STAT 714 fa 2023

Linear algebra review 6/6

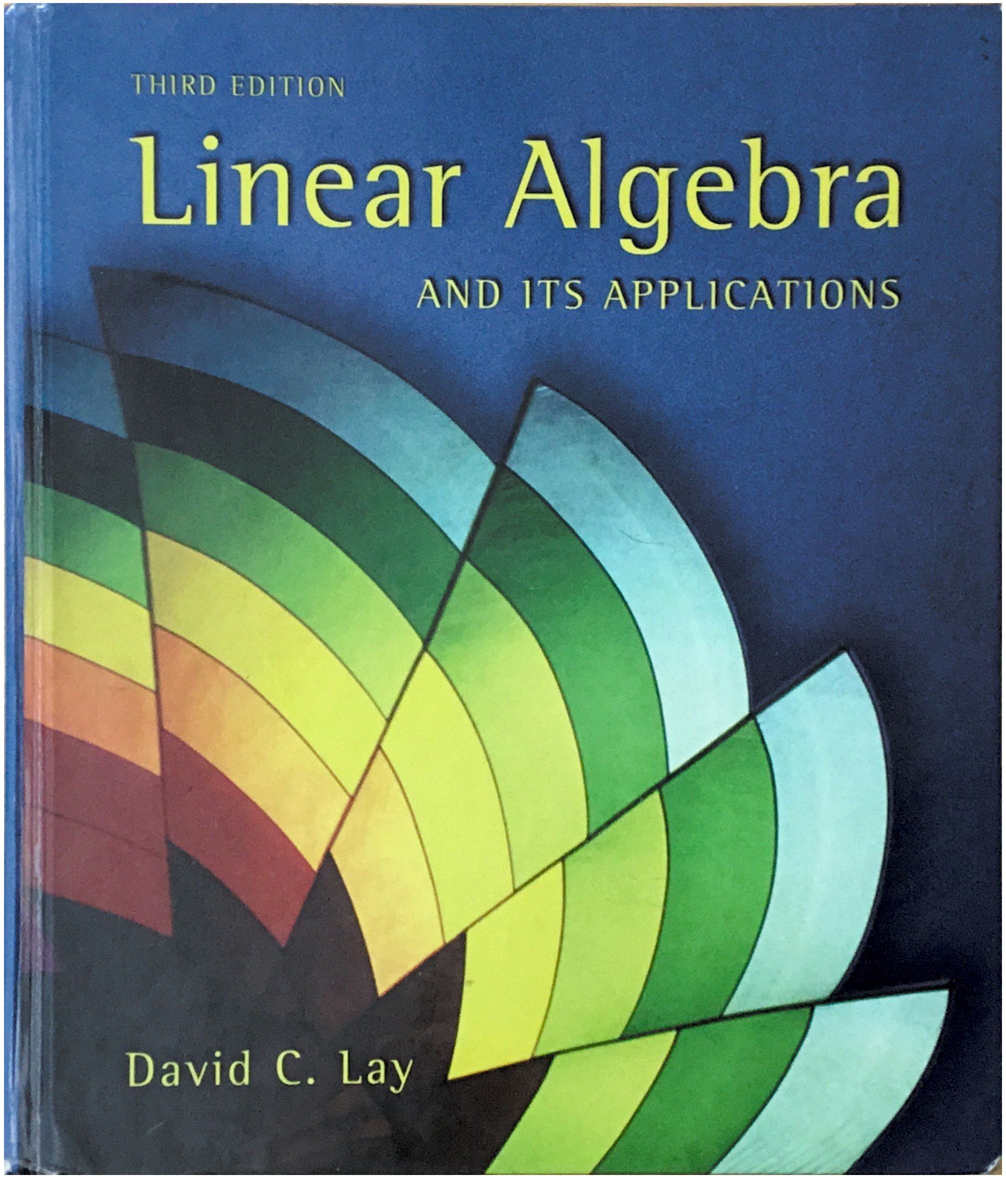
Symmetric matrices and quadratic forms

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

These notes include scanned excerpts from Lay (2003):



1. Spectral decomposition

2. Quadratic forms

3. Singular value decomposition

Some definitions for symmetric matrices

- A *symmetric* matrix $\mathbf{A}$ is a matrix such that $\mathbf{A} = \mathbf{A}^T$ (necessarily square).
- The set of eigenvalues of a symmetric matrix $\mathbf{A}$ is called the *spectrum* of $\mathbf{A}$.

If λ is an eigenvalue of A, then $\{\underset{\sim}{x} : A\underset{\sim}{x} = \lambda \underset{\sim}{x}\}$ is the eigenspace of A corresponding to λ.

Theorem (Spectral theorem for symmetric matrices)

Let $\mathbf{A}$ be an $n \times n$ symmetric matrix. Then:

1. *The eigenvalues of $\mathbf{A}$ are all real-valued.*
2. *For each eigenvalue, the dimension of the corresponding eigenspace is equal to the multiplicity of the eigenvalue as a root of the characteristic polynomial.*
3. *Eigenspaces corresponding to different eigenvalues are orthogonal.*
4. $\mathbf{A}$ *is orthogonally diagonalizable, i.e. we can write* $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, *where* $\mathbf{P}$ *is an orthogonal matrix and* $\mathbf{D}$ *is a diagonal matrix.*

Prove 3 and 4 (4 by way of Schur factorization).

Proof of (3)

Let $\lambda_1 \neq \lambda_2$ be two distinct eigenvalues of A, where $A = A^T$.

Let $\mathbf{v}_1, \mathbf{v}_2$ are corresponding eigenvectors.

$$\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 = (A\mathbf{v}_1) \cdot \mathbf{v}_2$$

$$= (A\mathbf{v}_1)^T \mathbf{v}_2$$

$$= \mathbf{v}_1^T A^T \mathbf{v}_2$$

$$= \mathbf{v}_1^T A \mathbf{v}_2$$

$$= \mathbf{v}_1^T (\lambda_2 \mathbf{v}_2)$$

$$= \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2 \implies \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0} \underbrace{\mathbf{v}_1 \cdot \mathbf{v}_2}_{} = 0$$

$$\implies \mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$$

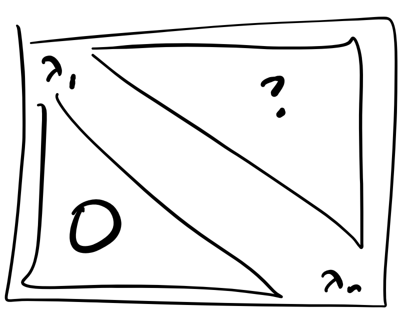
(4) **Schur Factorization:** Let A be an $n \times n$ with real eigenvalues $\lambda_1, \ldots, \lambda_n$, including multiplicities.

Then A admits the factorization

$$A = P R P^T,$$

- P has orthonormal columns
- R is upper triangular with diagonal entries $\lambda_1, \ldots, \lambda_n$.

R



If A symmetric, $A = PRP^T = (PRP^T)^T = PR^TP^T = A^T$.

$\implies R = R^T \implies R$ is diagonal.

Spectral decomposition

For a symmetric $n \times n$ matrix $\mathbf{A}$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ and corresponding unit-norm eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$, the representation $P = [\underset{\sim}{u}_1, \ldots, \underset{\sim}{u}_n]$

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^T = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T.$$

is called the *spectral decomposition* of $\mathbf{A}$.

Exercise: The matrix $\mathbf{A} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$ has unit eigenvectors

$$\mathbf{u}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

corresponding to the eigenvalues $\lambda_1 = 8$ and $\lambda_2 = 3$, respectively.

Check that $\mathbf{A} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T$.

Result (Rank of a symmetric matrix)

For a symmetric matrix, the rank is equal to the number of nonzero eigenvalues.

Discuss how the Spectral Theorem implies the above.

For an $n \times n$ symmetric matrix A:

$$\underbrace{\text{rank}\, A}_{\#\text{ nonzero eigenvalues}} + \underbrace{\dim \text{Null}\, A}_{\#\text{ eigenvalues equal to } 0} = n$$

There are n real eigenvalues, counting multiplicities (Spectral Theorem)

$\text{Null}\, A = \{ \underset{\sim}{x} : A\underset{\sim}{x} = \underset{\sim}{0} \} = \{ \underset{\sim}{x} : A\underset{\sim}{x} = 0 \cdot \underset{\sim}{x} \} =$ Eigenspace corresponding to eigenvalue $\lambda = 0$, if 0 is an eigenvalue.

$\dim \text{Null}\, A = \dim$ Eigenspace corresponding to $\lambda = 0$ = Multiplicity of $\lambda = 0$

Quadratic form

n×n, symmetric

A *quadratic form* on $\mathbb{R}^n$ is a function on $\mathbb{R}^n$ given by $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ for some symmetric matrix $\mathbf{A}$. The matrix $\mathbf{A}$ is called the *matrix of the quadratic form*.

Classifications of quadratic forms

A quadratic form $Q(\mathbf{x})$ is

1. *positive definite* if $Q(\mathbf{x}) > 0$ for all [redacted]. $\underset{\sim}{x} \neq \underset{\sim}{0}$
2. *negative definite* if $Q(\mathbf{x}) < 0$ for all [redacted]. $\underset{\sim}{x} \neq 0$
3. *positive semidefinite* if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
4. *negative semidefinite* if $Q(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$.
5. *indefinite* if $Q(\mathbf{x})$ takes both positive and negative values.

We apply the same terms to the matrices of quadratic forms: E.g. a *positive definite matrix* $\mathbf{A}$ is a symmetric matrix such that $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is positive definite.

Theorem (Principal axes theorem)

Let $\mathbf{A}$ *be an* $n \times n$ *symmetric matrix. Then there is an orthogonal change of variable* $\mathbf{x} = \mathbf{P}\mathbf{y} \iff \mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ *which transforms the quadratic form* $\mathbf{x}^T\mathbf{A}\mathbf{x}$ *into a quadratic form* $\mathbf{y}^T\mathbf{D}\mathbf{y}$ *in which* $\mathbf{D}$ *is a diagonal matrix.*

We call the columns of $\mathbf{P}$ the *principal axes* of the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$.

Prove the result.

$$A = PDP^{-1} = PDP^T, \quad P^{-1} = P^T$$

let $\underset{\sim}{y} = P^{-1}\underset{\sim}{x} = P^T \underset{\sim}{x} \iff \underset{\sim}{x} = P\underset{\sim}{y}$, I can write

$$\underset{\sim}{x}^T A \underset{\sim}{x} = (P\underset{\sim}{y})^T A P\underset{\sim}{y} = \underset{\sim}{y}^T P^T A P \underset{\sim}{y} = \underset{\sim}{y}^T P^T \underbrace{PDP^T}_{A} P \underset{\sim}{y} = \underset{\sim}{y}^T D \underset{\sim}{y}$$

Theorem (Quadratic forms and eigenvalues)

Let $\mathbf{A}$ *be a symmetric* $n \times n$ *matrix. Then the quadratic form* $\mathbf{x}^T\mathbf{A}\mathbf{x}$ *is*

1. *positive definite if and only if the eigenvalues of* $\mathbf{A}$ *are all positive.*
2. *negative definite if and only if the eigenvalues of* $\mathbf{A}$ *are all negative.*
3. *infinite if and only if* $\mathbf{A}$ *has both positive and negative eigenvalues.*

Prove the result.

p.d. $\underset{\sim}{x}^T A \underset{\sim}{x} > 0$ for all $\underset{\sim}{x} \in \mathbb{R}^n$

① let $A = PDP^T$, $P^TP = I$. Then for all $\underset{\sim}{x} \in \mathbb{R}^n$, I can write $\underset{\sim}{x}^T A \underset{\sim}{x} = \underset{\sim}{y}^T D \underset{\sim}{y}$, where $\underset{\sim}{y} = P^T \underset{\sim}{x}$.

$$\underset{\sim}{y}^T D \underset{\sim}{y} = (y_1 \cdots y_n)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 > 0$$

for all $\underset{\sim}{y} \in \mathbb{R}^n$

$\Leftrightarrow \quad \lambda_1, \dots, \lambda_n > 0.$

1. Spectral decomposition

2. Quadratic forms

3. Singular value decomposition

square

Not all matrices admit a factorization like $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$, with $\mathbf{D}$ diagonal.

But any matrix of any dimension $m \times n$ *can* be factored like $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{P}^{-1}$.

Singular value decomposition is a factorization of the latter type.

$\underset{m\times n}{A} =$

Singular values of an $m \times n$ matrix

The *singular values* of $\mathbf{A}$ are the non-negative square roots of the eigenvalues of $\mathbf{A}^T\mathbf{A}$.

Typically denote singular values by $\sigma_1 \geq \cdots \geq \sigma_n$.

Result (Setup for singular value decomposition)

Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be an orthogonal basis for $\mathbb{R}^n$ of unit eigenvectors of $\mathbf{A}^T\mathbf{A}$ arranged so that the corresponding eigenvalues of $\mathbf{A}^T\mathbf{A}$ satisfy $\lambda_1 \geq \cdots \geq \lambda_n$. Then:

1. *The singular values of $\mathbf{A}$ are the lengths of $\mathbf{A}\mathbf{v}_1, \dots, \mathbf{A}\mathbf{v}_n$.*
2. *If $\mathbf{A}$ has r nonzero singular values, then $\{\mathbf{A}\mathbf{v}_1, \dots, \mathbf{A}\mathbf{v}_r\}$ is an orthogonal basis for* Col $\mathbf{A}$ *and* rank $\mathbf{A} = r$.

Prove the results.

① $\|A\mathbf{v}_1\|^2 = (A\mathbf{v}_1)\cdot(A\mathbf{v}_1) = (A\mathbf{v}_1)^T A\mathbf{v}_1 = \mathbf{v}_1^T A^T A\mathbf{v}_1 = \lambda_1 \underbrace{\mathbf{v}_1^T\mathbf{v}_1} = \lambda_1$

$\|A\mathbf{v}_1\| = \sqrt{\lambda_1} = \sigma_1$

② Let A have r nonzero singular values.

(i) Show that $\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\}$ is orthogonal

(ii) Show that $\mathrm{Col}\, A = \mathrm{Span}\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\}$

Show (i): For $i \neq j$,

$$(A\mathbf{v}_j)\cdot(A\mathbf{v}_i) = \mathbf{v}_j^T A^T A\mathbf{v}_i = \lambda_i \mathbf{v}_j^T\mathbf{v}_i = 0.$$

Show (ii): $\mathrm{Span}\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\} \subset \mathrm{Col}\, A$ is immediate.

To show $\mathrm{Col}\, A \subset \mathrm{Span}\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\}$

Let $\mathbf{x} \in \mathrm{Col}\, A$. Then $\mathbf{x} = \underset{m\times n}{A}\mathbf{z}$ for some $\mathbf{z} \in \mathbb{R}^n$.

Since $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is a basis for $\mathbb{R}^n$, I can write

$$\mathbf{z} = c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n \quad \text{for some } c_1, \dots, c_n.$$

Then

$\|A\mathbf{v}_{r+1}\| = 0$

$\|A\mathbf{v}_n\| = 0$

$$\begin{aligned}\mathbf{x} = A\mathbf{z} &= A(c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n) \\ &= c_1 A\mathbf{v}_1 + \dots + c_r A\mathbf{v}_r + \boxed{c_{r+1}\underline{A\mathbf{v}_{r+1}} + \dots + c_n \underline{A\mathbf{v}_n}} \\ &\in \mathrm{Span}\{A\mathbf{v}_1, \dots, A\mathbf{v}_r\}.\end{aligned}$$

$= \mathbf{0}$

Theorem (Singular value decomposition)

For $\underset{m\times n}{\mathbf{A}}$ *with rank* r *there exist orthogonal matrices* $\underset{m\times m}{\mathbf{U}}$ *and* $\underset{n\times n}{\mathbf{V}}$ *such that*

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T, \qquad \textit{where} \qquad \underset{m\times n}{\boldsymbol{\Sigma}} = \begin{bmatrix} \underset{r\times r}{\mathbf{D}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\underset{r\times r}{\mathbf{D}}$ *is diag. with diag. entries* $\sigma_1 \geq \cdots \geq \sigma_r$ *the nonzero singular vals of* $\mathbf{A}$.

Note: The matrices $\mathbf{U}$ and $\mathbf{V}$ are not uniquely determined by $\mathbf{A}$, but $\boldsymbol{\Sigma}$ is.

The representation of $\mathbf{A}$ as $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ is the *singular value decomposition* of $\mathbf{A}$.

Columns in $\mathbf{U}$ and $\mathbf{V}$ are, respectively, the *left* and *right singular vectors* of $\mathbf{A}$.

$$\underset{m\times n}{\mathbf{A}} = \underset{m\times m}{\mathbf{U}}\underset{m\times n}{\mathbf{\Sigma}}\underset{n\times n}{\mathbf{V}^T}, \quad \text{where} \quad \underset{m\times n}{\mathbf{\Sigma}} = \begin{bmatrix} \underset{r\times r}{\mathbf{D}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

$$\Longleftrightarrow \quad A = U\Sigma V^T$$
$$AV = U\Sigma$$

$$U = \left[\tfrac{1}{\sigma_1} A\underset{\sim}{v}_1 \; \cdots \; \tfrac{1}{\sigma_r} A\underset{\sim}{v}_r \quad \underset{\sim}{u}_{r+1} \cdots \underset{\sim}{u}_m \right] \qquad D = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}$$

$$V = \left[\underset{\sim}{v}_1 \; \cdots \; \underset{\sim}{v}_n \right]$$

$$\begin{aligned} AV &= A\left[\underset{\sim}{v}_1 \cdots \underset{\sim}{v}_n\right] \\ &= \left[A\underset{\sim}{v}_1 \cdots A\underset{\sim}{v}_n\right] \\ &= \left[A\underset{\sim}{v}_1 \cdots A\underset{\sim}{v}_r \; A\underset{\sim}{v}_{r+1} \cdots A\underset{\sim}{v}_n\right] \\ &= \left[A\underset{\sim}{v}_1 \cdots A\underset{\sim}{v}_r \; \underset{\sim}{0} \cdots \underset{\sim}{0}\right] \end{aligned}$$

$$U\Sigma = \left[\tfrac{1}{\sigma_1} A\underset{\sim}{v}_1 \; \cdots \; \tfrac{1}{\sigma_r} A\underset{\sim}{v}_r \quad \underset{\sim}{u}_{r+1} \cdots \underset{\sim}{u}_m \right] \underset{m\times n}{\begin{bmatrix} \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix} & 0 \\ 0 & 0 \end{bmatrix}}$$

$$= \left[A\underset{\sim}{v}_1 \cdots A\underset{\sim}{v}_r \; \underset{\sim}{0} \cdots \underset{\sim}{0}\right]$$

Exercise: Find a SVD for $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

3×2

① Get $A^T A = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$

② Get eigenvalues of $A^T A$.

$$\begin{vmatrix} 9-\lambda & -9 \\ -9 & 9-\lambda \end{vmatrix} = 0 \iff (9-\lambda)^2 - 89 = 0$$

$$\iff \lambda_2 = 0, \quad \lambda_1 = 18$$

(3) Find orthogonal unit eigenvectors.

You find $\underset{\sim}{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ $\qquad \underset{\sim}{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

(4) Get singular values $\sigma_1 = \sqrt{18}$ $\qquad \sigma_2 = 0$.

$$\underset{2\times 2}{V} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \qquad \underset{3\times 3}{U} = \begin{bmatrix} \frac{1}{\sigma_1} A \underset{\sim}{v}_1 & \underset{\sim}{u}_2 & \underset{\sim}{u}_3 \end{bmatrix}$$

$$A\underset{\sim}{v}_1 = \frac{1}{\sqrt{18}} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -1/3 \\ 2/3 \\ -2/3 \end{bmatrix}$$

$$\underset{3\times 2}{\Sigma} = \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$A = U \Sigma V^T$$

$$\begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} -1/3 & & \\ 2/3 & \underset{\sim}{u}_2 & \underset{\sim}{u}_3 \\ -2/3 & & \end{bmatrix} \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T$$

$$= \begin{bmatrix} -1/3 \\ 2/3 \\ -2/3 \end{bmatrix} \sqrt{18} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \leftarrow \text{Reduced SVD}$$

The Invertible Matrix Theorem (concluded)

Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

u. $(\text{Col } A)^{\perp} = \{\mathbf{0}\}$.

v. $(\text{Nul } A)^{\perp} = \mathbb{R}^n$.

w. $\text{Row } A = \mathbb{R}^n$.

x. A has n nonzero singular values.

Recipe for reduced singular value decomposition

For an $m \times n$ matrix $\mathbf{A}$ with rank r:

1. Obtain orthonormal set $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of eigenvectors of $\mathbf{A}^T\mathbf{A}$ with corresponding eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n$, of which r are nonzero.
2. Obtain nonzero singular values $\sigma_1 = \sqrt{\lambda_1}, \dots, \sigma_r = \sqrt{\lambda_r}$.
3. Set $\mathbf{U}_r = [\mathbf{A}\mathbf{v}_1/\sigma_1 \ \cdots \ \mathbf{A}\mathbf{v}_r/\sigma_r]$ and $\mathbf{V}_r = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_r]$.
4. Set $\mathbf{D} = \operatorname{diag}(\sigma_1, \dots, \sigma_r)$.

Then the *reduced singular value decomposition* of $\mathbf{A}$ is $\mathbf{A} = \mathbf{U}_r\mathbf{D}\mathbf{V}_r^T$.

We can also construct a *low-rank approximation* to $\mathbf{A}$. For $1 \leq s < r$, set

$$\mathbf{A}_s = \mathbf{U}_s\mathbf{D}_s\mathbf{V}_s^T,$$

where $\mathbf{U}_s = [\mathbf{u}_1 \cdots \mathbf{u}_s]$, $\mathbf{D}_s = \operatorname{diag}(\sigma_1, \dots, \sigma_s)$, $\mathbf{V}_s = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_s]$.

Useful in image compression!

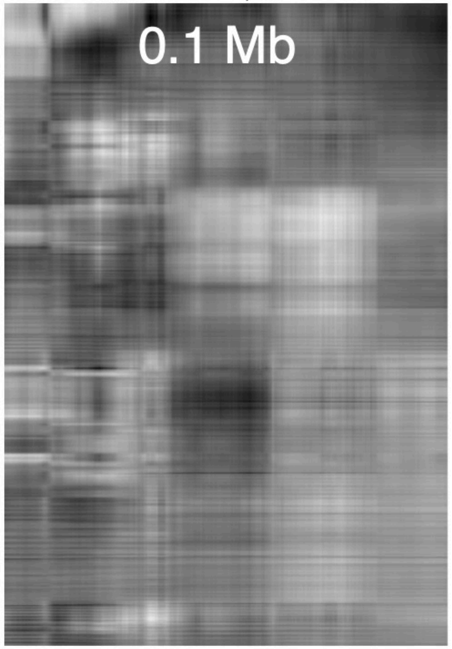
```
# read in color png
library(png)
im_col <- readPNG(source = "your_image.png")

# convert to grayscale with formula from
# https://www.had2know.org/technology/rgb-to-gray-scale-converter.html
im <- 0.299*im_col[,,1] + 0.587*im_col[,,2] + 0.114 * im_col[,,3]
m <- dim(im)[1]
n <- dim(im)[2]

# obtain singular value decomposition
im_svd <- svd(im)
s <- 10 # choose number of singular vectors to keep
Us <- im_svd$u[,1:s]
Ds <- diag(im_svd$d[1:s])
Vs <- im_svd$v[,1:s]

# construct low-rank approximation
im_approx <- Us %*% Ds %*% t(Vs)

# keep grayscale values in [0,1]
im_approx[im_approx > 1] <- 1
im_approx[im_approx < 0] <- 0

# display image
asp <- m/n
plot(NA,ylim = c(0,asp),xlim = c(0,1),type = "n",
     xaxt = "n",yaxt = "n",bty = "n",xlab = NA,ylab = NA)
rasterImage(im_approx,0,0,asp,asp,interpolate=FALSE)
```

$$\underset{n\times m}{A} = U_s D_s V_s^T$$

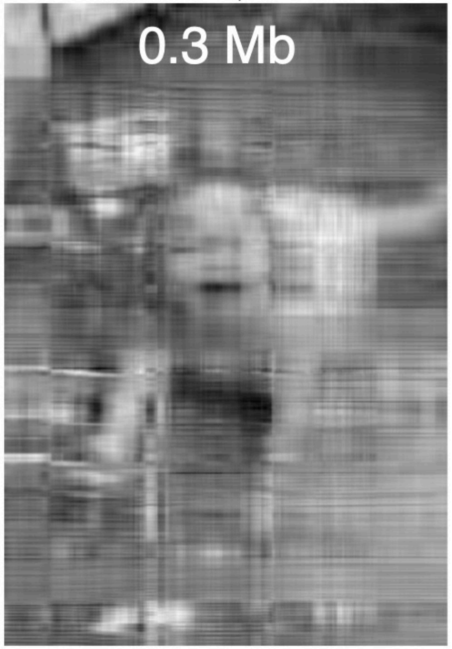
$m \times n$

Original 2091 x 1667 image

Reduced SVD, rank = 5



Reduced SVD, rank = 10



Reduced SVD, rank = 20

Reduced SVD, rank = 40

Reduced SVD, rank = 80

Lay, D. C. (2003). *Linear algebra and its applications. Third edition.* Pearson Education.