An Optimal, Democratic Diagonalization Technique from the Singular Value Decomposition

Scott Beaver - Western Oregon University

Humboldt State University Mathematics Department Colloquium

March 26, 2009

- We will discuss *orthogonalization* of a real full-rank matrix
- This can be accomplished in a variety of ways
- In introductory linear algebra courses it is typically done by way of the *Gram-Schmidt* procedure
- We'll focus on a very different orthogonalization method using the singular value decomposition (hereafter SVD)

Throughout, m and n will always denote natural numbers with $m \geq n$

Unless otherwise indicated, all vectors herein are interpreted as column vectors

$$u \in \mathbb{R}^n \qquad \Longrightarrow \qquad u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

Definition

Let $u, v \in \mathbb{R}^n$. Then the *inner product* of u and v, written $\langle u, v\rangle$ is defined as

$$\langle u, v\rangle \quad := \quad \sum_{j=1}^{n} u_j v_j$$

Definition

Let $u \in \mathbb{R}^n$. The (Euclidean) *norm* of u is defined as

$$\|u\|_2 := \left(\sum_{j=1}^{n} u_j^2\right)^{1/2} = \sqrt{\langle u, u\rangle}$$

A vector $u \in \mathbb{R}^n$ is a *unit vector* or *normal* if

$$\|u\|_2 \;=\; 1$$

Definition

Two vectors $u, v \in \mathbb{R}^n$ are *orthogonal* if

$$\langle u, v \rangle = 0$$

If u, v are each orthogonal and normal, then we say u and v are *orthonormal*

To avoid subscript ambiguity when considering a finite collection of vectors, denote by u_{i*} the i^{th} member of the collection

Definition

A collection of nonzero vectors $\{u_{1*}, \ldots, u_{n*}\}$ is *orthonormal* if $\langle u_{i*}, u_{j*} \rangle = 0$ when $i \neq j$ and if each u_{i*} is of unit length

- The orthogonality of the u_{i*}'s imply that the u_{i*}'s are *linearly independent* - none of them can be written as a linear combination of any of the others
- Further, any nonzero vector in $\mathbb{R}^n$ can be assembled as a linear combination (with coefficients not all zero) of the u_{i*}'s

These two facts, along with normality of the u_{*i}'s, mean that the u_{i*}'s form an *orthonormal basis* for $\mathbb{R}^n$

Though vectors are our building blocks, the primary focus of our discussion is the idea of *matrix* which is an extension of the idea of ordered n-tuples of real numbers (vectors), but with two directions of order, loosely speaking

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \\ \vdots & & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix}$$

Matrices are essentially the concatenation of vectors of identical length

Definition

Let $A = (a_{ij}) \in \mathbb{R}^{m \times n}$. The *transpose* A^T of A is the matrix $(a_{ji}) \in \mathbb{R}^{n \times m}$.

Example

$$\begin{pmatrix} 1 & 0 & 3 \\ 2 & -1 & -4 \end{pmatrix}^T = \begin{pmatrix} 1 & 2 \\ 0 & -1 \\ 3 & -4 \end{pmatrix}$$

Definition

(Matrix Multiplication) Let $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$. Then the product AB is defined element-wise as

$$(AB)_{ij} := \sum_{k=1}^{n} a_{ik} b_{kj}$$

and the matrix $AB \in \mathbb{R}^{m \times p}$.

The n-dimensional *identity matrix* is

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & & 1 \end{pmatrix}$$

We'll write I for the identity matrix when the size is clear from the context.

When $m \geq n$, it is possible that the columns of A, denoted A_i, $i = 1, \ldots, n$, can form a linearly independent set

Any linearly independent set of vectors can be transformed into an orthonormal set

This process is called orthogonalization

Definition

A square matrix Q is *orthogonal* if $Q^T Q = I$.

Definition

Let $A \in \mathbb{R}^{m\times n}$. Then the (full) SVD of A is $A = U\Sigma V^T =$

$$\left(\begin{array}{c|c|c|c} & & & \\ U_1 & U_2 & \cdots & U_m \\ & & & \end{array}\right)\begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \sigma_n \\ 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 \end{pmatrix}\left(\begin{array}{c} V_1^T \\ \hline V_2^T \\ \hline \vdots \\ \hline V_n^T \end{array}\right).$$

where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices and Σ is diagonal. The σ_i's are the *singular values* of A, by convention arranged in nonincreasing order

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$$

- The columns of U are termed *left singular vectors* of A; the columns of V are called *right singular vectors* of A
- Since U and V are orthogonal matrices, the columns of each form orthonormal bases for $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively
- We can use these bases to illuminate the fundamental property of the SVD

For the equation $Ax = b$, the SVD makes every matrix diagonal by selecting the right bases for the range and domain

Let $b, x \in \mathbb{R}^n$ such that $Ax = b$, and expand b in the columns of U and x in the columns of V to get

$$b' := U^T b, \qquad x' := V^T x$$

so that

$$b = Ax \qquad \Longleftrightarrow \qquad \begin{aligned} U^T b &= U^T A x \\ &= U^T (U \Sigma V^T) x \\ &= \Sigma x' \end{aligned}$$

or

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} \qquad \Longleftrightarrow \qquad \boldsymbol{b}' = \boldsymbol{\Sigma}\boldsymbol{x}'$$

Let $y \in \mathbb{R}^n$, then the action of left multiplication of y by A (computing $z := Ay$) is decomposed by the SVD into three steps

$$\begin{aligned} z &= Ay \\ &= (U\Sigma V^T)\,y \quad = U\Sigma(V^T y) \\ &= U\Sigma c \qquad\qquad (c := V^T y) \\ &= Uw \qquad\qquad (w := \Sigma c) \end{aligned}$$

Let $y \in \mathbb{R}^n$, then the action of left multiplication of y by A (computing $z := Ay$) is decomposed by the SVD into three steps

$$
\begin{aligned}
z &= Ay \\
&= (U\Sigma V^T)\,y \quad = U\Sigma(V^T y) \\
&= U\Sigma\, c \qquad\quad (c := V^T y) \\
&= U w \qquad\quad\;\; (w := \Sigma\, c)
\end{aligned}
$$

- $c = V^T y$ is the *analysis* step, in which the components of y, *in the basis of* $\mathbb{R}^n$ *given by the columns of* V, are computed

Let $y \in \mathbb{R}^n$, then the action of left multiplication of y by A (computing $z := Ay$) is decomposed by the SVD into three steps

$$\begin{aligned} z &= Ay \\ &= (U\Sigma V^T)\,y \quad = U\Sigma(V^T y) \\ &= U\Sigma c \qquad\quad (c := V^T y) \\ &= Uw \qquad\quad\; (w := \Sigma c) \end{aligned}$$

- $c = V^T y$ is the *analysis* step, in which the components of y, *in the basis of* $\mathbb{R}^n$ *given by the columns of* V, are computed
- $w = \Sigma c$ is the *scaling* step in which the components c_i, $i \in \{1, 2, \ldots, n\}$ are dilated

Let $y \in \mathbb{R}^n$, then the action of left multiplication of y by A (computing $z := Ay$) is decomposed by the SVD into three steps

$$
\begin{aligned}
z &= Ay \\
&= (U\Sigma V^T)\,y \quad = U\Sigma(V^T y) \\
&= U\Sigma c \qquad\quad (c := V^T y) \\
&= Uw \qquad\quad\ \ (w := \Sigma c)
\end{aligned}
$$

- $c = V^T y$ is the *analysis* step, in which the components of y, *in the basis of* $\mathbb{R}^n$ *given by the columns of* V, are computed
- $w = \Sigma c$ is the *scaling* step in which the components $c_i,\ i \in \{1, 2, \ldots, n\}$ are dilated
- $z = Uw$ is the *synthesis* step, in which z is assembled by scaling each of the $\mathbb{R}^n$-basis vectors U_i by w_i and summing

So how do we compute the matrices U, Σ, and V in the SVD of some square $A \in \mathbb{R}^{n \times n}$?

We have that $\quad V^T V = I = U^T U, \quad A = U \Sigma V^T \quad$ yields

$$\begin{aligned} AV &= U\Sigma & (1) \\ U^T A &= \Sigma V^T \\ A^T U &= V\Sigma & (2) \end{aligned}$$

Or, for each $j \in \{1, 2, \ldots, n\}$,

$$Av_j = \sigma_j u_j \tag{3}$$

Now we multiply Equation (3) by A^T to get

$$\begin{aligned} A^T A v_j &= A^T \sigma_j u_j \\ &= \sigma_j A^T u_j \\ &= \sigma_j^2 v_j \qquad \text{By (2)} \end{aligned}$$

So the v_j's are the eigenvectors of $A^T A$ with corresponding eigenvalues σ_j^2

If we denote the i^{th} row of A by ${}_iA$ and the j^{th} column of B by B_j we have

$$(AB)_{ij} \quad := \quad \langle ({}_iA)^T, B_j \rangle \quad = \quad {}_iAB_j$$

Note that $\qquad (A^TA)_{ij} \quad = \quad {}_iAA_j \qquad$ or, more succinctly

$$A^TA = \begin{pmatrix} {}_1A_1 & {}_1A_2 & \cdots & {}_1A_n \\ {}_2A_1 & {}_2A_2 & & \vdots \\ \vdots & & \ddots & \\ {}_nA_1 & \cdots & & {}_nA_n \end{pmatrix} \tag{4}$$

A^TA is a matrix of inner products of columns of A - often called the *Gram matrix of A*

We'll see the Gram matrix later when considering an application

Let's do an example

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \implies A^T = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

$$\implies A^T A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

To find the eigenvectors v and the corresponding eigenvalues λ for $A^T A$, we solve

$$A^T A v = \lambda v \qquad \Longleftrightarrow \qquad (A^T A - \lambda I) v = 0$$

for λ and v. The standard technique for finding such λ and v is to first seek the λ that make singular the matrix

$$A^T A - \lambda I = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} = \begin{pmatrix} 3-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix}$$

This is typically accomplished by solving $\det\left(A^T A - \lambda I\right) = 0$:

$$\begin{aligned}
\begin{vmatrix} 3-\lambda & 1 & 0 \\ 1 & 1-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} &= (3-\lambda)(1-\lambda)(2-\lambda) - 2 + \lambda \\
&= -\lambda^3 + 6\lambda^2 - 10\lambda + 4 \quad = \quad 0
\end{aligned}$$

which is solved by

$$\begin{aligned}
\lambda_1 &= \sigma_1^2 = 2 + \sqrt{2} \\
\lambda_2 &= \sigma_2^2 = 2 \\
\lambda_3 &= \sigma_3^2 = 2 - \sqrt{2}
\end{aligned}$$

Now (as a gentle first step) we find a vector v_2 so that $A^T A v_2 = 2 v_2$; we do this by finding a basis for the nullspace of

$$A^T A - 2I = \begin{pmatrix} 3-2 & 1 & 0 \\ 1 & 1-2 & 0 \\ 0 & 0 & 2-2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Certainly any vector of the form $\begin{pmatrix} 0 \\ 0 \\ t \end{pmatrix}$, $t \in \mathbb{R} \setminus \{0\}$, is mapped to zero by $A^T A - 2I$, so we are free to set $v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

To find v_1 we find a basis for the nullspace of

$$A^T A - (2+\sqrt{2})I = \begin{pmatrix} 1-\sqrt{2} & 1 & 0 \\ 1 & -1-\sqrt{2} & 0 \\ 0 & 0 & -\sqrt{2} \end{pmatrix}$$

$$\text{which row-reduces to } \begin{pmatrix} 1-\sqrt{2} & 1 & 0 \\ 0 & 0 & -\sqrt{2} \\ 0 & 0 & 0 \end{pmatrix}$$

So any vector of the form $\begin{pmatrix} s \\ (-1+\sqrt{2})s \\ 0 \end{pmatrix}$, $s \in \mathbb{R} \setminus \{0\}$, is mapped to the zero vector by $A^T A - (2+\sqrt{2})I$

Thus $v_1' := \begin{pmatrix} 1 \\ -1+\sqrt{2} \\ 0 \end{pmatrix}$ spans the nullspace of $A^T A - \lambda_1 I$, but $\| v_1' \| \neq 1$

So we set $v_1 = \frac{v_1'}{\| v_1' \|} = \frac{1}{\sqrt{4-2\sqrt{2}}} \begin{pmatrix} 1 \\ -1+\sqrt{2} \\ 0 \end{pmatrix}$

We *could* find v_3 in a similar manner, but in this particular case there's a quicker way...

$$v_3 = \begin{pmatrix} -(v_1)_2 \\ (v_1)_1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{4-2\sqrt{2}}} \begin{pmatrix} 1-\sqrt{2} \\ 1 \\ 0 \end{pmatrix}$$

Certainly $v_3 \perp v_2$ and by construction $v_3 \perp v_1$ - recall the theorem from linear algebra symmetric matrices must have orthogonal eigenvectors

We *could* find v_3 in a similar manner, but in this particular case there's a quicker way...

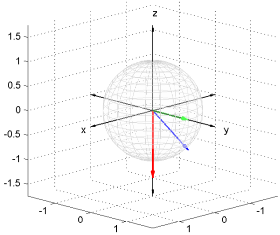
$$v_3 = \begin{pmatrix} -(v_1)_2 \\ (v_1)_1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{4-2\sqrt{2}}} \begin{pmatrix} 1-\sqrt{2} \\ 1 \\ 0 \end{pmatrix}$$

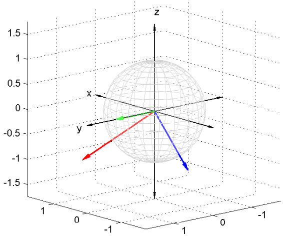
Certainly $v_3 \perp v_2$ and by construction $v_3 \perp v_1$ - recall the theorem from linear algebra symmetric matrices must have orthogonal eigenvectors

- Finally, the easy part: $A = U\Sigma V^T \quad \Longrightarrow \quad U = AV\Sigma^{-1}$

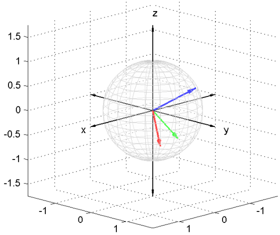
$$A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \sqrt{2+\sqrt{2}} & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{2-\sqrt{2}} \end{pmatrix}$$

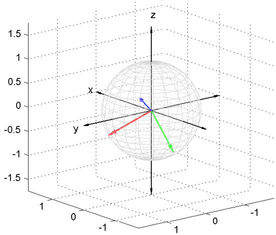
$$U = \begin{pmatrix} \frac{1}{2} & -\frac{1}{\sqrt{2}} & -\frac{1}{2} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & -\frac{1}{\sqrt{2}} & \frac{1}{2} \end{pmatrix} \qquad V^T = \begin{pmatrix} \frac{1}{\sqrt{4-2\sqrt{2}}} & \frac{-1+\sqrt{2}}{\sqrt{4-2\sqrt{2}}} & 0 \\ 0 & 0 & 1 \\ \frac{1-\sqrt{2}}{\sqrt{4-2\sqrt{2}}} & \frac{1}{\sqrt{4-2\sqrt{2}}} & 0 \end{pmatrix}$$



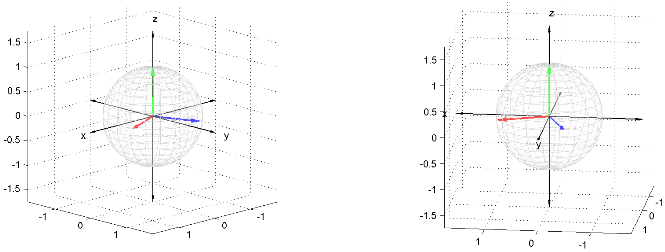


The columns of A in the unit sphere

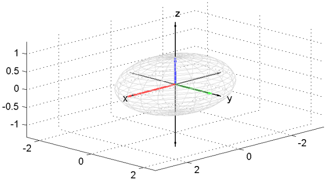


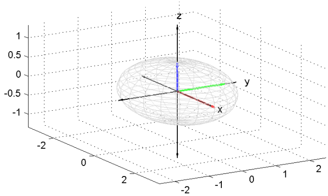


The columns of U in the unit sphere



The columns of V in the unit sphere





The columns of Σ in its ellipsoid

- For nonsingular A, the matrix $L := UV^T$ is called the *symmetric* or *Löwdin orthogonalization* of the matrix A
- L is unique since any sequence of sign choices for the columns of V determines a sequence of signs for the columns of U
- Like Gram-Schmidt orthogonalization, it takes as input a linearly independent set (the columns of A) and outputs an orthonormal set

- (Classical) Gram-Schmidt is unstable due to repeated subtractions; Modifed Gram-Schmidt remedies this
- An additional deficiency of Gram-Schmidt is that it requires a *choice* of an initial starting vector
- With some effort we can make an optimal choice
- But occasionally we want to disturb the original set of vectors as little as possible

Definition

The *Frobenius inner product* $A : B$ of $A, B \in \mathbb{R}^{m \times n}$ is

$$A : B := \sum_{i,j=1}^{n} A_{ij} B_{ij} = \operatorname{tr}(A^T B). \tag{5}$$

The norm induced by the Frobenius inner product is the *Frobenius norm*:

$$\|A\|_F := \sqrt{A : A} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}^2 \right)^{1/2} \tag{6}$$

which is just the Euclidean norm on the space $\mathbb{R}^{m \times n}$ (viewed as $\mathbb{R}^{mn}$).

- The Frobenius inner product satisfies the Cauchy-Schwarz inequality
$$|\mathrm{tr}(A^T B)| \leq \|A\|_F \|B\|_F. \tag{7}$$
Equality holds as expected, when $A = cB$, $c \in \mathbb{R}$.

- This equation permits us to associate an Euclidean geometry with $\mathbb{R}^{m\times n}$, augmenting its algebraic structures and the Euclidean geometry of the column space.

- In particular we can find "angles" between matrices by the familiar-looking
$$\cos\theta = \frac{\mathrm{tr}(A^T B)}{\|A\|_F \|B\|_F} \tag{8}$$
where we choose $\theta \in (-\pi, \pi]$

- We wish to find, for a given $A \in \mathbb{R}^{n \times n}$, the minimally-distant orthogonal $Q \in \mathbb{R}^{n \times n}$, measured by $\|A - Q\|_F$
- A nice way of simultaneously seeing what the solution must be and proving that it is correct is by minimizing the magnitude of the angle θ between A and Q over all orthogonal Q. Let the SVD of A be $U \Sigma V^T$. We have

$$\begin{aligned} \cos\theta &= \frac{A : Q}{\|A\|_F \|Q\|_F} = \frac{\operatorname{tr}(A^T Q)}{n\|A\|_F} \\ &= \frac{\operatorname{tr}(V \Sigma U^T Q)}{n\|A\|_F} = \frac{\operatorname{tr}(\Sigma U^T Q V)}{n\|A\|_F} \\ &= \frac{\sum_{i=1}^{n} \sigma_i (U^T Q V)_{ii}}{n\|A\|_F}. \end{aligned} \tag{9}$$

This expression leads to our main theorem

Theorem

Over all orthogonal matrices $Q \in \mathbb{R}^{n \times n}$, $\|A - Q\|_F$ *is minimized when* $Q = L$.

The proof is sufficiently elementary as to be included in a semester-long linear algebra course

Proof.

- Let Q be orthogonal, and denote $Y := U^T QV$, which is orthogonal with diagonal elements Y_{ii}, $i = 1, \ldots, n$.

Proof.

- Let Q be orthogonal, and denote $Y := U^T QV$, which is orthogonal with diagonal elements Y_{ii}, $i = 1, \ldots, n$.
- Since the denominator of $\frac{\sum_{i=1}^n \sigma_i (U^T QV)_{ii}}{n\|A\|_F}$ contains only positive constants, the task is to maximize $\sum_{i=1}^n \sigma_i Y_{ii}$.

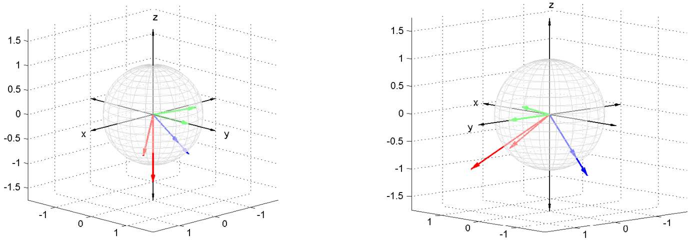
Proof.

- Let Q be orthogonal, and denote $Y := U^T QV$, which is orthogonal with diagonal elements Y_{ii}, $i = 1, \ldots, n$.
- Since the denominator of $\dfrac{\sum_{i=1}^{n} \sigma_i (U^T QV)_{ii}}{n\|A\|_F}$ contains only positive constants, the task is to maximize $\sum_{i=1}^{n} \sigma_i Y_{ii}$.
- By construction, the σ_i's are all positive, and the orthogonality of Y demands that $|Y_{ii}| \leq 1$, so the sum is maximized when all of the Y_{ii}'s equal one, which occurs if and only if $Y = U^T QV = I_n$.

Proof.

- Let Q be orthogonal, and denote $Y := U^T QV$, which is orthogonal with diagonal elements Y_{ii}, $i = 1, \dots, n$.
- Since the denominator of $\frac{\sum_{i=1}^{n} \sigma_i (U^T QV)_{ii}}{n\|A\|_F}$ contains only positive constants, the task is to maximize $\sum_{i=1}^{n} \sigma_i Y_{ii}$.
- By construction, the σ_i's are all positive, and the orthogonality of Y demands that $|Y_{ii}| \leq 1$, so the sum is maximized when all of the Y_{ii}'s equal one, which occurs if and only if $Y = U^T QV = I_n$.
- Since U and V are orthogonal, the conclusion $Q = L = UV^T$ is immediate.

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \qquad L = \begin{pmatrix} \dfrac{\frac{\sqrt{2}}{2}}{\sqrt{4-2\sqrt{2}}} & \dfrac{-1+\frac{\sqrt{2}}{2}}{\sqrt{4-2\sqrt{2}}} & -\dfrac{1}{\sqrt{2}} \\ \dfrac{\frac{2}{\sqrt{2}}-1}{\sqrt{4-2\sqrt{2}}} & \dfrac{1}{\sqrt{4-2\sqrt{2}}} & 0 \\ \dfrac{-\frac{\sqrt{2}}{2}}{\sqrt{4-2\sqrt{2}}} & \dfrac{1-\frac{\sqrt{2}}{2}}{\sqrt{4-2\sqrt{2}}} & -\dfrac{1}{\sqrt{2}} \end{pmatrix}$$



The columns of A and of L with the unit sphere

An alternative interpretation of L

- It can be shown that if A is nonsingular, then $L = (A^T A)^{-1/2} A$
- This is a generalization of the algorithm for producing a unit vector from an arbitrary nonzero vector in $\mathbb{R}^n$
- Useful for an application to wireless communications

- Orthogonal-Frequency-Division-Multiplexing (OFDM) is a candidate for the 4G wireless standard
- The carrier waves are orthogonal in the sense of an inner product defined by continuous-time integration
- But the carriers are square-waves, and require a “guard-interval” to preclude overlap due to multipath propagation
- Square-waves also have quite poor frequency-domain properties, so are susceptible to adjacent carrier interference due to the Doppler effect if present

- Gaussian waves are optimally robust regarding frequency interference but are not orthogonal
- Time-shifted and frequency-shifted Gaussians can be represented on a lattice in the time-frequency plane
- The Gram matrix (denoted R) of inner products of modulated waves is a strictly positive definite matrix
- This permits one to calculate $R^{-1/2}$ (via Taylor series or Newton's Algorithm, for example)

Let g be some function, and $T, F \in \mathbb{R}^+$. Denote $g_{i,j} := g(t - iT)e^{2\pi ijF}$.

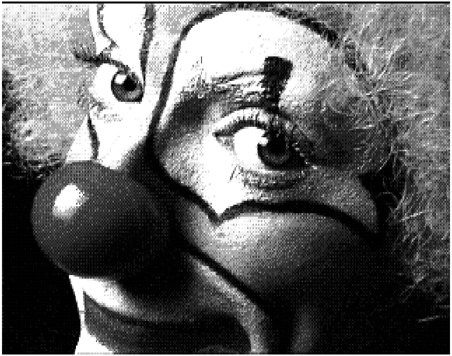
Theorem

If $g_{i,j}$ generates a linearly independent for its closed linear span, then

$$\text{Lö}(g) := \sum_{i,j,0,0} R_{i,j}^{-1/2} g_{i,j}$$

generates an orthonormal system for that span. Furthermore, if g is a gaussian then $\text{Lö}(g)$ is optimally robust against time- and frequency-dispersion.

Consider the 320-by-200-pixel image below



Probably a clown

- This is stored as a 320×200 matrix of grayscale values, between 0 (black) and 1 (white), denoted by A_{clown}
- We can take the SVD of A_{clown}
- Replacing the smallest $n-k$ singular values by 0 in the SVD of A_{clown} yields the best rank-k approximation, denoted $A_{\text{clown}}^{(k)}$, to A_{clown} as measured by the Frobenius norm
- Storage required for $A_{\text{clown}}^{(k)}$ is a total of $(320+200) \cdot k$ bytes for storing $\sigma_1 u_1$ through $\sigma_k u_k$ and v_1 through v_k
- $320 \cdot 200 = 64{,}000$ bytes required to store A_{clown} explicitly

Now consider the rank-20 approximation to the original image, and the difference between the images


Most of a clown, with the missing parts

The original image took 64 kb, while the low-rank approximation required $(320+200)\cdot 20 \;=\; 10.4$ kb, a compression ratio of .1625

Scott Beaver - Western Oregon University

beavers@wou.edu

www.wou.edu/~beavers