Vector Space and Linear Transform

♣ Vector space, Subspace, Examples

♣ Null space, Column space, Row space of a matrix

♣ Spanning sets and Linear Independence

♣ Basis and Dimension

♣ Rank of a matrix

♣ Vector norms and matrix norms

♣ Linear transform

- Projection, rotation, reflection
- Gauss transform
- Householder transform (Elementary reflector)
- Jacobi transform (Givens' rotation)

◇ Affine transform with applications

Definition of A Vector Space

Definition: A vector space V (over R) is a set on which the operations of addition $\oplus$ and scalar multiplication $\odot$ are defined. The set V associated with the operations of addition and scalar multiplication is said to form a *vector space* if the following axioms are satisfied.

(A1) $\mathbf{x} \oplus \mathbf{y} = \mathbf{y} \oplus \mathbf{x} \quad \forall\, \mathbf{x}, \mathbf{y} \in V$

(A2) $(\mathbf{x} \oplus \mathbf{y}) \oplus \mathbf{z} = \mathbf{x} \oplus (\mathbf{y} \oplus \mathbf{z}) \quad \forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

(A3) $\exists\, \mathbf{0} \in V$ such that $\mathbf{x} \oplus \mathbf{0} = \mathbf{x} \quad \forall\, \mathbf{x} \in V$

(A4) $\forall\, \mathbf{x} \in V,\ \exists\ -\mathbf{x} \in V$ such that $\mathbf{x} \oplus (-\mathbf{x}) = \mathbf{0}$

(A5) $\alpha \odot (\mathbf{x} \oplus \mathbf{y}) = (\alpha \odot \mathbf{x}) \oplus (\alpha \odot \mathbf{y}), \quad \forall\, \alpha \in R$ and $\mathbf{x}, \mathbf{y} \in V$

(A6) $(\alpha + \beta) \odot \mathbf{x} = (\alpha \odot \mathbf{x}) \oplus (\beta \odot \mathbf{x}), \quad \forall\, \alpha, \beta \in R$ and $\mathbf{x} \in V$

(A7) $(\alpha \cdot \beta) \odot \mathbf{x} = \alpha \odot (\beta \odot \mathbf{x}), \quad \forall\, \alpha, \beta \in R$ and $\mathbf{x} \in V$

(A8) $1 \odot \mathbf{x} = \mathbf{x}$ for a $1 \in R$ *and* $\forall\, \mathbf{x} \in V$

Examples

(1) R^n (over R), in particular, $n = 2, 3$

(2) $C[a, b]$, for example, C[0,1]

(3) $P_n \ = \ \{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 |\ a_j \in R\}$

(4) $R^{m \times n}$ $=$ the set of all m by n real matrices

Exercises for Vector Space

1. Mark ◯ if the vector addition and scalar multiplication forms a vector space, otherwise mark $\times$.

($\times$) **(a)** For $(R^3, \oplus, \odot)$, the set of all triples of real numbers $[x, y, z]$ with the operations

$$[u, v, w] \oplus [x, y, z] = [u + x, v + y, w + z] \ \ and \ \ \alpha \odot [x, y, z] = [\alpha x, y, z]$$

($\times$) **(b)** For $(R^3, \oplus, \odot)$, the set of all triples of real numbers $[x, y, z]$ with the operations

$$[u, v, w] \oplus [x, y, z] = [u + x, v + y, w + z] \ \ and \ \ \alpha \odot [x, y, z] = [0, 0, 0]$$

($\times$) **(c)** For $(R^2, \oplus, \odot)$, the set of all paris of real numbers $[x, y]$ with the operations

$$[u, v] \oplus [x, y] = [u + x, v + y] \ \ and \ \ \alpha \odot [x, y] = [2\alpha x, 2\alpha y]$$

($\times$) **(d)** For $(R^2, \oplus, \odot)$, the set of all paris of real numbers $[x, y]$ with the operations

$$[u, v] \oplus [x, y] = [u + x + 1, v + y + 1] \ \ and \ \ \alpha \odot [x, y] = [\alpha x, \alpha y]$$

(◯) **(e)** For $(V, \oplus, \odot)$, where $V = \{[1, y] | y \in R\}$, the set of all paris of real numbers $[1, y]$ with the operations

$$[1, x] \oplus [1, y] = [1, x + y] \ \ and \ \ \alpha \odot [1, y] = [1, \alpha y]$$

(◯) **(f)** For $(V, \oplus, \odot)$, where $V = \{x \in R | x > 0\}$, $\alpha \in R$,

$$x \oplus y = xy \ \ and \ \ \alpha \odot x = x^{\alpha}$$

(◯) **(g)** For $(V, \oplus, \odot)$, where $V = \{a + bx | a, b \in R\}$,

$$(a + bx) \oplus (c + dx) = (a + c) + (b + d)x \ \ and \ \ \alpha \odot (c + dx) = (\alpha c) + (\alpha d)x$$

Subspaces of Vector Space

Definition: A subspace U of a vector space V is a *nonempty* subset satisfying

$$\mathbf{x} \oplus \mathbf{y} \in U \quad and \quad \alpha \odot \mathbf{x} \in U \quad \forall\, \mathbf{x}, \mathbf{y} \in U;\ \alpha \in R$$

Examples

The set of lower-Δ (upper-Δ) matrices

The set of tridiagonal (diagonal, Hessenberg) matrices

Let $A \in R^{m\times n}$, $A = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]$, and $A^t = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_m]$, then

$Null(A) = \{\mathbf{x} \in R^n |\ A\mathbf{x} = \mathbf{0}\} \subset R^n$ (**Nullspace**)

$R(A) = \{\sum_{j=1}^{n} \alpha_j \mathbf{a}_j |\ \alpha_j \in R\} \subset R^m$ (**Column space**)

$R(A^t) = \{\sum_{i=1}^{m} \beta_i \mathbf{b}_i |\ \beta_i \in R\} \subset R^n$ (**Row space**)

Theorem: The system $A\mathbf{x} = \mathbf{b}$ is solvable iff the vector $\mathbf{b}$ can be expressed as a linear combination of the columns of A

$$A = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]$$

$$A\mathbf{x} = \mathbf{b} \quad iff \quad \sum_{i=1}^{n} x_i \mathbf{a}_i = \mathbf{b}$$

Overdetermined, Underdetermined, Homogeneous Systems

$$\begin{array}{ccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
. & & . & & . & & . & & . \\
. & & . & & . & & . & & . \\
a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
\end{array}$$

Definition: A linear system is said to be *overdetermined* if there are more equations than unknowns $(m > n)$, *underdetermined* if $m < n$, *homogeneous* if $b_i = 0,\ \forall\ 1 \leq i \leq m$.

$$(A)\quad \begin{array}{rcrcr}
x & + & y & = & 1 \\
x & - & y & = & 3 \\
-x & + & 2y & = & -2
\end{array}
\qquad
(B)\quad \begin{array}{rcrcr}
x & + & y & = & 3 \\
x & - & y & = & 1 \\
2x & + & y & = & 5
\end{array}
\qquad
(C)\quad \begin{array}{rcrcr}
x & + & y & = & 2 \\
2x & + & 2y & = & 4 \\
-x & - & y & = & -2
\end{array}$$

(A) has no solution, (B) has unique solution, (C) has infinitely many solutions

$$(D)\quad \begin{array}{rcrcrcr}
x & + & 2y & + & z & = & -1 \\
2x & + & 4y & + & 2z & = & 3
\end{array}
\qquad
(E)\quad \begin{array}{rcrcrcr}
x & + & 2y & + & z & = & 5 \\
2x & - & y & + & z & = & 3
\end{array}$$

(D) has no solution, (E) has infinitely many solutions

Solutions of m Equations in n Unknowns

Theorem: $\forall\, A \in R^{m\times n}$, there corresponds a permutation matrix P, a unit lower-Δ matrix L, and an $m \times n$ upper trapezoidal matrix U such that $PA = LU$

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 6 \\ 2 & 5 & 6 & 8 \\ 1 & 3 & 4 & 5 \end{bmatrix} \;\Rightarrow\; PA = P_{34}P_{23}A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 8 \\ 1 & 3 & 4 & 5 \\ 1 & 2 & 3 & 6 \end{bmatrix}$$

$$PA = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 8 \\ 1 & 3 & 4 & 5 \\ 1 & 2 & 3 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{bmatrix} = LU$$

Linear Span

Definition: Let $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ be vectors in a vector space V. A sum of the form $\sum_{i=1}^{n} c_i \mathbf{v}_i$, where $c_i's$ are scalars, is called a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$. The *linear span* is the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ and is denoted by

$$span(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n)$$

□ In R^3, $span(\mathbf{e}_1, \mathbf{e}_2) = \{[a, b, 0]^t \mid a, b \in R\}$

□ The nullspace could be span$([1, -2, 1, 0]^t, [-1, 1, 0, 1]^t)$, where

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ \\ 2 & 1 & 0 & 1 \end{bmatrix}$$

□ $Null(A) = span([1, -2, 1, 0]^t, [-1, 1, 0, 1]^t)$

□ $Null(A) = span([1, -2, 1, 0]^t, [0, -1, 1, 1]^t)$

Theorem: If $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are elements of a vector space V, span$(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n)$ is a subspace of V.

Proof: Show that $a\mathbf{u} + b\mathbf{v} \in V, \ \forall\ a, b \in R; \ \mathbf{u}, \mathbf{v} \in V$

Spanning Sets

Definition: The set of $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$ is *a spanning set* for V iff each $\mathbf{v} \in V$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$.

(1) $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, [1,2,3]^t\}$ is a spanning set of R^3.

(2) $\{[1,1,1]^t, [1,1,0]^t, [1,0,0]^t\}$ is a spanning set of R^3.

(3) $\{[1,0,1]^t, [0,1,0]^t\}$ is not a spanning set of R^3.

(4) $\{[1,2,4]^t, [2,1,3]^t, [4,-1,1]^t\}$ is not a spanning set of R^3.

(5) $span(1, x, x^2) = span(1 - x^2, x + 2, x^2)$, where $P_2 = \{ax^2 + bx + c | \ a, b, c \in R\}$

Linear Independence

Definition: The vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ are said to be *linearly independent* if $\sum_{i=1}^{n} c_i\mathbf{v}_i = \mathbf{0}$ implies that $c_i = 0$ for $1 \leq i \leq n$. Otherwise, they are said to be linearly dependent.

$$\alpha \begin{bmatrix} 1 \\ \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} \;\Rightarrow\; \alpha = \beta = 0$$

Let

$$A = \begin{bmatrix} 3 & 4 & 2 \\ 0 & 1 & 5 \\ 0 & 0 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 2 & 4 \\ 4 & 8 \end{bmatrix}$$

(1) The column vectors of A are *linearly independent.*

(2) The column vectors of B are *linearly dependent.*

(3) The column vectors of C are *linearly dependent.*

Theorem: A set of n vectors in R^m must be linearly dependent if $n > m$

Basis and Dimension

Definition: *A basis* for a vector space is a set of vectors satisfying two properties: (1) it is linearly independent, (2) it spans the vector space.

• $\{\mathbf{e}_1, \mathbf{e}_2\}$ is not a basis for R^3 since $span(\mathbf{e}_1, \mathbf{e}_2) \neq R^3$

• The vectors $[1,0]^t, [0,1]^t, [2,1]^t$ spans R^2 but are not linearly independent so it is not a basis for R^2

Definition: Any two bases for a vector space V contain the same number of vectors. This number, shared by all bases and expresses the number of freedom of the space, is called the *dimension* of V.

Theorem: Suppose that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$ and $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_n$ are both bases for the same vector space S, then $m = n$.

Theorem: Any linearly independent set in a vector space V can be extended to a basis by adding more vectors if necessaary. Any spanning set in V can be reduced to a basis by discarding vectors if necessaary.

Example: Let $A \in R^{64\times 17}$ be a matrix of rank 11.

(1) $6 = (17 - 11)$ independent vectors $\mathbf{x}$ satisfy $A\mathbf{x} = \mathbf{0}$

(2) $53 = (64 - 11)$ independent vectors $\mathbf{y}$ satisfy $A^t\mathbf{y} = \mathbf{0}$

The Rank of A Matrix

□ The rank of a matrix $A \in R^{m\times n}$ can be defined as the number of linear independent columns. In Matlab command:

```
rank(A)
```

.

$$A = \begin{bmatrix} 1 & 2 & 3 & -2 \\ 2 & -1 & 1 & 1 \\ 1 & 1 & 2 & -1 \end{bmatrix}$$

Suppose the Gaussian elimination reduces A$\mathbf{x}$=$\mathbf{b}$ to U$\mathbf{x}$=$\mathbf{c}$ with r pivots, i.e., the last $m-r$ rows are zero. Then, there is a solution only if the last $m-r$ components of $\mathbf{c}$ are also zero. If $m=r$, there is always a solution. The general solution is the sum of a particular solution (with all free variables zero) and a homogeneous solution (with $n-r$ free variables as independent parameters). If $r=n$, there are no free variables and the nullspace contains only $\mathbf{x}$=$\mathbf{0}$. The number r is called the rank of matrix A.

Suppose $\mathbf{x}_p$ satisfies $A\mathbf{x}_p = \mathbf{b}$ and $\mathbf{x}_h$ satisfies $A\mathbf{x}_h = \mathbf{0}$

Then $\mathbf{x}_g = \mathbf{x}_p + \mathbf{x}_h$ satisfies $A\mathbf{x}_g = A\mathbf{x}_p + A\mathbf{x}_h = \mathbf{b} + \mathbf{0} = \mathbf{b}$

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$

Four Fundamental Subspaces from a Matrix

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} \Rightarrow U = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow E = \begin{bmatrix} 1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

matrix $\Rightarrow$ row echelon form $\Rightarrow$ reduced row echelon form

♣ **Fundamental Theorem of Linear Algebra:** Let $A \in R^{m\times n}$ have rank r,

(1) $R(A)$: the column space of A, $dim(R(A)) = r$

(2) $N(A)$: the nullspace of A, $dim(N(A)) = n - r$

(3) $R(A^t)$: the row space of A (the column space of A^t), $dim(R(A^t)) = r$

(4) $N(A^t)$: the left nullspace of A (the column space of A^t), $dim(N(A^t)) = m - r$

- $N(A) = \{\mathbf{x} |\ A\mathbf{x} = \mathbf{0}\}$
- $R(A) = \{\sum_{j=1}^{n} t_j \mathbf{a}_j |\ A = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]\}$

The row space of A has the same dimension r as the row space of U because $R(A^t) = R(U^t)$. The nullspace $N(A)$ has dimension $n - r$.

(1) $dim(R(A)) + dim(N(A)) = r + (n - r) = n$

(2) $dim(R(A^t)) + dim(N(A^t)) = r + (m - r) = m$

Example: $A \in R^{3\times 4}$

$$A = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \Rightarrow U = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \begin{aligned} dim(R(A)) &= 2 \\ dim(N(A)) &= 4 - 2 \\ dim(R(A^t)) &= 2 \\ dim(N(A^t)) &= 3 - 2 \end{aligned}$$

Vector Norms

Definition: A vector norm on R^n is a function

$$\tau \ : \ R^n \ \rightarrow \ R^+ = \{x \geq 0 | \ x \in R\}$$

that satisfies

(1) $\tau(\mathbf{x}) > 0 \ \ \forall \ \mathbf{x} \neq \mathbf{0}, \ \tau(\mathbf{0}) = 0$

(2) $\tau(c\mathbf{x}) = |c|\tau(\mathbf{x}) \ \forall \ c \in R, \ \mathbf{x} \in R^n$

(3) $\tau(\mathbf{x}+\mathbf{y}) \leq \tau(\mathbf{x}) + \tau(\mathbf{y}) \ \forall \ \mathbf{x}, \mathbf{y} \in R^n$

Hölder norm (p-norm) $\|\mathbf{x}\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p} \ \ for \ p \geq 1.$

(p=1) $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ (Mahattan or City-block distance)

(p=2) $\|\mathbf{x}\|_2 = (\sum_{i=1}^n |x_i|^2)^{1/2}$ (Euclidean distance)

(p=∞) $\|\mathbf{x}\|_\infty = max_{1\leq i\leq n}\{|x_i|\}$ (∞-norm)

Matrix Norms

Definition: A matrix norm on $R^{m\times n}$ is a function

$$\tau \ : \ R^{m\times n} \ \rightarrow \ R^{+} = \{x \geq 0 |\ x \in R\}$$

that satisfies

(1) $\tau(A) > 0 \ \ \forall\ A \neq O,\ \tau(O) = 0$

(2) $\tau(cA) = |c|\tau(A)\ \forall\ c \in R,\ \ A \in R^{m\times n}$

(3) $\tau(A+B) \leq \tau(A) + \tau(B) \ \ \forall\ A, B \in R^{m\times n}$

Consistency Property: $\tau(AB) \leq \tau(A)\tau(B) \ \ \forall\ A, B$

(a) $\tau(A) = max\{|a_{ij}|\ |\ 1 \leq i \leq m,\ 1 \leq j \leq n\}$

(b) $\|A\|_F = \left[\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right]^{1/2}$ (Frӧbenius norm)

Subordinate Matrix Norm: $\|A\| = max_{\|\mathbf{x}\|\neq \mathbf{0}}\{\|A\mathbf{x}\|/\|\mathbf{x}\|\}$

(1) If $A \in R^{m\times n}$, then $\|A\|_1 = max_{1\leq j\leq n}\left(\sum_{i=1}^{m}|a_{ij}|\right)$

(2) If $A \in R^{m\times n}$, then $\|A\|_\infty = max_{1\leq i\leq m}\left(\sum_{j=1}^{n}|a_{ij}|\right)$

(3) Let $A \in R^{n\times n}$ be real symmetric, then $\|A\|_2 = max_{1\leq i\leq n}|\lambda_i|$, where $\lambda_i \in \lambda(A)$

Linear Transformation

Definition: A mapping L from a vector space V to a vector space W is said to be a linear transform (transformation) or a linear operator if

$$L((\alpha \odot_V \mathbf{v}_1) \oplus_V (\beta \odot_V \mathbf{v}_2)) = (\alpha \odot_W L(\mathbf{v}_1)) \oplus_W (\beta \odot_W L(\mathbf{v}_2)), \quad \forall\, \alpha, \beta \in R, \;\; \mathbf{v}_1, \mathbf{v}_2 \in V$$

Examples: Projection, Scaling, Rotation, Reflection on $V = R^2$

(a) $L(\mathbf{x}) = \mathbf{u}^t\mathbf{x}$, for $\mathbf{u} \in V$

(b) $L(\mathbf{x}) = s\mathbf{x}$, for $s \in R$

(c) $L(\mathbf{x}) = R_\theta \mathbf{x}$, where $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

(d) $L(\mathbf{x}) = \mathbf{y}$, where $y_1 = -x_1$ and $y_2 = x_2$

(e) $L(f) = F$, where $f \in C[a,b]$ and $F(x) = \int_a^x f(t)dt$

(f) $L(f) = f'$, where $f \in C^1[a,b]$ and $f'(x) = \frac{d}{dx} f(x)$

(g) $L(\mathbf{x}) = A\mathbf{x}$, where $\mathbf{x} \in R^n$, $A \in R^{m\times n}$

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

Image and Kernel

Let $L : V \rightarrow W$ be a linear transfrom, and let $S \subset V$ be a *subspace* of V.
The *kernel* of L, denoted by $Ker(L)$, is defined by

$$Ker(L) = \{\mathbf{v} \in V |\ L(\mathbf{v}) = \mathbf{0}\}$$

The *image* of S under L, denoted by $L(S)$, is defined by

$$L(S) = \{\mathbf{w} \in W |\ \mathbf{w} = L(\mathbf{v})\ for\ some\ \mathbf{v} \in S\}$$

Theorem: Let $L : V \rightarrow W$ be a linear transfrom, and let $S \subset V$ be a *subspace* of V, then

(a) $Ker(L)$ is a subspace of V

(b) $L(S)$ is a subspace of W

Changing Coordinates in R^2

$$\{\mathbf{e}_1, \mathbf{e}_2\} \quad \Rightarrow \quad \{\mathbf{v}_1, \mathbf{v}_2\}$$

Any vector in $\mathbf{w} \in R^2$ can be expressed as $\mathbf{w} = x\mathbf{e}_1 + y\mathbf{e}_2 = [x, y]^t$, suppose that we want to express $\mathbf{w}$ as a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ as $\mathbf{w} = x'\mathbf{v}_1 + y'\mathbf{v}_2 = [x', y']^t$. What are $\{x, y\}$ and $\{x', y'\}$ related?

$$\begin{bmatrix} x \\ \\ y \end{bmatrix} = x\mathbf{e}_1 + y\mathbf{e}_2 = x'\mathbf{v}_1 + y'\mathbf{v}_2 = [\mathbf{v}_1, \mathbf{v}_2] \begin{bmatrix} x' \\ \\ y' \end{bmatrix}$$

Then

$$\begin{bmatrix} x' \\ \\ y' \end{bmatrix} = [\mathbf{v}_1, \mathbf{v}_2]^{-1} \begin{bmatrix} x \\ \\ y \end{bmatrix}$$

Example: $\mathbf{v}_1 = [1, 1]^t$, $\mathbf{v}_2 = [-1, 1]^t$, then

$$[\mathbf{v}_1, \mathbf{v}_2] = \begin{bmatrix} 1 & -1 \\ \\ 1 & 1 \end{bmatrix} \quad \Rightarrow \quad [\mathbf{v}_1, \mathbf{v}_2]^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}$$

Gauss Transform

Define an elementary matrix as

$$E_{ik}(r) = I - r\mathbf{e}_i\mathbf{e}_k^t, \quad i > k \quad \Rightarrow \quad E_{ik}(r)^{-1} = I + r\mathbf{e}_i\mathbf{e}_k^t$$

A *Gauss* transform is a matrix of the form

$$\prod_{i=n}^{k+1} E_{ik} = E_{nk}E_{n-1,k}\cdots E_{k+1,k}$$

which can annihilate the components of a vector $\mathbf{x}$ after index k.

Examples

$$G = E_{31}(-1)E_{21}(2) = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} \quad \Rightarrow \quad G\mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

Householder Transform (Elementary Reflector)

$$H = I - 2\mathbf{u}\mathbf{u}^t, \quad \textit{where } \mathbf{u} \in R^n \textit{ with } \|\mathbf{u}\|_2 = 1$$

$$H^t = H \quad \textit{and} \quad H^{-1} = H$$

Let $\mathbf{x} = [3, 1, 5, 1]^t$, then $\|\mathbf{x}\|_2 = \sqrt{3^2 + 1^2 + 5^2 + 1^2} = 6$.

Define $\mathbf{v} = \mathbf{x} + \|\mathbf{x}\|_2 \mathbf{e}_1$, and let $\mathbf{u} = \mathbf{v}/\|\mathbf{v}\|_2$, then

$$H = I - 2\mathbf{u}\mathbf{u}^t = \frac{1}{54}\begin{bmatrix} -27 & -9 & -45 & -9 \\ -9 & 53 & -5 & -1 \\ -45 & -5 & 29 & -5 \\ -9 & -1 & -5 & 53 \end{bmatrix}, \quad \textit{and} \quad H\mathbf{x} = \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Jacobi Transform (Givens' Rotation)

$$J(i,k;\theta) = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & c & \cdot & -s & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & s & \cdot & c & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & 0 & \cdot & 0 & \cdot & 1 \end{bmatrix}$$

$J_{hh} = 1$ if $h \neq i$ or $h \neq k$, where $i < k$

$J_{ii} = J_{kk} = c = \cos\theta$

$J_{ki} = s = \sin\theta,\ J_{ik} = -s = -\sin\theta$

Let $\mathbf{x}, \mathbf{y} \in R^n$, then $\mathbf{y} = J(i,k;\theta)\mathbf{x}$ implies that

$y_i = cx_i - sx_k$

$y_k = sx_i + cx_k$

$c = \frac{x_i}{\sqrt{x_i^2+x_k^2}}, \quad s = \frac{-x_k}{\sqrt{x_i^2+x_k^2}},$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}, \quad then \quad J(2,4;\theta)\mathbf{x} = \begin{bmatrix} 1 \\ \sqrt{20} \\ 3 \\ 0 \end{bmatrix}$$

Affine Transform with Applications

$$\mathbf{y} = A\mathbf{x} + \mathbf{t} \quad \Rightarrow \quad \begin{bmatrix} x_{n+1} \\ \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} a_i & b_i \\ \\ c_i & d_i \end{bmatrix} \begin{bmatrix} x_n \\ \\ y_n \end{bmatrix} + \begin{bmatrix} e_i \\ \\ f_i \end{bmatrix}$$

w	a	b	c	d	e	f	$\lvert ad - bc \rvert$
1	0	0	0	0.16	0	0	0.01
2	0.85	0.04	-0.04	0.85	0	1.60	0.85
3	0.20	-0.26	0.23	0.22	0	1.60	0.07
4	-0.15	0.28	0.26	0.24	0	0.44	0.07

Table 1: **An IFS consisting of 4 affine transforms for Fern**

Textures Generated by Fractal Models

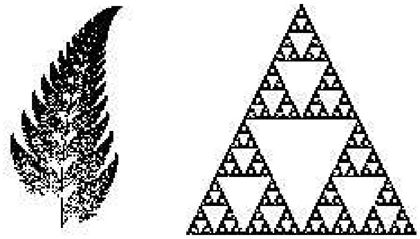
Fractal models used to generate such textures as *ferns*, *Sierpinski triangles*, and *snowflakes* have recently received attention in many image compression field. Synthesis is based on the iterated function system (IFS) codes [1,2,3], which are nothing but a set of affine transformations. Let $A \in R^{2\times 2}$ and $\mathbf{t} \in R^2$. An affine transform on $\mathbf{x} \in R^2$ is defined as $A\mathbf{x} + \mathbf{t}$. To describe the algorithm, we denote $A_i = \begin{bmatrix} a_i & b_i \\ c_i & d_i \end{bmatrix}$, with $p_i = |a_i d_i - b_i c_i| \neq 0$, and let $\mathbf{t}_i$ denote $\begin{bmatrix} e_i \\ f_i \end{bmatrix}$. An algorithm based on IFS codes with K affine transforms is listed below. Experiments conducted using two sets of affine transformations to generate textures are given. The parameters of this fractal model are given in Tables 1 and 2, respectively. Two such synthesized textures are shown in the following Figure.

Table 1. IFS codes for a fern

i	a_i	b_i	c_i	d_i	e_i	f_i
1	0	0	0	0.16	0	0
2	0.85	0.04	-0.04	0.85	0	1.60
3	0.20	-0.26	0.23	0.22	0	1.60
4	-0.15	0.28	0.26	0.24	0	0.44

Table 2. IFS codes for Sierpinski triangles

i	a_i	b_i	c_i	d_i	e_i	f_i
1	0.5	0	0	0.5	0	0
2	0.5	0	0	0.5	1.0	0
3	0.5	0	0	0.5	0.5	0.5



(a) (b)

Figure 1: Textures synthesized using IFS codes.

Fractal Generating Algorithm

(1) Set m=0 and randomly pick up an initial point $\mathbf{x}^{(0)} \in R^2$.

(2) Select $\{A_j, \mathbf{t}_j\}$ according to the probability distribution of $\{r_j,\ 1 \le j \le K\}$, where $r_j = p_j / \sum_{i=1}^{K} p_i$ for $1 \le j \le K$.

(3) $\mathbf{x}^{(m+1)} \longleftarrow A_j \mathbf{x}^{(m)} + \mathbf{t}_j$.

(4) $m \longleftarrow m + 1$.

(5) Repeat steps 2, 3, 4 until "*convergence,*" for example, $m = 1000$, is achieved.

(6) Plot $\{\mathbf{x}^{(i)}\}$ for $i = L$ to 1000, say L=100.

Convergence of this algorithm was studied by Barnsley [1], and Chu and Chen [3]. An IFS code consisting of two to five contractive affine transforms has been suggested [2,3] which generates self-similar images.

References

1. M.F. Anderson and L.P. Hurd, "Fractal image compression," Addison Wesley, 1993.

2. C.C. Chen and C.C. Chen, Texture Synthesis: A Review and Experiments, Journal of Information Science and Engineering, vol.19, no.2, 371-380, 2003.

3. H.T. Chu and C.C. Chen, "A Fast Algorithm for Generating Fractals," The Proceedings of IEEE Int'l Conference on Pattern Recognition, Barcelona, Spain, vol.3, 306-309, 2000.