Reading for This Lecture

- **Primary**
  - Miller and Boxer, Pages 124-128
  - Forsythe and Mohler, Sections 1 to 8
Real Vector Spaces

- A real vector space is a set $V$, along with
  - an addition operation that is commutative and associative.
  - an element $0 \in V$ such that $a + 0 = a, \forall a \in V$.
  - an additive inverse operation such that $\forall a \in V, \exists a' \in V$ such that $a + a' = 0$.
  - a scalar multiplication operation such that $\forall \lambda, \mu \in \mathbb{R}, a, b \in V$
    - $\lambda(a + b) = \lambda a + \lambda b$
    - $(\lambda + \mu)a = \lambda a + \mu a$
    - $\lambda(\mu a) = (\lambda\mu)a$
    - $1a = a$
Norms on Vector Spaces

- A norm on a vector space is a function \( \| \cdot \| : \mathcal{V} \rightarrow \mathbb{R} \) satisfying
  - \( \| v \| \geq 0 \) \( \forall v \in \mathcal{V} \)
  - \( \| v \| = 0 \) if and only if \( v = 0 \)
  - \( \| v + w \| \leq \| v \| + \| w \| \) \( \forall v, w \in \mathcal{V} \)
  - \( \| \lambda v \| = |\lambda| \cdot \| v \| \)

- Norms are used for measuring the "size" of an object or the "distance" between two objects in a vector space.

- These are the normal properties you would expect such a measure to have.
Examples of Vector Spaces

- $\mathbb{R}^n$
- $\mathbb{Z}^n$
- $\mathbb{R}^{n \times n}$
- $\{ y \in \mathbb{R}^m : Ax = y, \exists x \in \mathbb{R}^n \}$
Matrix and Vector Norms

• Unless otherwise indicated, we will use the $L_2$ norm for vectors and the corresponding norm for matrices.
• We will denote this by $\| \cdot \|$.
• Note the following definitions and properties
  - $|x^T y| \leq \| x \| \cdot \| y \|$
  - $\| A \| = \max \{ \| Ax \|/\| x \|, \ x \neq 0 \}$
  - $\| Ax \| \leq \| A \| \cdot \| x \|$
  - $\| AB \| \leq \| A \| \cdot \| B \|$
Matrix Multiplication

• The standard sequential algorithm for multiplying matrices is $O(n^3)$.

• Strassen's Algorithm is a divide and conquer approach.

• Analysis of Strassen's Algorithm
  
  - $T(n) = 7T(n/2) + dn^2$
  
  - $T(n) = O(n^\log(7)) = O(n^{2.81...})$

• Every algorithm must be $\Omega(n^2)$.

• The best known algorithm to date is $O(n^{2.376...})$.

• Can we parallelize Strassen's Algorithm?
Parallel Matrix Multiplication

- Assume a CREW shared-memory architecture with $n^3$ processors.
- Label processors as $P_{111}$ through $P_{nnn}$.
- Processor $P_{ijk}$ calculates $a_{ik} \cdot b_{kj}$.
- The remaining sums can be computed in $O(\log n)$ using a semigroup operation.
- The running time is $O(\log n)$.
- Cost optimality?
Matrix Multiplication on a Mesh

• Assume a $2n \times 2n$ mesh computer.
• Assume each processor initially stores one entry.
• Algorithm

• Analysis

• Optimality