matrix of eigenvectors of the problem $A\mathbf{x_i} = \lambda_i\mathbf{x}_i$. If $\det X \neq 0$, then we can diagonalize the matrix A by a similarity with the matrix X: $X^{-1}AX = \Lambda$ where Λ = diag (λ_1, λ_2.....).

Fact: When the λ_i are all different, we know from theorem 15C that $\det X \neq 0$ so then $X^{-1}AX = \Lambda$.

The Block Diagonal Form Theorem: Any matrix A with its spectrum of eigenvalues and algebraic multiplicities can be brought completely into block-diagonal form (which is overall an upper triangular form) by some NxN similarity transformation X. The eigenvalue spectrum can be obtained from the secular equation $\det(A - \lambda I) = 0$. All blocks have the upper triangular form shown in M&M 10-41 where the eigenvalue for the block appears on all the diagonal elements of the block. Obviously when algebraic multiplicity = 1, the block is 1x1 and consists of nothing but the eigenvalue. The blocks are of course all square, being k x k where k is the algebraic multiplicity of that eigenvalue. The issue of the geometric multiplicity m within each block can be studied locally within that block, and as we know, it can range anywhere from 1 up to the algebraic multiplicity k for that block.

Comment: Thus, any matrix A can be brought to *triangular form* by a similarity. Later we learn that the similarity can be found by the QR Algorithm which I think differs from what M&M present and which I call the pivot and twist. A symmetric matrix A can then be brought to *diagonal form* by a similarity. In either case, you cannot just "write down" the similarity as a simple function of the elements of A, you have to find it by an iterative method. (Yet M&M's method seemed not to require iteration.)

10.16 Congruent Transformations **********

Lemma: The product of two upper triangular matrices is upper triangular.

Theorem 16A: A symmetric matrix A can be diagonalized by a congruence transformation $B = Q^T A Q$ where B is diagonal and where Q is an upper triangular matrix.

Fact: This is a useful fact if you are trying to compute the quadratic form x^TAx, because the result is then simply $\Sigma B_{ii} y_i^2$ where $x = Qy$.

Fact: A congruence transformation does NOT preserve eigenvalues because $\det(Q^T Q) \neq 1$ in general. But in this section, we don't care about the eigenvalue problem!

Fact: As they note, the Q and B_{ii} you obtain are not unique! The *value* of the quadratic form however *will* be unique since it is always going to be equal to $x^T Ax$ regardless of Q. So the terms in the sum change, but the sum stays the same.

10.17 Orthogonal Transformations **********

Theorem 17A: if R is real orthogonal, meaning $R^{-1} = R^T$, then $\det(R) = \pm 1$.

Theorem 17B: R ro preserves the length of a vector, so that $\| y \|^2 = y^T y = \| x \|^2$.

Theorem 17C: The eigenvalues of a real orthogonal (or Hermitian) matrix all have $| \lambda | = 1$. Complex eigenvalues must occur in pairs.

Theorem 17D: Any real orthogonal matrix R can be diagonalized to Λ by a unitary matrix U. The eigenvalues are complex but unimodular.

10.18 Hermitian Vector Space **********

10.19 Hermitian Matrices **********

Theorem 19A: (x,Ax) is real if A is Hermitian

Theorem 19B: The eigenvalues of A are real if A is Hermitian

Theorem 19C: A Hermitian matrix can be diagonalized by a unitary similarity.

Theorem 19D: Hermiticity is conserved under a similarity by any unitary matrix.
Just the way symmetry is conserved under similarity by a real orthogonal.

Theorem 19E: If two matrices commute, they can both be diagonalized by the *same* unitary similarity. The converse is also true.

10.20 Unitary Matrices **********

Theorem 20A: Any unitary U can be diagonalized to Λ by a unitary matrix V. The eigenvalues are complex but unimodular.

********** **M&M Support Notes** **********

Projection Theorem S1: Suppose space A of dimension N has a subspace S of dimension M. Then any vector in A can be decomposed as v = s + r where s is in S, and r is in $S^{\perp}$. We can write $A = S + S^{\perp}$. We

know that r•s = 0. We could define projections such that P_S v = s and $P_{S\perp}$ v = r and of course $P_S + P_{S\perp} = I$. But in general $S^{\perp}$ will not also be a subspace.

Projection Theorem S2 (Generalization). For a general matrix A that has several subspaces (such as nullspaces of A in this case, Ax = 0) and a residual, you could perhaps say that $B = S_1 + S_2 + S_3 + R$ where we have three subspaces and whatever is left over. We could then arrange our vectors in this space to match and have the form $x = s_1 + s_2 + s_3 + r$. We could make a projection operator into each chunk. Notice that in general, R is not likely to be a subspace because elements in it might map into a combination of spaces under the action of $(A-\lambda I)x$, but we could still wrote projections for each section. When B is Hermitian and $A = (B - \lambda I)$, the nullspaces of A are the eigenmanifolds of A and there is no residual piece R.

Theorem S3: If you map a set of N vectors $\mathbf{x_i}$ to a new set $\mathbf{x_i'}$ using $R\mathbf{x_i} = \mathbf{x_i'}$ where $\det R \neq 0$, then if the original vectors were linearly independent, then so are the final vectors.

Theorem S4: If $\det R \neq 0$, then $R\mathbf{e_I} = \mathbf{0}$ is not possible.

Theorem S5: Consider an NxN eigenvalue problem $A\mathbf{x_i} = \lambda_i\mathbf{x_i}$ in which it happens that the N eigenvectors form an orthonormal basis. If you transform this problem in the usual similarity way to $A'\mathbf{x_i}' = \lambda_i\mathbf{x_i}'$ by a transformation $\mathbf{x_i}' = R\mathbf{x_i}$ with the only restriction being that $\det R \neq 0$, then in the new problem, the vectors $\mathbf{x_i}'$ still form a basis, but they will in general be neither normalized nor orthogonal. If it happens that R is a real orthogonal transformation (a "rotation"), then the new basis will still be orthonormal.

Theorem S6: Consider the world of 2x2 matrices. If the two eigenvalues are different, there are two distinct non-zero eigenvectors, and geometric multiplicity = algebraic multiplicity = 1 for each eigenmanifold. If the two eigenvalues are the same, there is exactly one non-zero eigenvector, so geometric multiplicity = 1 and algebraic multiplicity = 2. There are no 2x2 matrices which have geometric multiplicity = 0 (nor are there of any size!)

Theorem S7: Consider the problem $Ax_i = \lambda_i x_i$. Find the eigenvalues as usual and their algebraic multiplicities. Within each eigenmanifold, here is how we can compute the geometric multiplicity for that manifold

geometric multiplicity = nullity(Ax - λI) = N - rank(Ax - λI)

********** **Some Matrix Theorems (doc)** **********

Example: Compute eigenvalues/vectors for symmetric real 2x2 matrix.

Theorem 0: A real matrix A can have complex eigenvalues the complex eigenvectors.

Definition: If Ax=0, the **nullity** of A is the dimensionality of the nullspace of A, which is in turn the number of distinct non-zero vectors that solve Ax = 0.

Definition: The **geometric multiplicity** of an eigenvalue λ of the problem $Ax=\lambda x$ is the nullity of $(A-\lambda E)$, that is, it is the dimensionality of the nullspace of $(A-\lambda E)$. It is thus the number of independent non-zero solutions x of the equation $(A-\lambda E)x = 0$.

Definition: The **algebraic multiplicity** of an eigenvalue is the number of times it appears in the factored secular equation.

Theorem 1 : Geometric multiplicity $m \leq$ Algebraic multiplicity k.

Corollary: For any matrix A of order N, if λ is an eigenvalue of algebraic multiplicity k, then there can be at most k distinct eigenvectors associated with λ.

Theorem 2: For any NxN matrix A, if λ is an eigenvalue of algebraic multiplicity k, the number m of distinct eigenvectors associated with λ is given by $m = N - \text{rank}(A-\lambda E)$ and $1 \leq m \leq k$.

Corollary: There is always at least one eigenvector for a given λ. (just because $m \geq 1$) A matrix always has N eigenvalues, but they can clump in sameness groups, eg, all could be the same.

Theorem 3: Let A be an NxN complex matrix which has only one eigenvalue λ and its corresponding single eigenmanifold. Matrix A therefore has algebraic multiplicity $k = N$. In general, the geometric multiplicity m is $m \leq N$ (another theorem). **However, if A is Hermitian, then m =k = N.**

Theorem 4: The only Hermitian matrix having just one eigenvalue λ is $A = \lambda E$.

Theorem 5: If A is Hermitian, then **each eigenmanifold has k = m**. That is to say, for each eigenmanifold, the geometric multiplicity has its maximal value and it equal to the algebraic multiplicity.

Theorem 5A. If A is Hermitian, then the nullspaces of $(A - \lambda E)x=0$ *partition* the whole space. You can then decompose any vector f into components each of which is in one of these nullspaces. These nullspaces are of course the eigenmanifolds of $Ax = \lambda x$. There are no vectors x in the whole space that are *not* elements of one of the nullspaces. For A non-Hermitian, we showed a 2x2 example earlier where we *could* find elements that were in non of the nullspaces.

Theorem 6: An NxN Hermitian matrix has **N distinct eigenvectors.**

Theorem 7: All eigenvalues of a Hermitian matrix **are all real**.

Theorem 8: For any Hermitian matrix A in complex space, the eigenvectors corresponding to different eigenvalues are **orthogonal**.

Theorem 9: A Hermitian NxN matrix A has N eigenvectors which are **linearly independent** and thus form a basis in $E^{N.}$ It is possible to make the eigenvectors form an **orthogonal** or **orthonormal** set.

Theorem 10: The components of the eigenvectors of a Hermitian matrix A are in general complex.

Theorem 11: An NxN Hermitian matrix A can be diagonalized by a unitary similarity X, and X is unique only if there are no degenerate eigenmanifolds for A. A can also be diagonalized by a huge set of non-unitary X similarities.

Definition: The **rank** of a matrix is the size of the smallest non-zero sub-determinant. If detA $\neq 0$, then the smallest non-zero det is the whole NxN, so rank = N. If A has all zeros, then rank = 0.

Theorem R1: If the **rank** of an NxN matrix A is N-k, then there are only k eigenvectors of Ax=0.

Theorem R2: If a row in a matrix is linearly dependent on the other rows, then any segment of that row will be linearly dependent on the corresponding segments of the other rows.

Theorem R3: If the rank of an NxN matrix A is N-k, then k rows of A are linearly dependent on other rows. Only N-k rows of A are linearly independent. The same is true for the columns. This means that there are k "relations" among the rows, if k are linearly dependent on the others.

Theorem R4: $\dim(R_A)$ = rank [the dimensionality of the range in Ax=f is equal to the rank of A.]

Theorem R5: **(Alternative Theorem)** $R_A^{\perp} = N_{A\dagger}$. (The perp-space of the range of A is equal to the nullspace of $A^{\dagger}$.)

Theorem R6: For any NxN matrix A, rank(A) + nullity(A) = N.
Also, nullity(A) = nullity($A^{\dagger}$) and rank(A) = rank($A^{\dagger}$)

********** **Matrix Research (doc)** **********

Theorem M1: If C = AB, then r(C) $\leq$ min[r(A) , r(B)], A and B any conformable matrices. [r = rank]

Theorem M2: If C = AB where A is non-singular, then r(C) = r(B) and vice versa A $\leftrightarrow$ B.

Definition: **ERO** = Elementary Row Operation : Scale a row by nonzero, swap rows, or add multiple of a row. None of these things changes truth of the set of equations Ax = y.

Definition: **Row Echelon Form**.
This is what you get by applying ERO's to a matrix to clear out the lower triangle, the thing we called Gauss elimination in Scheid. Zero rows (if any) are put on the bottom. The leftmost entry in any row is a 1. The remaining entries in each row could be anything. You swap rows in an effort to maintain the leading diagonal of ones, but if you fail, then it jogs over.

Definition: **Reduced Row Echelon Form**. (aka **Normal Form**)
If you first go to the Row Echelon form and then you continue to clear out the upper triangle as well, this is what you get. I called this Gauss-Jordan elimination in my Scheid notes. In this form, each column has only a single entry. Happy columns have a sole 1, but where you had to jog over in doing Gauss elimination, you can have something other than a 1.

Definition: A is **Row Equivalent** to B means that you can get from A to B using EROs, B = EA.

Definition: A is **Equivalent** to B means A is either row or column equivalent. Write **A.eq.B**. Notice that this is a different meaning from saying that

Theorem M3: The Reduced Row Echelon form of a matrix A is unique.

Theorem M4: EROs do not change the rank of a matrix.

Theorem M5: EROs do change the determinant of a matrix in a predictable manner, but cannot cause a non-vanishing determinant to vanish.

Theorem M6: An ERO is invertible and thus non-singular.

Theorem M7: The rank of a matrix equals the number of non-zero rows in its reduced row echelon form.

Theorem M8. Row and Column Operations on A can be done by A' = RAC, and R and C are non-singular.

Theorem M9. (**The LU Factorization**.)

Theorem M10: (**PLU Factorization**). Any non-singular square matrix A can be written as PLU where P is a permutation matrix, L is lower-triangular with 1's on the diagonal, and U is upper triangular. The matrix P is a some permutation of the unit column vectors. [Maple does this with LUdecomp]

Definition: A and B are **equivalent** if there exist non-singular P and Q where A = PBQ.

Theorem M11: Equivalent matrices have the same rank.

Theorem M12: (**Submatrix Theorem**) Let matrix B be a horizontal slice of NxN matrix A, where B is S rows high. Then $r(A) \geq r(B) \geq r(A) + S - N$.

Theorem M13: (**Sylvester's Law of Nullity**) If A and B are NxN, then $r(AB) \geq r(A) + r(B) - N$. In terms of nullity, $n(AB) \leq n(A) + n(B)$. Our previous theorem said $r(AB) \leq \min[r(A),r(B)]$ so the combined result is then

$$r(A) + r(B) - N \leq r(AB) \leq \min[r(A),r(B)]$$

Rewrite as

$$\max[n(A),n(B)] \leq n(AB) \leq n(A) + n(B)$$ // triangle rule!

Theorem M14: $r(A+B) \leq r(A) + r(B)$

Definition: A **Householder Matrix**: given any real vector u , the matrix is $H = 1 - \sigma uu^T$ where the number $\sigma = (2/u^Tu)$. Think of H as H(u). [aka an **Elementary Reflector**] Can write as $H = 1 - 2ww^T$ if use w that is normalized.

Theorem M15: The following are true for a Householder Matrix H:
(1) $H = H^T$ which is true for any constant σ
(2) $HH = 1$ (or $H^{-1} = H$) which is true only for the specific σ shown in the definition
(3) $H^T = H^{-1}$ so H is real orthogonal (follows from 1 and 2)
(4) $(uu^T)^2 = (u^Tu)(uu^T)$ is true for any u
(5) $(uu^T)v = (u^Tv)u$ is true for any u and any v

Theorem M16: The product of a Householder matrix and a real orthogonal matrix is a real orthogonal matrix.

Corollary: The product of two Householder matrices is a real orthogonal matrix.

Theorem M17: One can find a Householder matrix H such that $Hv = (\beta,0,0...)$ for any vector v, where the value of $\beta = \pm \text{sqrt}(v_1^2 + v_2^2 + ... + v_N^2) = \pm |v|$. Usually one takes β to have the same sign as v_1.

Theorem M18: (**QR Factorization**). Any square matrix A can be written as A = QR where Q is real orthogonal, and R is upper triangular.

Theorem M19: (QR factorization for non-square matrices) $A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$ or $A = Q (R_1, R_2)$ as sketched below.

Theorem M20: omitted because it is wrong.

Theorem M21: $\text{rank}(A \otimes B) = \text{rank}(A)\text{rank}(B)$

Theorem M22: $\text{rank}(A \oplus B) = \text{rank}(A) + \text{rank}(B)$

Theorems M23-27: repeats of earlier Householder things with Hermitian instead of symmetric.

Theorem M28: (**complex LQ Factorization**). Any square matrix A can be written as A = LQ where Q is unitary, and L is lower triangular.

Theorem M29: (LQ factorization for non-square matrices) $A = \begin{pmatrix} L1 \\ L2 \end{pmatrix} Q$ or $A = LQ_1$ as sketched below. (see doc for sketch)

Theorem M30: The inverse T^{-1} of an upper triangular matrix T is upper triangular, and the diagonal elements of T^{-1} are the inverses of those of T. This only makes sense if all diagonal elements of T are non-zero, otherwise $\det T = 0$ and T^{-1} does not exist.

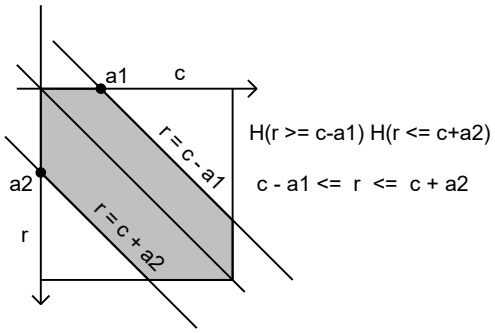
Theorem M31: The product of two upper triangular matrices is upper triangular.

Theorem M32: Let A = triangular, B = partially triangular in same sense. Then C = AB and D = BA are both partially triangular in the same sense as B. { same theorem below in terms of Hessenberg n }

Theorem M33: If you multiply two (partially) triangular matrices of the same sense, the result has the shape of the factor matrix having the fewest zeros, ie, the one that is least triangular.

Theorem M34: In the QR iteration scheme we factor A = QR, where Q is real orthogonal and R is upper triangular, and then define A' = RQ = Q^{-1} A Q = Q^T A Q. The transformation from A to A' preserves Hessenberg form.

Definition. A **banded matrix** A with c - $a_1 \le r \le c + a_2$ means that $A_{rc} = 0$ for r,c outside the range shown. Here is a picture of this banded matrix A (integers a_1 and a_2 are assumed non-negative)



Example #1: An **upper triangular** (**UT**) matrix A has a1 = N and a2 = 0

Example #2: An **generalized upper Hessenberg** (**UHn**) matrix has a1 = N and a2 = n. A standard issue upper Hessenberg has n=1 and we write as UH1.

Example #3: A **tridiagonal** matrix has a1 = a2 = 1.

Theorem M35: If A is a banded matrix with c - $a_1 \le r \le c + a_2$, then if $a_1 \ge N$ we can ignore the left inequality, and if $a_2 \ge N$ we can ignore the right inequality.

Theorem M36 (Banded Matrix Theorem): Consider C = AB where A is banded by c - $a_1 \le r \le c + a_2$ and where B is banded by c - $b_1 \le r \le c + b_2$. Then C is banded by c - $c_1 \le r \le c + c_2$, where $c_1 = a_1 + b_1$, $c_2 = a_2 + b_2$.

Corollary M36.1. $UH_n * UH_m = UH_{n+m}$

Corollary M36.2: $UH_0 * UH_m = UH_m$ or UT*UH_m = UH_m

Corollary M36.3: $UH_0 * UH_0 = UH_0$ or UT*UT = UT

Corollary M36.4. UT * Tridiag = UH_1

Corollary M36.5. Let C = AB where A is general and B is tridiagonal. Then C is general.

Theorem M37: The QR Transformation preserves UHn.

Theorem M38: The QR Transformation preserves tridiagonality if A is symmetric

Theorem M39: The inverse of a UHn matrix is UHn.

Corollary M39.1: The inverse of a UT matrix is a UT matrix.

Theorem M40: If matrix A has an eigenvalue 0, then A is singular.

Theorem M41: Let $C(\lambda) = A(\lambda)B(\lambda)$ where $A(\lambda)$ and $B(\lambda)$ are polynomials with matrix coefficients. Then if $B(D) = 0$, it follows that $C(D) = 0$ [and this is true even though we know that $C(D) \neq A(D)B(D)$]

Theorem M42: Let $X(\lambda) = A(\lambda)B(\lambda).....F(\lambda)$ where all are polys with matrix coefficients. Then
(a) It is NOT true that $X(D) = A(D)B(D).....F(D)$ for any matrix D.
(b) It is true that $F(D) = 0 \Rightarrow X(D) = 0$.

Corollary M42.1: Suppose $P(\lambda) = Q(\lambda)(A - \lambda I)$ where P and Q are polynomials in λ but with square matrix coefficients, and of course $(A - \lambda I)$ is a matrix as is A. If we then construct P(A) by extending $P(\lambda)$ to a matrix argument as described above, then we find that $P(A) = 0$.

Theorem M43 (**Cayley-Hamilton**). If $K(\lambda) = 0$ is the secular equation for matrix A, then $K(A) = 0$. This says that every matrix A satisfies its own secular equation.

Corollary to Cayley-Hamilton. For any NxN matrix A, any positive power of A can be written as a linear combination of the N matrices $A^0 = I$, $A^1 = A$, A^2, A^3, ...A^{N-1}. If A is invertible, then the conclusion also applies to negative powers of A. [Note: there may be some $r < N$ for which this is also true.]

Theorem M44: $N_0 = N - r$; $N_+ = (r+s)/2$, $N_- = (r-s)/2$.
If two matrices are related by a non-singular congruence, they have the same N and r.
Thus, (same signature) $\Leftrightarrow$ (same inertia).

Theorem M45: (**Sylvester's Law of Inertia**) Signature is invariant under a congruence transformation.
My PDF notes state it as : (A is congruent with B) $\Leftrightarrow$ (A and B have the same inertia).
The inertia of a symmetric matrix is not affected by a congruence transformation on that matrix.