Matrices and Their Adjoints
An Inclusive Approach to Basic Linear Algebra

Jack W. Rogers, Jr.*

January 7, 2008

The object of this talk is to show how to unify the treatment of R^n with that of more general vector spaces, such as P_n, the space of polynomials of degree $\leq n$, and $C[a,b]$, the space of continuous functions defined on the interval $[a,b]$, in a standard first course in linear algebra. To this end, we begin by generalizing the definition of a matrix.

 Department of Mathematics, Auburn University, AL 36849; e-mail: rogerj2@auburn.edu.
Slides presented at the Special Session on Innovative and Effective Ways to Teach Linear Algebra at the 2008 Joint Sessions in San Diego, January 7, 2008.

1 Matrices

Definition 1.1 *A* matrix *is a* $1 \times n$ *array*

$$V = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \in V^{1\times n}$$

of vectors in a vector space Σ *for some positive integer* n*. For each* $\vec{c} \in \mathbf{R}^n$*, the matrix-vector product* $V\vec{c}$ *is the linear combination*

$$V\vec{c} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n.$$

Examples of matrix-vector products are

$$\begin{bmatrix} \vec{1} & \vec{t} & \overrightarrow{t^2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \overrightarrow{1 + 2t + t^2},$$

where $\overrightarrow{f(t)}\,(a) = f(a)$, and

$$\begin{bmatrix} \overrightarrow{\sin^2 t} & \overrightarrow{\cos^2 t} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \overrightarrow{\sin^2 t} + \overrightarrow{\cos^2 t} = \vec{1}.$$

The array $A \in \mathbf{R}^{m\times n}$ satisfies this definition if we consider it to be the $1 \times n$ array of its columns,

$$A = \begin{bmatrix} \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} & \cdots & \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix}.$$

Note, however, that the $n \times 1$ array of its rows

$$\begin{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{m1} & \cdots & a_{mn} \end{bmatrix} \end{bmatrix}$$

is not a matrix under this definition. We will return to this mathematical object later.

We usually speak of the image and kernel of a matrix, when we mean the image and kernel of the associated linear transformation T_A. We extend this identification of a matrix with its associated linear transformation, unless clarity demands more care, as the following examples show.

Definition 1.2 .
1) The columns of the matrix $V : \mathbf{R}^n \to \mathbf{\Sigma}$ span $\mathbf{\Sigma}$ *if and only V is onto.*
2) The columns of V are independent *if and only if V is one-to-one.*
3) The columns of V form a basis *for $\mathbf{\Sigma}$ if and only if V is an isomorphism (i.e., both onto and one-to-one) or, equivalently, the columns of V are independent and span $\mathbf{\Sigma}$.*

Each of these definitions relates attributes of a sequence of vectors to attributes of the associated linear transformation.

1.1 The matrix-matrix product

Definition 1.3 *If $A\colon \mathbf{R}^n \to \mathbf{R}^k$ and $V: \mathbf{R}^k \to \mathbf{\Sigma}$ are matrices, the matrix-matrix product VA is the matrix for the functional composition $V \circ A$, with diagram*

$$\begin{array}{ccc} \mathbf{R}^n & \stackrel{A}{\longrightarrow} & \mathbf{R}^k \\ & {\scriptstyle VA} \searrow & \downarrow_V \\ & & \mathbf{\Sigma} \end{array}$$

The product satisfies

$$\begin{aligned} VA &= V \begin{bmatrix} \vec{a}_1 & \cdots & \vec{a}_n \end{bmatrix} \\ &= \begin{bmatrix} V\vec{a}_1 & \cdots & V\vec{a}_n \end{bmatrix} \end{aligned}$$

For example,

$$\begin{bmatrix} \overrightarrow{\sin^2 t} & \overrightarrow{\cos^2 t} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \overrightarrow{\sin^2 t} + \overrightarrow{\cos^2 t} & -\overrightarrow{\sin^2 t} + \overrightarrow{\cos^2 t} \end{bmatrix} = \begin{bmatrix} \vec{1} & \overrightarrow{\cos 2t} \end{bmatrix}$$

The standard treatments of dimension and coordinates relative to a basis work for these matrices, but there is no analog for the LU decomposition. Instead, we have the following decomposition.

1.2 The BR decomposition

A BR decomposition of a matrix $V\colon \mathbf{R}^n \to \mathbf{\Sigma}$ consists of a one-to-one matrix $B\colon \mathbf{R}^k \to \mathbf{\Sigma}$ and an onto matrix $R\colon \mathbf{R}^n \to \mathbf{R}^k$ in reduced row echelon form such that $V = BR$.

$$\begin{tikzcd} \mathbf{R}^n \arrow[r, "R\text{ (onto)}"] \arrow[dr, "V"'] & \mathbf{R}^k \arrow[d, "B\text{ (1-1)}"] \\ & \operatorname{im} V \subset \mathbf{\Sigma} \end{tikzcd}$$

Since R is onto, $\operatorname{im} V = \operatorname{im} B$, so the columns of B form a basis for $\operatorname{im} V$. Hence the name of the decomposition.

Example 1.4 *$\overrightarrow{\sin^2 t}$ and $\overrightarrow{\cos^2 t}$ are independent, since scalar multiples have the same roots, but $\vec{1} = (1)\overrightarrow{\sin^2 t} + (1)\overrightarrow{\cos^2 t}$, yielding the decomposition*

$$V = \begin{bmatrix} \overrightarrow{\sin^2 t} & \overrightarrow{\cos^2 t} & \vec{1} \end{bmatrix} = \underbrace{\begin{bmatrix} \overrightarrow{\sin^2 t} & \overrightarrow{\cos^2 t} \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}_{R}.$$

Clearly, the decomposition is unique. Also, since B is one-to-one, $\ker V = \ker R = \operatorname{span} \begin{bmatrix} -1 & -1 & 1 \end{bmatrix}^T$.

For an array $A \in \mathbf{R}^{m\times n}$, reduction to reduced row echelon form yields $\hat{R}$. The matrix B consists of the pivot columns of A, and the matrix R consists of the *nonzero* rows of $\hat{R}$.

If W is a standard basis for $\mathbf{\Sigma}$, then the matrix V_W satisfying $V = WV_W$ is easily found, and Gaussian elimination techniques yield $V_W = BR$. Then $V = WV_W = (WB)\,R$ is the BR decomposition of V, since both W and B are one-to-one.

Example 1.5 *The standard basis for P_2 is $W = \left(\vec{1}, \vec{t}\right)$. If*

$$V = \left[\ \overrightarrow{1+2t}\ \ \overrightarrow{2+4t}\ \ \overrightarrow{2+5t}\ \ \overrightarrow{1+t}\ \right], \textit{ then}$$

$$V_W = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 4 & 5 & 1 \end{bmatrix} \sim \cdots \sim \begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \textit{ so}$$

$$V_W = \underbrace{\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}}_{R}, \textit{ and}$$

$$V = \underbrace{\left[\ \overrightarrow{1+2t}\ \ \overrightarrow{2+5t}\ \right]}_{WB} \underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{bmatrix}}_{R}.$$

2 Adjoints

From this point on, all spaces are inner product spaces, and $\mathbf{R}^n$ and $C[a,b]$ have the standard inner products defined by $\langle \vec{v}, \vec{w} \rangle_{\mathbf{R}^n} = \vec{v} \cdot \vec{w}$ and $\langle f, g \rangle_{C[a,b]} = \int_a^b f(t)g(t)\ dt$. The transpose of an array satisfies

$$(A\vec{c}) \cdot \vec{d} = (A\vec{c})^T \vec{d} = \vec{c}^T \left(A^T \vec{d}\right) = \vec{c} \cdot \left(A^T \vec{d}\right).$$

The analog for general linear transformations between inner product spaces is the adjoint.

Definition 2.1 *If $T\colon \mathbf{\Theta} \to \mathbf{\Sigma}$ is a linear transformation, the linear transformation $T^*\colon \mathbf{\Sigma} \to \mathbf{\Theta}$ is an* adjoint *for T if and only if*

$$\left\langle T\left(\vec{t}\right), \vec{s} \right\rangle_{\mathbf{\Sigma}} = \left\langle \vec{t}, T^*(\vec{s}) \right\rangle_{\mathbf{\Theta}}$$

for all $\vec{s} \in \mathbf{\Sigma}$ and $\vec{t} \in \mathbf{\Theta}$.

As expected, the adjoint shares many of the properties of the transpose, e.g., $(T^*)^* = T$, $(TS)^* = S^*T^*$, etc.

2.1 The adjoint of a matrix

If $V : R^n \to \mathbf{\Sigma}$ is a matrix, then

$$\begin{aligned}
\langle V\vec{c}, \vec{s}\rangle_{\mathbf{\Sigma}} &= c_1 \langle \vec{v}_1, \vec{s}\rangle_{\mathbf{\Sigma}} + \cdots + c_n \langle \vec{v}_n, \vec{s}\rangle_{\mathbf{\Sigma}} \\
&= \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \cdot \begin{bmatrix} \langle \vec{v}_1, \vec{s}\rangle_{\mathbf{\Sigma}} \\ \vdots \\ \langle \vec{v}_n, \vec{s}\rangle_{\mathbf{\Sigma}} \end{bmatrix} = \langle \vec{c}, V^*\vec{s}\rangle_{\mathbf{R}^n}\text{, where}
\end{aligned}$$

$$V^* = \begin{bmatrix} \langle \vec{v}_1, \cdot\rangle_{\mathbf{\Sigma}} \\ \vdots \\ \langle \vec{v}_n, \cdot\rangle_{\mathbf{\Sigma}} \end{bmatrix}$$

Example 2.2 *If* $V : \mathbf{R}^2 \to C\,[0,1] = \begin{bmatrix} \vec{1} & \vec{t} \end{bmatrix}$*, and* $\vec{s} = \vec{t^2}$*, then* $V^*\vec{s} = \begin{bmatrix} \vec{1} & \vec{t} \end{bmatrix}^* \vec{s} =$

$$\begin{bmatrix} \vec{1}^*\overrightarrow{t^2} \\ \\ \vec{t}^*\overrightarrow{t^2} \end{bmatrix} = \begin{bmatrix} \left\langle \vec{1}, \overrightarrow{t^2} \right\rangle_{C[0,1]} \\ \\ \left\langle \vec{t}, \overrightarrow{t^2} \right\rangle_{C[0,1]} \end{bmatrix} = \begin{bmatrix} \int_0^1 \vec{1}\ \overrightarrow{t^2}\ dt \\ \\ \int_0^1 \vec{t}\ \overrightarrow{t^2}\ dt \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/4 \end{bmatrix}.$$

2.2 Products of matrices and matrix-adjoints

matrix-matrix VA	adjoint-adjoint A^*V^*
$\begin{array}{ccc} \mathbf{R}^n & \xrightarrow{A} & \mathbf{R}^k \\ & VA \searrow & \downarrow V \\ & & \boldsymbol{\Sigma} \end{array}$	$\begin{array}{ccc} \mathbf{R}^n & \xleftarrow{A^*} & \mathbf{R}^k \\ & A^*V^* \nwarrow & \uparrow V^* \\ & & \boldsymbol{\Sigma} \end{array}$
$\begin{bmatrix} V\vec{a}_1 & \cdots & V\vec{a}_n \end{bmatrix}$	$\begin{bmatrix} \vec{a}_1^*V^* \\ \vdots \\ \vec{a}_k^*V^* \end{bmatrix}$
adjoint-matrix V^*W	**matrix-adjoint VW^***
$\begin{array}{ccc} \mathbf{R}^n & \xrightarrow{V^*W} & \mathbf{R}^k \\ & W \searrow & \uparrow V^* \\ & & \boldsymbol{\Sigma} \end{array}$	$\begin{array}{ccc} & & \mathbf{R}^k \\ & W^* \nearrow & V \downarrow \\ \boldsymbol{\Theta} & \xrightarrow{VW^*} & \boldsymbol{\Sigma} \end{array}$
$\begin{bmatrix} \vec{v}_1^*\vec{w}_1 & \cdots & \vec{v}_1^*\vec{w}_n \\ \vdots & & \vdots \\ \vec{v}_k^*\vec{w}_1 & \cdots & \vec{v}_k^*\vec{w}_n \end{bmatrix}$	$\vec{v}_1\vec{w}_1^* + \cdots + \vec{v}_k\vec{w}_k^*,$

In each case, we multiply the rows of the first object with the columns of the second object. However, all these forms coalesce into different ways of computing the same product only when all matrices are rectangular arrays of numbers.

2.3 The four fundamental spaces

Throughout this section, $T\colon \mathbf{\Theta} \to \mathbf{\Sigma}$ is linear with an adjoint $T^*\colon \mathbf{\Sigma} \to \mathbf{\Theta}$. There are four fundamental spaces associated with T: $\operatorname{im} T$ and $\ker T^*$ in $\mathbf{\Sigma}$, and $\ker T$ and $\operatorname{im} T^*$ in $\mathbf{\Theta}$.

Theorem 2.3 $\ker T = (\operatorname{im} T^*)^{\perp}$.

Proof.

$$\begin{aligned}
\vec{t} \in \ker T &\iff T\vec{t} = \vec{0} \\
&\iff \left\langle T\vec{t}, \vec{s}\right\rangle_{\mathbf{\Sigma}} = 0 \textit{ for all } \vec{s} \in \mathbf{\Sigma} \\
&\iff \left\langle \vec{t}, T^*\vec{s}\right\rangle_{\mathbf{\Theta}} = 0 \textit{ for all } \vec{s} \in \mathbf{\Sigma} \\
&\iff \vec{t} \in (\operatorname{im} T^*)^{\perp}
\end{aligned}$$

■

While many of the fundamental theorems involving images, kernels and orthogonality do not hold for all linear transformations between inner product spaces, they do hold for matrices, e.g., $\operatorname{im} V^* = (\ker V)^{\perp}$, $\left((\operatorname{im} V)^{\perp}\right)^{\perp} = \operatorname{im} V$, etc.

2.4 A least-square fit

The treatment of discrete and continuous least-square fits is the same. For example, the continuous first-degree least-square fit in $\mathbf{\Sigma} = C\,[0,1]$ to $\vec{s} = \overrightarrow{t^2}$ is the orthogonal projection of $\vec{s}$ onto the image of $V = \left[\begin{array}{cc} \vec{1} & \vec{t} \end{array}\right]$. As in the discrete case, we first solve the normal equation $V^*V\vec{n} = V^*\vec{s}$ for $\vec{n}$.

$$[V^*V \,\vdots\, V^*\vec{s}] = V^*\,[V \,\vdots\, \vec{s}] = \left[\begin{array}{c} \vec{1}^* \\ \vec{t}^* \end{array}\right]\left[\begin{array}{ccc} \vec{1} & \vec{t} & \vdots & \overrightarrow{t^2} \end{array}\right] =$$

$$\left[\begin{array}{cccc} \int_0^1 1\ dt & \int_0^1 t\ dt & \vdots & \int_0^1 t^2\ dt \\ \int_0^1 t\ dt & \int_0^1 t^2\ dt & \vdots & \int_0^1 t^3\ dt \end{array}\right] = \left[\begin{array}{cccc} 1 & \frac{1}{2} & \vdots & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \vdots & \frac{1}{4} \end{array}\right]$$

$$\sim \cdots \sim \left[\begin{array}{cccc} 1 & 0 & \vdots & -\frac{1}{6} \\ 0 & 1 & \vdots & 1 \end{array}\right] \Longrightarrow \vec{n} = \left[\begin{array}{c} -\frac{1}{6} \\ 1 \end{array}\right].$$

So

$$P\overrightarrow{t^2} = V\vec{n} = \left[\begin{array}{cc} \vec{1} & \vec{t} \end{array}\right]\left[\begin{array}{c} -\frac{1}{6} \\ 1 \end{array}\right] = \overrightarrow{t - \frac{1}{6}}.$$

The algorithm is the same as in the discrete case; only the computation of the inner product is different. Here are plots of both the parabola and the continuous least square fit,

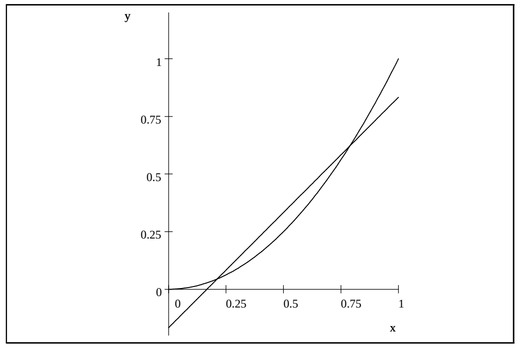


Figure 1: A least square fit.

3 Matrix decompositions

We next show that several important decompositions of a matrix V can be obtained from decompositions of V^*V.

Theorem 3.1 *Suppose $O\colon \Theta \to \Gamma$, $B\colon \Gamma \to \Theta$, and $T\colon \Theta \to \Sigma$ are linear transformations. If $T^*T = BO$ is a decomposition of T^*T and O is onto, then there is a unique linear transformation W which maps Γ onto $\operatorname{im} T$ so that the following diagram commutes. If $B\colon \Gamma \to \operatorname{im} T^*T$ is an isomorphism, so is $W\colon \Gamma \to \operatorname{im} T$.*

$$\begin{tikzcd}
& \Gamma \arrow[dr, "B"] \arrow[dd, "W"] & \\
\Theta \arrow[ur, "O\ \text{(onto)}"] \arrow[dr, "T"'] & & \operatorname{im} T^*T \subset \Theta \\
& \operatorname{im} T \subset \Sigma \arrow[ur, "T^*"'] &
\end{tikzcd} \tag{3.1}$$

3.1 The BR decomposition

If $V : \mathbf{R}^n \to \mathbf{\Sigma}$ is a matrix, then $V^*V \in \mathbf{R}^{n\times n}$, so Gaussian elimination yields $V^*V = BR$, Diagram 3.1 becomes

$$\begin{tikzcd}
& \mathbf{R}^k \arrow[dr, "B"] \arrow[d, "\bar{B}"] & \\
\mathbf{R}^n \arrow[ur, "R \text{ (onto)}"] \arrow[dr, "V"'] & & \text{im}\, V^*V \subset \mathbf{R}^n \\
& \text{im}\, V \subset \mathbf{\Sigma} \arrow[ur, "V^*"'] &
\end{tikzcd}$$

$B = \begin{bmatrix} V^*\vec{v}_{i_1} & \cdots & V^*\vec{v}_{i_k} \end{bmatrix}$ consists of the pivot columns of V^*V, and $\bar{B} = \begin{bmatrix} \vec{v}_{i_1} & \cdots & \vec{v}_{i_k} \end{bmatrix}$, consists of the columns of V corresponding to the pivot columns of V^*V.

The columns of R^T form a basis for im V^*, so the BR decomposition provides a basis for three of the four fundamental spaces for V. We only lack a basis for ker V^*, which we are not going to get in general, since ker V^* may be infinite dimensional.

3.2 Isometries and orthogonal matrices

Definition 3.2 *An* isometry $Q\colon \Gamma \to \boldsymbol{\Sigma}$ *is a linear transformation from* Γ *onto* $\boldsymbol{\Sigma}$ *that preserves distances, i.e.*

$$\|Q(\vec{v}-\vec{w})\| = \|Q\vec{v} - Q\vec{w}\| = \|(\vec{v}-\vec{w})\|$$

for all $\vec{v}, \vec{w} \in \Gamma$.

Theorem 3.3 Q *is an isometry if and only if* $Q^*Q = id_\Gamma$.

Theorem 3.4 *Suppose* $S\colon \mathbf{R}^n \to \mathbf{R}^k$ *and* $T\colon \Theta \to \boldsymbol{\Sigma}$ *are linear with adjoints satisfying* $S^*S = T^*T$*, and* S *is onto. Then there is an isometry* $Q\colon \mathbf{R}^k \to \operatorname{im} T$ *that makes the following diagram commutative.*

$$\begin{tikzcd} & \mathbf{R}^k \arrow[d, "Q"] \arrow[dr, "S^*"] & \\ \mathbf{R}^n \arrow[ur, "S\ (\text{onto})"] \arrow[dr, "T"'] & \operatorname{im} T \subset \boldsymbol{\Sigma} \arrow[r, "T^*"'] & \operatorname{im} T^*T \subset \mathbf{R}^n \\ & & \end{tikzcd}$$

For the following two resulting decompositions, we plot the columns of S to obtain a picture of the columns of $V = \begin{bmatrix} \sin^2 t & \cos^2 t \end{bmatrix}$ with $\langle f, g\rangle_{\boldsymbol{\Sigma}} = \int_{-\pi}^{\pi} f(t)g(t)\, dt$.

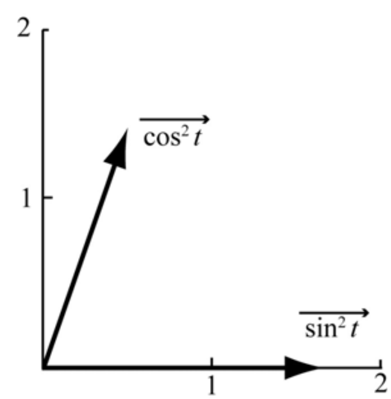
3.3 The QR decomposition

If $V : \mathbf{R}^n \to \mathbf{\Sigma}$ is an isomorphism, then $V^*V : \mathbf{R}^n \to \mathbf{R}^n$ is a positive definite array, so it has an LDL^T decomposition,

$$V^*V = LDL^T = \left(\sqrt{D}L^T\right)^T \left(\sqrt{D}L^T\right) = R^T R.$$

Theorem 3.4 applies to yield the isometry Q so that $V = QR$.

$$\begin{tikzcd}
 & \mathbf{R}^n \arrow[dr, "R^T = L\sqrt{D}"] \arrow[dd, "Q"] & \\
\mathbf{R}^n \arrow[ur, "R = \sqrt{D}L^T"] \arrow[dr, "V"'] & & \mathbf{R}^n \\
 & \mathbf{\Sigma} \arrow[ur, "V^*"'] &
\end{tikzcd}$$



The columns of V.

3.4 The singular value decomposition

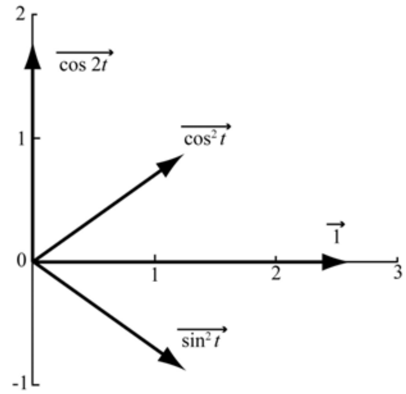
If $V : \mathbf{R}^n \to \boldsymbol{\Sigma}$ is one-to-one, then V^*V is symmetric and positive definite, so it has an orthogonal diagonalization

$$(V^*V) = PDP^T = \left(\boldsymbol{\Sigma} P^T\right)^T \left(\boldsymbol{\Sigma} P^T\right),$$

where $\boldsymbol{\Sigma} = \sqrt{D}$, yielding the diagram below. The resulting orthogonal matrix Q satisfies $V = Q\boldsymbol{\Sigma} P^T$, which is called the *singular value decomposition* of V.

$$\begin{tikzcd}
 & \mathbf{R}^n \arrow[dr, "(\boldsymbol{\Sigma} P^T)^T"] \arrow[dd, "Q"] & \\
\mathbf{R}^n \arrow[ur, "\boldsymbol{\Sigma} P^T"] \arrow[dr, "V"] & & \mathbf{R}^n \\
 & \boldsymbol{\Sigma} \arrow[ur, "V^*"] &
\end{tikzcd}$$

Example 3.5 *For* $V = \left[\ \overrightarrow{\sin^2 t}\ \ \overrightarrow{\cos^2 t}\ \right]$*, the columns of* $Q = VP^{-T}\boldsymbol{\Sigma}^{-1}$ *turn out to be multiples of* $\vec{1}$ *and* $\overrightarrow{\cos 2t}$*. Plotting the columns of* $\boldsymbol{\Sigma} P^T$ *yields this figure.*



The geometry of four vectors in $\operatorname{im} V$.

The standard identities, $\vec{1} = \overrightarrow{\cos^2 t} + \overrightarrow{\sin^2 t}$ *and* $\overrightarrow{\cos 2t} = \overrightarrow{\cos^2 t} - \overrightarrow{\sin^2 t}$ *are evident in the figure.*