Todd Young

Department of Mathematics

OHIO UNIVERSITY

Math Club, Athens, November 2017

# Rank

Definition: The *rank*, $r$, of an $m \times n$ matrix $A$ is the number of independent columns (or rows).

$A$ has *full rank* if $r = \min\{m, n\}$.

Rank can be determined by Gaussian Elimination:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

1 non-zero row implies $A$ has rank 1.

# Rank

Definition: The *rank*, $r$, of an $m \times n$ matrix $A$ is the number of independent columns (or rows).

$A$ has *full rank* if $r = \min\{m, n\}$.

Rank can be determined by Gaussian Elimination:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$$

1 non-zero row implies $A$ has rank 1.

But, *Rank is numerically unstable.*

$$B = \begin{pmatrix} 1 & 2 \\ 2 & 4.000001 \end{pmatrix}$$

$B$ has rank 2. A random matrix has full rank.

# Partial Differential Equations

Heat Equation:

$$\frac{\partial}{\partial t} u = c \frac{\partial^2}{\partial x^2} u$$

$u(t, x)$ - temperature as a function of time and space.

# Partial Differential Equations

Heat Equation:

$$\frac{\partial}{\partial t} u = c \frac{\partial^2}{\partial x^2} u$$

$u(t, x)$ - temperature as a function of time and space.

How is this related to RANK?

# Products of Vectors

Dot Product: $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$ $( = |\mathbf{u}||\mathbf{v}| \cos\theta )$.

# Products of Vectors

Dot Product: $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n \ ( = |\mathbf{u}||\mathbf{v}| \cos\theta \ )$.
In matrix notation:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

# Products of Vectors

Dot Product: $\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n$ $( = |\mathbf{u}||\mathbf{v}| \cos\theta )$.
In matrix notation:

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$$

The *Outer Product* is defined to be:

$$\mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^T = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix}$$

# Outer products are rank 1 matrices

Example:

$$\mathbf{u}\mathbf{v}^T = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 \\ 8 & 10 & 12 \\ 12 & 15 & 18 \end{pmatrix}$$

Each column is a multiple of $\mathbf{u}$ and each row a multiple of $\mathbf{v}$.

$\mathbf{u}\mathbf{v}^T$ has rank 1.

**Theorem:** A matrix $A$ has rank 1 iff $A = \mathbf{u}\mathbf{v}^T$ for some $\mathbf{u}$, $\mathbf{v}$.

This representation is not unique since $\mathbf{u}\mathbf{v}^T = (c\mathbf{u})(\frac{1}{c}\mathbf{v}^T)$.

$A = a\mathbf{u}\mathbf{v}^T$ with $|\mathbf{u}| = |\mathbf{v}| = 1$ is unique, where $a = \|A\|_2$.

# Heat Equation is Separable

Temperature $u(t,x)$ satisfies $\dfrac{\partial}{\partial t}u = c\dfrac{\partial^2}{\partial x^2}u$.

Assume an ansatz: $u(t,x) = T(t)X(x)$ - "separable".

$$T'(t)X(x) = T(t)X''(x) \Longrightarrow \frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)}$$

# Heat Equation is Separable

Temperature $u(t,x)$ satisfies $\frac{\partial}{\partial t}u = c\frac{\partial^2}{\partial x^2}u$.

Assume an ansatz: $u(t,x) = T(t)X(x)$ - "separable".

$$T'(t)X(x) = T(t)X''(x) \Longrightarrow \frac{T'(t)}{T(t)} = \frac{X''(x)}{X(x)} = C.$$

# Heat Equation is Separable

Temperature $u(t,x)$ satisfies $\dfrac{\partial}{\partial t}u = c\dfrac{\partial^2}{\partial x^2}u$.

Assume an ansatz: $u(t,x) = T(t)X(x)$ - "separable".

$$T'(t)X(x) = T(t)X"(x) \Longrightarrow \frac{T'(t)}{T(t)} = \frac{X"(x)}{X(x)} = C.$$

$T' = CT \Longrightarrow T(t) = T(0)e^{Ct}$

$X" = CX \Longrightarrow X(X) = A\sin(\sqrt{C}x) + B\cos(\sqrt{C}x)$

$$u(t,x) = e^{Ct}\left(A\sin(\sqrt{C}x) + B\cos(\sqrt{C}x)\right)$$

$u = T(t)X(x)$ is an "outer product".

# Discretization of $u(t,x)$

In applications, solutions of differential equations are often approximated numerically rather than solved symbolically.

Numerical solutions are discrete. $t \mapsto \{t_j\}$, $x \mapsto \{x_i\}$.

$$u(t_j, x_i) \approx U_{ij}.$$

Think of $(U_{ij})$ as a matrix $U$. If $u$ is separable,

$$u(t,x) = T(t)X(x) \Longrightarrow U = (T(t_j))(X(x_i))^T = \overline{T}\,\overline{X}^T.$$

$\{u(t,x)$ is separable$\}$ corresponds to $\{U$ is an outer product$\}$.

# Complexity and Storage

Suppose $m$ grid points in $x$ direction & $n$ grid points in $t$ direction.

$U$ is a $m \times n$ matrix $\Longrightarrow$ $mn$ entries.

If $U$ is an outer product of $\overline{T}$ and $\overline{X}$, then you only really need to store $m + n$ numbers, i.e. the entires of $\overline{T}$ and $\overline{X}$

Example: $m = 100$ and $n = 100$,

$mn = 10,000$ but $m + n = 200$.

# The Singular Value Decomposition

**SVD theorem:** Given any $m \times n$ matrix $A$,

$$A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = U \Sigma V^T$$

- $\{\mathbf{u}_i\}$, $\{\mathbf{v}_i\}$ are orthonormal sets of vectors,
- $\{\sigma_i\}$ are positive real numbers called *singular values*,
- $r \leq \min(m, n)$ is the rank of $A$.

May order: $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.

# The SVD Approximation Theorem

Principle: The larger singular values are more important.

For $k < r$ let

$$A_k \equiv \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T$$

**Theorem** (Eckart-Young-Mirsky): $A_k$ is the best rank-$k$ approximation of $A$ with respect to either the Frobenius norm or the 2-norm. Further,

$$\|A - A_k\|_2 = \sigma_{k+1}, \qquad \|A - A_k\|_F = \sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_r.$$

$$\|A - A_k\|_2 \equiv \max_{|\mathbf{x}|=1} |A\mathbf{x}|, \qquad \|A - A_k\|_F \equiv \sqrt{\sum_{i,j=1,1}^{m,n} a_{ij}^2}$$

# Application to Image Compression

An image is recorded in an array of RGB values from 0 to 255.
Example: 500 by 500 pixel image is $500 \times 500 \times 3 = 750,000$.

Suppose $R$ is the $500 \times 500$ matrix of red values and let
$R = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T = U \Sigma V^T$ be the SVD of $R$.

Let $R_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ the best rank-$k$ approximation of $R$.

Storing $R_k$ requires $k(m + n + 1)$ numbers. If $k < n/2$ then this is fewer numbers to store.
In the above example, with $k = 100$, then $3k(2n + 1) = 300,300$.

- Maybe $R_k$ is good enough?

# Back to Partial Differential Equations

$u(t,x) = T(t)X(x)$ being separable depended on the linearity of the PDE.

What if:

$$\frac{\partial}{\partial t} u = c \frac{\partial^2}{\partial x^2} u + g(u)$$

where $g(u)$ is a small nonlinearity.

Maybe $u(t,x) \approx T_1(t)X_1(x) + T_2(t)X_2(x)$?
Or at worst:

$$u(t,x) \approx T_1(t)X_1(x) + T_2(t)X_2(x) + \ldots + T_k(t)X_k(x),$$

where $k$ is small. Then storing an approximation $U_{ij}$ would require $k(m+n+1)$ numbers rather than $mn$.

Working efficiently with low rank representations is a subject of current research by Mohlenkamp and Young.

# The Curse of dimensionality

Example: Electrons in an atom

Hydrogen - 1 electron $\times$ 3 dim $=$ 3 variables, linear equation - separable.
Helium - 2 electrons $\times$ 3 dim $=$ 6 variables, nonlinear - not separable.
Carbon - 6 electrons $\times$ 3 dim $=$ 18 variables, nonlinear - not separable.

Suppose we wish to represent the carbon "wave function."
With 100 grid points in each variable then

$$100^{18} = 10^{24} \text{ Terabytes} \Longrightarrow \textbf{Intractable.}$$

With a rank 100 approximation instead:

$$100(18 \times 100) = 180,000 \text{ Numbers} \Longrightarrow \text{No Problem.}$$

# Conclusions

Thank you for your attention!

Study more Linear Algebra!