Quadratic Forms and Matrix Norms

Practical Linear Algebra | Lecture 13

Quadratic Forms

- A quadratic form is a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(x) = x^T A x = \sum_{i,j=1}^{n} A_{ij} x_i x_j$$

for some particular $A \in \mathbb{R}^{n \times n}$

Quadratic Forms

- The choice of $A \in \mathbb{R}^{n \times n}$ is not unique
    - Makes sense – we're going from several variables and coefficients to just one scalar
- Since $x^T A x$ is a scalar:

$$\begin{aligned} x^T A x &= (x^T A x)^T \\ &= x^T A^T (x^T)^T \\ &= x^T A^T x \end{aligned}$$

Quadratic Forms

$$x^T A x = x^T A^T x$$

$$x^T A x + x^T A x = x^T A^T x + x^T A x$$

$$2x^T A x = x^T (A^T + A) x$$

$$x^T A x = x^T \left(\frac{A^T + A}{2} \right) x$$

$\dfrac{A^T + A}{2}$ is a symmetric matrix (called the symmetric part of A)

Quadratic Forms

- An infinite number of matrices have the same symmetric part
- When working with quadratic forms, we can assume A is symmetric without loss of generality
  - Why do we care? Real symmetric matrix means we can use the spectral theorem...
- Given some $a \in \mathbb{R}$:
  - $\{x \mid x^T A x \leq a\}$ is a quadratic region
  - $\{x \mid x^T A x = a\}$ is a quadratic surface
  - Ellipsoid is a special case of a quadratic region!

Quadratic Forms

- If $A = A^T \in \mathbb{R}^{n \times n}$, then via spectral theorem:

$$\begin{aligned} x^T A x &= x^T Q \Lambda Q^T x \\ &= (Q^T x)^T \Lambda (Q^T x) \\ &= \sum_{i=1}^{n} \lambda_i (q_i^T x)^2 \\ &\leq \sum_{i=1}^{n} \lambda_1 (q_i^T x)^2 \end{aligned}$$

Quadratic Forms

$$
\begin{aligned}
x^T A x &\leq \sum_{i=1}^{n} \lambda_1 (q_i^T x)^2 \\
&= (Q^T x)^T \lambda_1 I (Q^T x) \\
&= \lambda_1 (Q^T x)^T (Q^T x) \\
&= \lambda_1 x^T Q Q^T x \\
&= \lambda_1 x^T x \\
&= \lambda_1 \|x\|^2
\end{aligned}
$$

Quadratic Forms

$$x^T A x \leq \lambda_1 \|x\|^2$$

Similar logic gives: $$x^T A x \geq \lambda_n \|x\|^2$$

Together: $$\lambda_1 \|x\|^2 \geq x^T A x \geq \lambda_n \|x\|^2$$

- If the eigenvalues are sorted such that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, then we can denote λ_1 by $\lambda_{\max}$ and λ_n by $\lambda_{\min}$
- We can always rearrange the diagonal entries of Λ and the columns of Q to sort the eigenvalues

Definite Quadratic Forms

- A matrix $A = A^T \in \mathbb{R}^{n \times n}$ is called positive semidefinite (PSD) if

$$x^T A x \geq 0$$

  for all $x \in \mathbb{R}^n$
- We use the notation $A \geq 0$ or $A \succeq 0$
  - $\succeq$ is read "succeeds or is equal to"
- This means $\lambda_{\min} \geq 0$ for PSD matrices

Definite Quadratic Forms

- A matrix $A = A^T \in \mathbb{R}^{n \times n}$ is called positive definite if

$$x^T A x > 0$$

  for all $x \in \mathbb{R}^n$
- Denoted $A > 0$ or $A \succ 0$
- A is negative semidefinite if $-A \geq 0$
- A is negative definite if $-A > 0$
- A is definite if it's any of the above
- Otherwise A is indefinite

Matrix Inequalities

- We can use definite quadratic forms in matrix inequalities
  - $A \geq B$ means $A - B \geq 0$ (the matrix $A - B$ is PSD)
  - Many scalar inequalities also hold for matrices
    - $A \geq B,\ C \geq D \Rightarrow A + C \geq B + D$
    - $B \leq 0 \Rightarrow A + B \leq A$
- Matrix inequality is a partial order
  - It's possible for $A \ngeq B$ and $A \nleq B$ to both be true

Ellipsoids

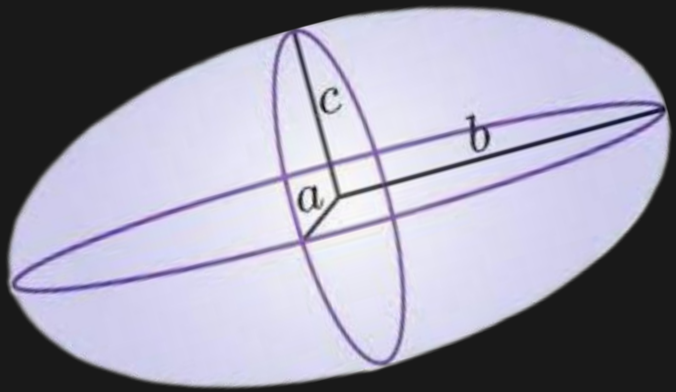
- If $A = A^T > 0$, then the set

$$\{x \mid x^T A x \leq 1\}$$

is an ellipsoid in $\mathbb{R}^n$ centered at 0

Ellipsoids

- Semiaxes are given by $\dfrac{q_i}{\sqrt{\lambda_i}}$



From https://en.universaldenker.org/formulas/731

Matrix Norms

- Many different matrix norms have been defined
- We'll discuss two:
  - Operator norm
  - Frobenius norm

Operator Norm

- The operator norm or spectral norm is defined as

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \qquad \text{for } A \in \mathbb{R}^{m \times n}$$

- This is the maximum scaling or “gain” that A applies to any input vector

Operator Norm

$$\begin{aligned}
\max_{x\neq 0} \frac{\|Ax\|^2}{\|x\|^2} &= \max_{x\neq 0} \frac{(Ax)^T(Ax)}{x^Tx} \\
&= \max_{x\neq 0} \frac{x^TA^TAx}{x^Tx} \\
&= \frac{\lambda_1(A^TA)q_1^Tq_1}{q_1^Tq_1} \\
&= \lambda_1(A^TA) = \lambda_{\max}(A^TA)
\end{aligned}$$

Operator Norm

$$\|A\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sqrt{\lambda_{\max}(A^T A)}$$

- We can use similar logic to show that the minimum gain is

$$\min_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sqrt{\lambda_{\min}(A^T A)}$$

Operator Norm

- Minimum gain is always real and nonnegative for any matrix
  - $A^T A$ is symmetric since $A^T A = (A^T A)^T$
  - $A^T A$ is PSD since $x^T A^T A x = \|Ax\|^2 \geq 0$
  - This means $\lambda_{\min} \geq 0$

Operator Norm

- Operator norm shares many properties with the ordinary vector norm
  - $\|cA\| = |c|\|A\|$ for $c \in \mathbb{R}$
  - Triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$
  - $\|A\| = \|A^T\|$
  - $\|A\| = 0 \iff A = 0$
  - Operator and vector norms are identical if the "matrix" is a vector
    - $\sqrt{\lambda_{\max}(a^T a)} = \sqrt{a^T a} = \sqrt{\|a\|^2} = \|a\|$ for $a \in \mathbb{R}^n$

Frobenius Norm

- The Frobenius norm is defined as

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^2 \right)^{\frac{1}{2}}$$

- Fact: $\|A\| \leq \|A\|_F$

Frobenius Norm

- The trace of a square matrix is the sum of the elements along its main diagonal

$$\mathbf{trace}(B) = \sum_{i=1}^{n} B_{ii} \qquad \text{for } B \in \mathbb{R}^{n \times n}$$

- We can rewrite the Frobenius norm using the trace

$$\|A\|_F = \left(\mathbf{trace}(A^T A)\right)^{\frac{1}{2}}$$

Next Time

- Singular Value Decomposition