Linear Algebra Review

- **Objects**
  - Scalars $\alpha \in \mathbb{R}$ or $\mathbb{C}$, Vectors $x \in \mathbb{R}^n$ (column), Matrices $A \in \mathbb{R}^{m \times n}$
  - Components: $(x)_i = x_i$, $(A)_{ij} = a_{ij}$, columns of $A$: $a_j$, rows of $A$: $a_i$

- **Actions**
  - Transpose: $(A^T)_{ij} = (A)_{ji}$
  - Scalar Product: $(\alpha x)_i = \alpha (x)_i$, $(\alpha A)_{ij} = \alpha (A)_{ij}$
  - Inner or Dot Product: $x \cdot y = x^T y = \sum_{i=1}^n x_i y_i$
    Vectors $x$ and $y$ are orthogonal if $x \cdot y = 0$.
  - Matrix-Vector Product: $(Ax)_i = a_i x = \sum_{j=1}^n a_{ij} x_j$
  - Matrix-Matrix Product: $(AB)_{ij} = a_{ik} b_{kj}$
    Number of columns of $A$ must be the same as number of rows of $B$.
    Typically $AB \neq BA$
  - Identity Matrix: $I \in \mathbb{R}^{n \times n}$ (or $I_n$); $(I)_{ij} = \delta_{ij}$ where $\delta_{ij} = 1$ if $i = j$, 0 otherwise; columns are $e_j$
  - Outer or Tensor Product: $(xy^T)_{ij} = (x \otimes y)_{ij} = x_i y_j$
    $(I + xy^T)z = z + (xy^T)z = z + x(y^T z) = z + (y \cdot z)x$

- **Sets of Vectors**: $\{x_i\}_{i=1}^m$, $x_i \in \mathbb{R}^n$
  - Linear Combination: $v = \sum_{i=1}^m \alpha_i x_i$
  - Convex Combination: same as LC but, $\alpha_i \geq 0$ and $\sum \alpha_i = 1$
  - Span: set of all linear combinations
  - Convex Hull: set of all convex combinations
  - Linearly Independent: if $\sum_1^m \alpha_i x_i = 0$ if and only if $\alpha_i = 0$ for all $i$; Linearly Dependent otherwise
    If $m > n$ then they must be linearly dependent.
  - Basis: if linearly independent they form a basis for the span, i.e. every vector in the span has a unique representation in the form $\sum_1^m \alpha_i x_i$
    If $m = n$ and linearly independent, the span is $\mathbb{R}^n$ and the vectors form a basis for $\mathbb{R}^n$
    If the vectors are non-zero, pairwise orthogonal, i.e. $x_i \cdot x_j = 0$ if $i \neq j$, then they are linearly independent and form an orthogonal basis
    If, in addition, $x_i \cdot x_i = 1$, they form an orthonormal basis

- **Norms**
  - Vector Norm: for $x \in \mathbb{R}^n$, length or norm of $x$ is denoted by $\|x\|$.
    It has the following properties:
    1. $\|x\| \geq 0$ for all $x$ and $\|x\| = 0$ if and only if $x = 0$.
    2. $\|\alpha x\| = |\alpha| \|x\|$ for all $x$ and $\alpha$.
    3. $\|x + y\| \leq \|x\| + \|y\|$ for all $x$ and $y$. 

Examples: \( \|x\|_2 = \sqrt{\sum x_i^2}, \|x\|_1 = \sum |x_i|, \|x\|_\infty = \max_i |x_i| \)

Theorem: All vector norms on \( \mathbb{R}^n \) are equivalent, i.e. there are constants \( c_1 \) and \( c_2 \) such that 
\[ c_1 \|x\| \leq \|x\|' \leq c_2 \|x\| \text{ for all } x \]

Property: For non-zero \( x, y \in \mathbb{R}^n \), the angle \( \theta \) between these vectors is defined by
\[ \cos(\theta) = \frac{x \cdot y}{\|x\|_2 \|y\|_2} \]

- **Matrix Norm**: Same notation and same properties as vector norm. May add the condition that \( \|AB\| \leq \|A\| \|B\| \)
  
  Example: \( \|A\|_F = \sqrt{\sum_{ij} a_{ij}^2} \) (Frobenius Norm)

  Induced or Subordinate Matrix Norm: based on a vector norm, we have
  \[ \|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \]

  Also has the property that \( \|Ax\| \leq \|A\|\|x\| \)

  Examples: \( \|A\|_\infty = \max_{i} \sum_{j} |a_{ij}|, \|A\|_2 = \sqrt{\text{largest eigenvalue of } A^T A} \)

- **Linear Systems**: \( A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m \), find \( x \in \mathbb{R}^n \) such that \( Ax = b \)
  
  - When \( m > n \) (overdetermined system); typically no solution; solve instead least-squares problem: Find \( x \in \mathbb{R}^n \) which minimizes \( \|Ax - b\|_2^2 \); the solution satisfies the normal equations \( A^T A x = A^T b \)
  
  - When \( m < n \) (underdetermined system); typically a family of solutions; solve instead by adding additional condition like solution must be of minimum norm, i.e. \( \|x\|_2 \) is as small as possible
  
  - When \( m = n \); typically one unique solution (when \( A \) is invertible, non-singular, \( \det A \neq 0 \), \( A \) has full rank, \( Ax = 0 \) has only \( x = 0 \) as a solution, or other conditions)
    
    Solution is \( x = A^{-1} b \), but except in the \( 2 \times 2 \) case, we never compute \( x \) this way, instead we use some computationally stable method like Gaussian Elimination to solve the system \( Ax = b \)

- **Eigenvalue and Eigenvectors**
  
  - For \( A \in \mathbb{R}^{n \times n}, \lambda \in \mathbb{C} \) and \( v \in \mathbb{C}^n, v \neq 0 \), if \( Av = \lambda v \) then \( \lambda \) is an eigenvalue of \( A \) and \( v \) is a corresponding eigenvector
    
    If \( v \) is an eigenvector for \( \lambda \), then so is \( \alpha v \) for any \( \alpha \neq 0 \)
  
  - Eigenvalues: \( A \) has \( n \) eigenvalues which are the roots of the polynomial \( \det(A - \lambda I) = 0 \)
  
  - Eigenvector: for \( \lambda \), solve \( (A - \lambda I)v = 0 \) for the eigenvector \( v \)
  
  - Symmetry: \( A \) is symmetric if \( A^T = A \); if \( A \) is symmetric then all the eigenvalues are real and the eigenvectors can be chosen in such a way that they form an orthonormal basis for \( \mathbb{R}^n \)

- **Quadratic Form**: From \( A \in \mathbb{R}^{n \times n}, A \text{ symmetric}, b \in \mathbb{R}^n \) and \( c \in \mathbb{R} \), we form the quadratic form
  \[ \phi(x) = \frac{1}{2} x^T A x - x^T b + c \]

  \[ x^T A x = \sum_{ij} a_{ij} x_i x_j \]

  \( \hat{x} = A^{-1} b \) is a minimizer of \( \phi \) if and only if all the eigenvalues of \( A \) are positive (i.e. \( A \) is positive definite)