Math240; Linear algebra review

1. Matrix addition and multiplication.

- Size constraints: $(m\times n)+(m\times n)\mapsto m\times n$, $(m\times p)\cdot(p\times n)\mapsto m\times n$.
- Distributivity and commutativity: $(A+B)C = AC+BC$, $C(A+B) = CA+CB$, $A(BC) = (AB)C$ when defined.
- It is possible that $AB \neq BA$ or $AB = 0$ for non-zero 2×2 matrices. For example, $AB = 0 \neq BA$ where

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

- Rules for the transpose: $(A+B)^T = A^T + B^T$,$(AB)^T = B^T A^T$
- Any linear system may be written as $A\vec{x} = \vec{b}$; the system is homogeneous when $\vec{b} = 0$.

2. Gaussian and Gauss-Jordan elimination.

Elimination preserves

- the rank of a matrix (i.e. the number of linearly independent rows, or that is the same, the number of linearly independent columns);
- the solution set of a linear system

Determinant is not preserved by elimination, but the change is easily described. For example, switching two rows changes the sign.

These are the main cases when one should apply elimination:

- To solve a linear system $A\vec{x} = \vec{b}$ apply Gauss-Jordan to $(A|b)$.
- To find $\mathrm{rk}(A)$ apply Gauss to A.
- To find $|A|$ apply Gauss to A and keep track of what happens to the determinant.
- To find A^{-1} apply Gauss-Jordan to $(A|I)$.

A homogeneous $n\times n$ system $A\vec{x} = 0$ has a non-zero solution iff applying Gauss elimination to A we obtain a zero row, i.e. $\mathrm{rk}(A) < n$ (or A is singular).

3. Inverse matrix.

- If $AB = I$ for $n\times n$ matrices, then $BA = I$ and $B = A^{-1}, A = B^{-1}$.
- Properties: $(A^{-1})^{-1} = A$, $(A^T)^{-1} = (A^{-1})^T$,$(AB)^{-1} = B^{-1}A^{-1}$.
- The solution of a linear system $A\vec{x} = \vec{b}$ with invertible A is $x = A^{-1}b$.
- An $n\times n$ matrix is invertible (or non-singular) iff $\mathrm{rk}(A) = n$.

4. Determinants.

- $\det(A) = |A|$ is defined for any $n\times n$ matrix, it vanishes iff A is singular.
- Properties: $|AB| = |A||B|$, $|A^{-1}| = |A|^{-1}$, $|A^T| = |A|$, $|A^k| = |A|^k$ for any integer k.
- The determinant of a triangle matrix is the product of its diagonal entries.

- The determinant may be found by applying elementary operations on the columns (similarly to the usual elimination process which works with rows).

5. Cofactors.

- A cofactor C_{ij} equals to $(-1)^{i+j}$ times the determinant of the matrix obtained from A by removing i-th row and j-th column.
- i-th row decomposition: $|A| = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}$.
- j-th column decomposition: $|A| = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$.
- $A^{-1} = \frac{1}{|A|}(C_{ij})^T$ where (C_{ij}) is the matrix formed by the cofactors of A.
- Cramer's rule: a linear system $A\vec{x} = \vec{b}$ may be solved by $x_i = \frac{|A_i|}{|A|}$, where A_i is obtained from A by replacing the i-th column with b.

6. Characteristic polynomial and eigenvalues.

- The characteristic polynomial of an $n \times n$ matrix A is defined as $f_A(\lambda) = |A - \lambda I_n| = (-1)^n\lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_0$.
- $f_A(\lambda)$ is computed as a usual determinant, though diagonal entries are linear polynomials in λ: either cofactor or elimination method, whichever is easier.
- Cayley-Hamilton: $0 = f_A(A) = (-1)^n A^n + c_{n-1}A^{n-1} + \cdots + c_0 I_n$.
- $f_A(\lambda) = (-1)^n(\lambda - \lambda_1)\ldots(\lambda - \lambda_n)$, where $\lambda_1, \ldots, \lambda_n$ are the roots. If some λ_k is complex, say $a + bi$, then $\overline{\lambda}_k = a - bi$ is also a root.
- $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A taking account of multiplicities, these are exactly the values of λ such that the homogeneous system $(A - \lambda I)\vec{x} = 0$ has a non-zero solution, equivalently, $A\vec{x} = \lambda\vec{x}$.
- Change of coordinates via invertible matrix P preserves the characteristic polynomial, namely $f_A(\lambda) = f_{P^{-1}AP}(\lambda)$.
- $|A| = c_0 = \lambda_1 \ldots \lambda_n$ is the product of the eigenvalues with multiplicities.

7. Eigenvectors and eigenvalues.

- If $A\vec{x} = \lambda\vec{x}$ (or, that is the same, $(A - \lambda I_n)\vec{x} = 0$) with $\vec{x} \neq 0$, then $\vec{x}$ is an eigenvector and λ is an eigenvalue.
- There are exactly n eigenvalues if one takes into account the multiplicities.
- If λ is an eigenvalue of multiplicity m, then its eigenvectors form a space $\mathbf{R}^k$ where $1 \leq k \leq m$ (i.e. there are k linearly independent eigenvectors with eigenvalue λ). It is the solution set of the homogeneous linear system $(A - \lambda I_n)\vec{x} = 0$ (and k is the number of parameters).
- If $k < m$ (for some choice of λ), then there is no basis of eigenvectors of A. For example

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

has $f_A(\lambda) = \lambda^2$ and one multiple eigenvalue $\lambda_1 = \lambda_2 = 0$, but only one eigenvector $\vec{x} = \binom{1}{0}$ (that satisfies $A\vec{x} = 0$).

8. Diagonalization.

- A is called diagonalizable if $D = P^{-1}AP$ is diagonal for an invertible matrix P called the diagonalizing matrix.
- The characteristic polynomials of A and D are equal, in particular, the eigenvalues of A are $d_{11}, \ldots, d_{nn}$ and $|A| = d_{11} \ldots d_{nn}$.
- $P = (P_1, \ldots, P_n)$ diagonalizes A iff its columns $P_1, \ldots, P_n$ are linearly independent eigenvectors of A; then P is invertible and $D = P^{-1}AP$ is diagonal with $AP_i = d_{ii}P_i$.
- So, A is diagonalizable iff there exists a basis (i.e. n linearly independent vectors) formed by eigenvectors of A. In particular, if A has no multiple eigenvalues then A is diagonalizable.
- The 2×2 matrix from section 7 is not diagonalizable.

Practical diagonalization of A.

- Find the characteristic polynomial $f_A(\lambda) = |A - \lambda I_n|$.
- Find the eigenvalues by factoring $f_A(\lambda)$ as $(-1)^n(\lambda - \lambda_1) \ldots (\lambda - \lambda_n)$. In some cases, guess a root, say λ_1, find $(-1)^n(\lambda - \lambda_2) \ldots (\lambda - \lambda_n) = f_A(\lambda)/(\lambda - \lambda_1)$ by long division of polynomials, then guess another root, etc.
- Find the eigenvectors: for each eigenvalue λ of multiplicity m solve the linear system $(A - \lambda I_n)\vec{x} = 0$. If there are no m linearly independent solutions (i.e. less than m parameters) then A is not diagonalizable. Choose m linearly independent solutions otherwise.
- Set $P = (P_1, \ldots, P_n)$ where P_i's are the found eigenvectors, then $P^{-1}AP$ is diagonal, and its entries are the eigenvalues.

9. Orthogonal matrices.

- A is orthogonal if $A^{-1} = A^T$.
- $A = (A_1, \ldots, A_n)$ is orthogonal iff A_i's form an orthonormal basis, i.e. $A_i \cdot A_j = A_i^T A_j$ vanishes for $i \neq j$ and is 1 for $i = j$.
- If A and B are orthogonal then A^{-1} and AB are orthogonal, and $|A| = \pm 1$.
- A defines an orthogonal transformation on $\mathbf{R}^n$, i.e. it preserves the dot product: $(Ax) \cdot (Ay) = x \cdot y$.

10. Orthogonalization.

Any orthogonal basis $\vec{v}_1, \ldots, \vec{v}_n$ (i.e. a basis with $\vec{v}_i \cdot \vec{v}_j = 0$ for $i \neq j$) can be made orthonormal by scaling $\vec{v}_i$'s (namely, dividing $\vec{v}_i$ by their lengths).

An arbitrary basis can be orthogonalized by Gramm-Schmidt method, which uses the transformation $(\vec{u}, \vec{v}) \mapsto (\vec{u}' = \vec{u}, \vec{v}' = \vec{v} - \frac{\vec{u} \cdot \vec{v}}{\vec{u} \cdot \vec{u}}\vec{u})$. Notice that $\vec{u}'$ and $\vec{v}'$ are orthogonal, so re-scaling them one gets an orthonormal basis of the plane spanned by $\vec{u}$ and $\vec{v}$.

11. Symmetric matrices and orthogonal diagonalization.

- A is symmetric if $A = A^T$

- If A is symmetric, then A^n is symmetric for any integer n.
- Product of symmetric matrices may be not symmetric, for example
$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
- Any symmetric A is diagonalizable and all its eigenvalues are real.
- If $\vec{x}, \vec{y}$ are eigenvectors corresponding to different eigenvalues then $\vec{x} \cdot \vec{y} = 0$ (they are orthogonal).
- Thus, if A has no multiple eigenvalues then its eigenvectors form an orthogonal basis, and it becomes an orthonormal basis after appropriate re-scaling.
- A is called orthogonally diagonalizable if it may be diagonalized by an orthogonal matrix.
- A is orthogonally diagonalizable iff it is symmetric.
- Orthogonal diagonalization of A reduces to finding an orthonormal basis $P_1, \dots, P_n$ of eigenvectors of A.

Practical orthogonal diagonalization of A.

- Find a basis $P_1, \dots, P_n$ of eigenvectors of A as in usual diagonalization.
- If there are no multiple eigenvalues, simply re-scale P_i's.
- Otherwise apply to P_i's the orthogonalization procedure from section 10. You may work separately with each set of eigenvectors corresponding to the same eigenvalue.