Lineær algebra – 2
Ligningssystemer og vektorrum

Lineær algebra og dynamiske systemer

Troels B. Sørensen /Gilberto Berardinelli

Linear systems of equations

n : number of unknowns !

$$\left\{\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n = b_1 \\ \vdots \qquad\qquad\qquad\qquad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots a_{mn}x_n = b_m \end{array}\right\}$$

Or, shorthand

$$\bar{\bar{A}}\,\bar{x} = \bar{b}$$

$\bar{b} \neq \bar{0}$ nonhomogeneous system (inhomogen)

$\bar{b} = \bar{0}$ homogeneous system (homogen)

Coefficient matrix

Solution (unknown)

Linear system of equations

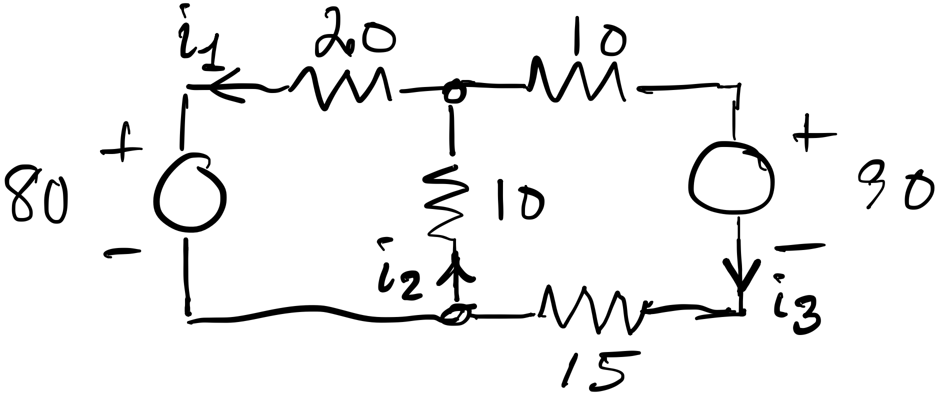
The augmented matrix is defined as [total matrix]

$$\tilde{\bar{\bar{A}}} = \left[\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right] = \left[\bar{\bar{A}} \mid \bar{b}\right]$$

and determines the system completely (all information about the linear system)

Gauss elimination leaves a row equivalent system, with the essential vectors (rows), and same rank, hence same solutions

Ex.: Solution by Gauss elimination..



$$\left.\begin{aligned} i_1 - i_2 + i_3 &= 0 \\ -i_1 + i_2 - i_3 &= 0 \end{aligned}\right\} \text{KCL}$$

$$\left.\begin{aligned} 10 i_2 + 25 i_3 &= 90 \\ 20 i_1 + 10 i_2 \quad &= 80 \end{aligned}\right\} \text{KVL}$$

$\bar{x} = ?$

$$\tilde{\bar{A}} = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ -1 & 1 & -1 & 0 \\ 0 & 10 & 25 & 90 \\ 20 & 10 & 0 & 80 \end{array}\right] \xrightarrow{\text{Gauss}} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & -95 & -190 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

using partial pivoting (swapping rows)

... and Backsubstitution

Backsubstitution

$$\tilde{\bar{\bar{A}}} = \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 10 & 25 & 90 \\ 0 & 0 & -95 & -190 \\ 0 & 0 & 0 & 0 \end{array}\right] \qquad \begin{aligned} \bar{i}_1 &= \bar{i}_2 - \bar{i}_3 = 2 \\ \bar{i}_2 &= (90 - 25\bar{i}_3)/10 = 4 \\ i_3 &= -190/-95 = 2 \\ 0 &= 0 \end{aligned}$$

- A unique solution (since $\text{rank}(\bar{\bar{A}}) = n$)
- An overdetermined system since $m > n$
- A consistent system since solutions exist

Possible linear systems

At the end of Gauss elimination

$$\tilde{\tilde{A}}' = \left.\left[\begin{array}{cccccc|c} a'_{11} & a'_{12} & a'_{13} & \cdots & \cdots & a'_{1n} & b'_1 \\ 0 & a'_{22} & a'_{23} & \cdots & \cdots & a'_{2n} & b'_2 \\ \vdots & & & & & \vdots & \vdots \\ 0 & & a'_{rr} & \cdots & \cdots & a'_{rn} & b'_r \\ 0 & & & & & 0 & b'_{r+1} \\ \vdots & & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & b'_m \end{array}\right]\right\} \text{rank } r$$

modified augmented matrix

n : number of unknowns

Possible linear systems (cont.)

a) no solutions if $r < m$ and $\{b'_{r+1} \ldots b'_m\} \neq 0$, otherwise

inconsistent

b) precisely one solution if $r = n$ and $\{b'_{r+1} \ldots b'_m\} = 0$ consistent

c) infinitely many solutions if $r < n$ and $\{b'_{r+1} \ldots b'_m\} = 0$ consistent

Possible linear systems (cont.)

- Overdetermined if $m > n$ $\quad r \leq n < m$
  [overbestemt]
- Determined if $m = n$ $\quad r \leq m = n$
  [bestemt]
- Underdetermined if $m < n$ $\quad r \leq m < n$
  [underbestemt]

"How does that relate to the number of solutions?"

Existence of solutions

$$\underline{\underline{A}}\,\underline{\underline{x}} = \underline{\underline{b}} \qquad \tilde{\underline{\underline{A}}} = [\underline{\underline{A}} \mid \underline{\underline{b}}]$$

$(m \times n)$ $\qquad$ $(m \times (n+1))$

Fundamental Theorem:

Solutions exist if $\text{rank}(\underline{\underline{A}}) = \text{rank}(\tilde{\underline{\underline{A}}})$

If $r = \text{rank}(\underline{\underline{A}}) = \text{rank}(\tilde{\underline{\underline{A}}})$

Consistent ?

1) $r = n$, precisely one solution

2) $r < n$, infinitely many solutions

Homogeneous system

$$\bar{\bar{A}}\,\bar{x} = \bar{0}$$

always one solution

If $r = \mathrm{rank}(\bar{\bar{A}})\ (= \mathrm{rank}(\tilde{\bar{\bar{A}}})) = n$

only one solution $\bar{x} = \bar{0}$ (obviously)

the trivial solution

Otherwise, infinitely many (r < n)

↳ we need this for non-trivial solutions ($\det(\bar{\bar{A}}) = 0$!)

Homogeneous system (cont.)

These solutions $+(\bar{x}=\bar{0})$ forms a vector space (will get to that) with dimension

$$\boxed{\text{nullity}(\bar{\bar{A}}) = n - \text{rank}(\bar{\bar{A}}) = n - r}$$

This solution space is named the null space of $\bar{\bar{A}}$ since all solutions $\bar{x}$ are vectors in this space and fulfills $\bar{\bar{A}}\,\bar{x} = \bar{0}$ maps to zero! (as does $\bar{x}=\bar{0}$)

Non-homogeneous system

Whereas linearity holds for homogeneous systems, i.e., $\bar{x} = c_1\bar{x}_1 + c_2\bar{x}_2$ is a solution if $\bar{x}_1$ and $\bar{x}_2$ are solutions, this is not the case for non-homogeneous systems. In this case the solution is

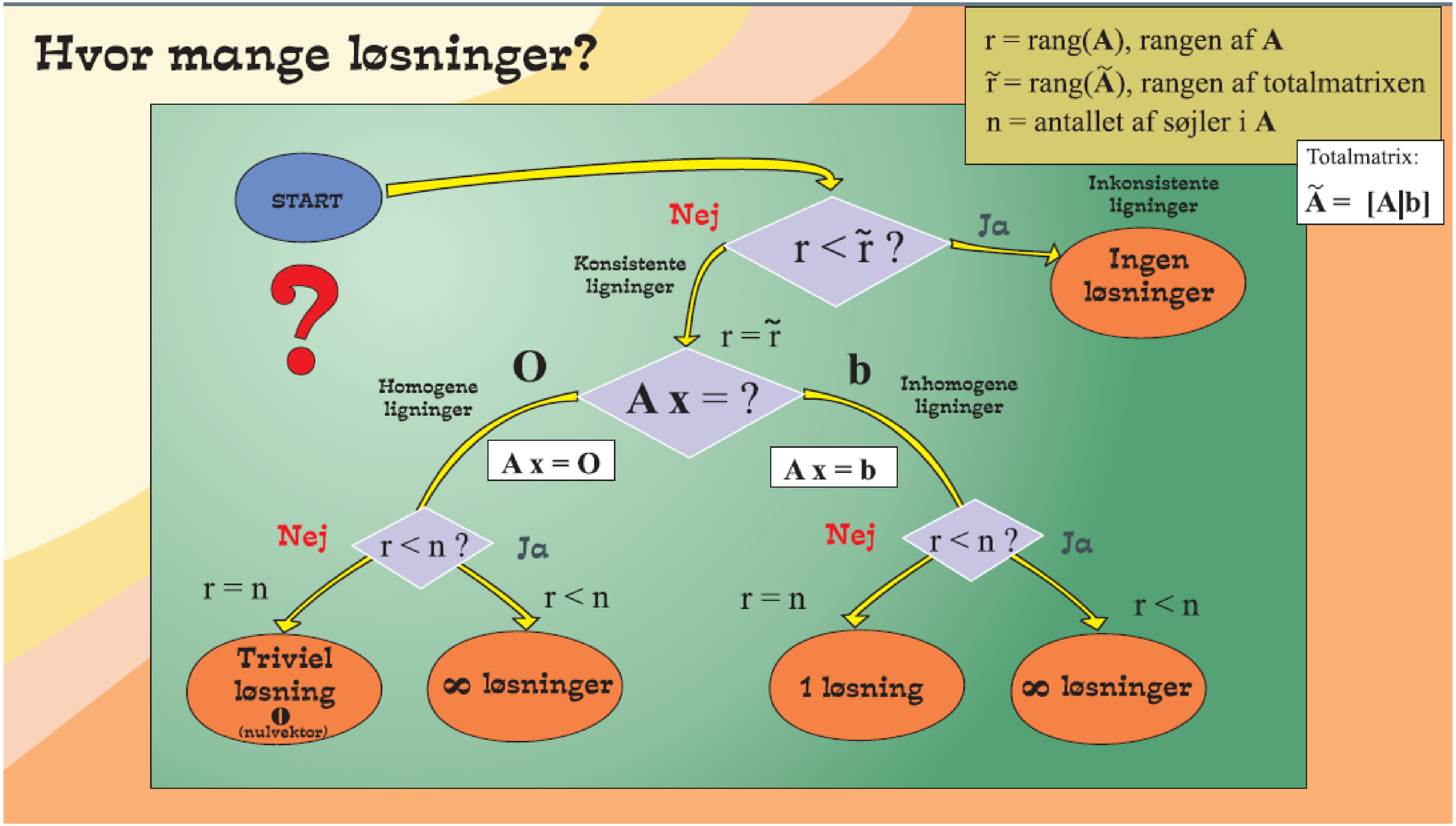
$$\bar{x} = \bar{x}_h + \bar{x}_p$$ fixed (particular) solution

where $\bar{\bar{A}}\bar{x}_h = \bar{0}$ and x_p is one solution to $\bar{\bar{A}}\bar{x}_p = \bar{b}$

Also, if $\bar{x}$ is any solution of $\bar{\bar{A}}\bar{x} = \bar{b}$, then infinitely many !

$$\bar{\bar{A}}(\underline{\bar{x} - \bar{x}_p}) = \bar{\bar{A}}\bar{x} - \bar{\bar{A}}\bar{x}_p \; (= \bar{\bar{A}}\bar{x}_h) = \bar{0}$$

Solutions to linear systems



Solution approaches

Gauss elimination $(m \times n)$:

solves any case

Use of inverse (algebraic):

$$\bar{\bar{A}}\,\bar{x} = \bar{b} \iff \bar{x} = \bar{\bar{A}}^{-1}\,\bar{b} \quad \text{IF}$$

$\bar{\bar{A}}$ is square and non-singular

Use of Cramer's rule:

$$\bar{x}_k = \det(\bar{\bar{A}}_{(k)}) / \det(\bar{\bar{A}}) \quad \text{IF}$$

Solution approaches (cont.)

$$\bar{\bar{A}}_k = [\bar{b}_1, \ldots, \bar{b}_{k-1}, \bar{b}, \bar{b}_{k+1}, \ldots \bar{b}_n]$$

where $\bar{b}_i$'s are column vectors of $\bar{\bar{A}}$

Use of LU factorization:

$$\bar{\bar{A}} = \bar{\bar{L}}\,\bar{\bar{U}}$$, where

IF $\bar{\bar{A}}$ is square and non-singular

$$\bar{\bar{L}} = \begin{bmatrix} 1 & & \bar{0} \\ & \ddots & \\ c_{m1} & & 1 \end{bmatrix} \qquad \bar{\bar{U}} = \begin{bmatrix} u_{11} & \cdots & u_{1n} \\ & \ddots & \vdots \\ \bar{0} & & u_{nn} \end{bmatrix}$$

lower triangular — upper triangular

Solution approaches (cont.)

The c_{ij}'s, $i>j$, are the multipliers in the Gauss elimination and u_{ij}'s, $j \geq i$, the system at the end of Gauss elimination.

(can be derived by multiplying out)

$$\bar{\bar{A}}\bar{x} = \bar{\bar{L}}(\bar{\bar{U}}\bar{x}) = \bar{b} \quad \text{OR} \quad \bar{\bar{L}}\bar{y} = \bar{b}, \; \bar{\bar{U}}\bar{x} = \bar{y}$$

thus, solution by simple backsubstitution

Solution approaches (cont.)

Row interchanges in $\bar{\bar{A}}$ might be needed to derive LU factorization. This solution approach is also known as Crout's and/or Doolittle's method.

If $\bar{\bar{A}}$ is symmetric we can choose $\bar{\bar{U}} = \bar{\bar{L}}^T$ and get what is known as Cholesky's method

$$\bar{\bar{A}} = \bar{\bar{L}}\,\bar{\bar{L}}^T \qquad (n \times n)$$

Numerical Gauss elimination

Numerical precision leads to a concern about scaling – we don't want to magnify round-off errors.

Hence, we always pick the pivot equation to be the one with the absolutely largest

pivot a_{jk} in column k (on or below the main diagonal) or, in fact the one for which

$|a_{jk}|/|A_j|$ is the largest; $|A_j|$ is the largest absolute value in row j ∇ partial pivoting

Iterative solutions (indirect)

We may rearrange the system so that $a_{ii} \neq 0$

$$\underset{(n \times n)}{\bar{\bar{A}}} = \bar{\bar{I}} + \bar{\bar{L}} + \bar{\bar{U}}$$

← contains upper triangle of $\bar{\bar{A}}$

contains lower triangle of $\bar{\bar{A}}$

after division by a_{ii} diagonals become 1

hence $\bar{\bar{A}}\bar{x} = (\bar{\bar{I}} + \bar{\bar{L}} + \bar{\bar{U}})\bar{x} = \bar{b}$

$$\Leftrightarrow \bar{x} = \bar{b} - \bar{\bar{L}}\bar{x} - \bar{\bar{U}}\bar{x}$$

OR

$$\underset{(n \times n)}{\bar{\bar{A}}} = \bar{\bar{I}} + (\bar{\bar{A}} - \bar{\bar{I}})$$

hence $\bar{\bar{A}}\bar{x} = (\bar{\bar{I}} - (\bar{\bar{A}} - \bar{\bar{I}}))\bar{x} = \bar{b}$

$$\Leftrightarrow \bar{x} = \bar{b} + (\bar{\bar{I}} - \bar{\bar{A}})\bar{x}$$

Iteration methods (sparse $\bar{\bar{A}}$)

If we start from $\bar{x}_0$ (a guess) and always update with the most recent values, x_i, we get:

Gauss-seidel

$$\bar{x}^{(m+1)} = \bar{b} - \underbrace{\bar{\bar{L}}}_{\text{"new" (lower)}} \bar{x}^{(m+1)} - \underbrace{\bar{\bar{U}}}_{\text{"old" (upper)}} \bar{x}^{(m)}$$

If instead we stick to the old values until all x_i's have been updated

Jacobi

$$\bar{x}^{(m+1)} = \bar{b} + (\bar{\bar{I}} - \bar{\bar{A}})\bar{x}^{(m)}$$

Ill-conditioning

We want the residual $\bar{b} - \bar{\bar{A}}\bar{x}$ to be small, and $\bar{0}$ if there is a solution. For this the system needs to be well-conditioned, which is the case when $\bar{\bar{A}}$ has large main diagonals. (in absolute sense) compared to other entries.

Ill-conditioning (cont.)

It can be "measured" by the condition number $\kappa(\bar{\bar{A}})$:

$$\kappa(\bar{\bar{A}}) = \|\bar{\bar{A}}\| \cdot \|\bar{\bar{A}}^{-1}\| \quad , \bar{\bar{A}} \; (n \times n)$$

singular ?

If small (< 100 and preferably << 100) the system is well-conditioned.

$\|\bar{\bar{A}}\|$ is some matrix norm of $\bar{\bar{A}}$

Matrix/Vector norms

Norm of vector $\bar{x}$: "generalized length"

$\|\bar{x}\| \geq 0$ and $\|\bar{x}\| = 0 \Leftrightarrow \bar{x} = \bar{0}$

A p-norm is :

$$\|\bar{x}\|_p = (|x_1|^p + |x_2|^p + \dots |x_n|^p)^{1/p}$$

e.g.

$$\|\bar{x}\|_1 = |x_1| + \dots |x_n| \quad l_1\text{-norm}$$

$$\|\bar{x}\|_2 = (|x_1|^2 + \dots |x_n|^2)^{1/2} \quad l_2\text{-norm (Euclidean)}$$

$$\|\bar{x}\|_\infty = \max_i |x_i| \quad l_\infty\text{-norm}$$

Matrix/vector norm (cont.)

Norm of a matrix is defined based on the vector norm of $\bar{\bar{A}}$ $(n \times n)$:

$$\|\bar{\bar{A}}\| = \max \frac{\|\bar{\bar{A}}\bar{x}\|}{\|\bar{x}\|}, \quad \bar{x} \neq \bar{0}$$

If:

l_1 : $\|\bar{\bar{A}}\|_1 = \max\limits_{k} \sum\limits_{j=1}^{n} |a_{jk}|$ "column sum"

l_∞ : $\|\bar{\bar{A}}\|_\infty = \max\limits_{j} \sum\limits_{k=1}^{n} |a_{jk}|$ "row sum"

l_2 : $\|\bar{\bar{A}}\|_2 = ?$

Vector space V

An algebraic structure with vectors

$\bar{v}_i \in R^n$ (or C^n) such that: R : real
C : complex

if $\bar{v}_i \in V$, $i = 1..k$ then

$$c_1 \bar{v}_1 + \dots c_k \bar{v}_k \in V, \quad c_i \text{ scalars}$$

Thus, two algebraic operations defined:

- vector addition
- scalar multiplication

Vector space V (cont.)

$$\bar{v}_1 + \bar{v}_2 = \bar{v}_2 + \bar{v}_1$$ commutative

$$\bar{v}_1 + (\bar{v}_2 + \bar{v}_3) = (\bar{v}_1 + \bar{v}_2) + \bar{v}_3$$ associative

$$\bar{v} + \bar{0} = \bar{v}$$ zero vector

$$\bar{v} + (-\bar{v}) = \bar{0}$$ anti-element

$$c(\bar{v}_1 + \bar{v}_2) = c\bar{v}_1 + c \cdot \bar{v}_2$$ distributive

$$(c+k)\bar{v} = c\bar{v} + k\bar{v}$$ (wrt. add./mult.)

$$c(k\bar{v}) = (ck)\bar{v}$$ associative

$$1\bar{v} = \bar{v}$$ identity

Example signal space

Eksempel Et (diskret) signal er en “dobbelt uendelig” følge af reelle tal

$$\bar{x} = \{x_k\}_{k\in\mathbb{Z}} = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$$

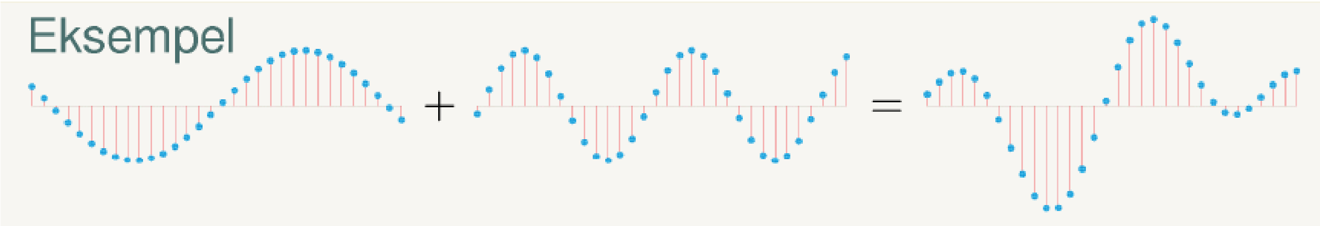
Signaler kan lægges sammen:

$$\{x_k\} + \{y_k\} = \{x_k + y_k\} = (\ldots, x_{-1} + y_{-1}, x_0 + y_0, x_1 + y_1, \ldots)$$

og ganges med skalarer (reelle tal):

$$c\{x_k\} = \{c\, x_k\} = (\ldots, cx_{-1}, cx_0, cx_1, \ldots)$$

Mængden $\mathbb{S}$ af signaler er et vektorrum.



Vector sub-space

A vector space $U \subset V$, $U \neq \emptyset$

which is closed wrt. the two algebraic operations, thus

$$c_1 \bar{u}_1 + \dots c_k \bar{u}_k \in U \quad \text{if } \bar{u}_i \in U$$

and which contains $\bar{0}$ (obviously)

Examples $R^n, C^n, \bar{A}_{(2\times 2)}, R^n$ with constraints

Vector space and basis

Dimension of V ($\dim(V) = D$)

the maximum number D of linearly independent vectors in V

Basis of V ($\{\bar{v}_1, \bar{v}_2, \ldots \bar{v}_D\} = B$)

the maximum linearly independent set B of vectors in V

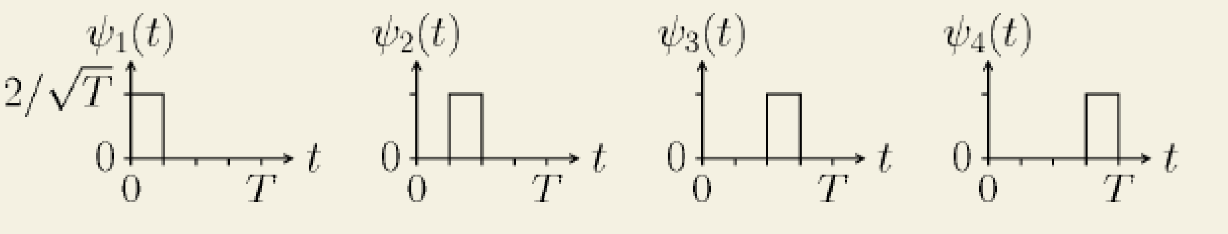
B spans V

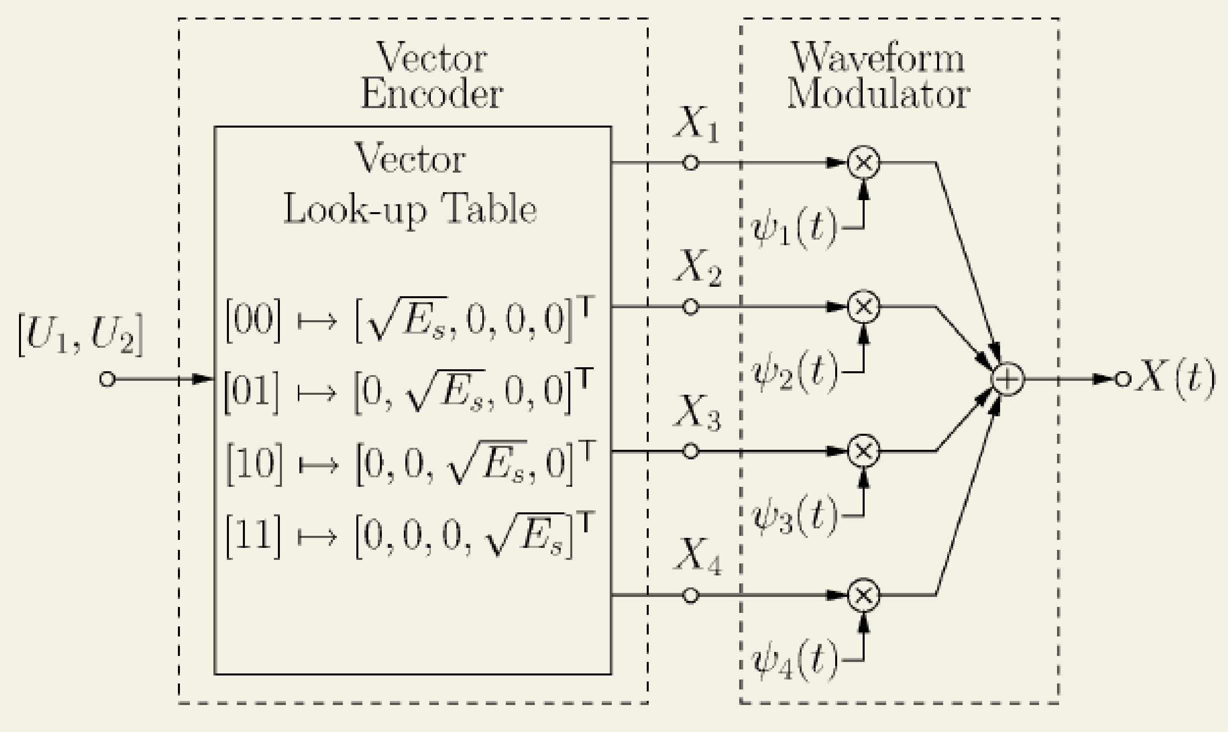
Span of vectors

all linear combinations of the vectors

Example signal space w. basis

The four orthonormal basis functions are





Row and column space

The column and row vectors of $\bar{\bar{A}}$ span a vector space of dimension $r = \text{rank}(\bar{\bar{A}})$ — as per def. of rank

Row space: a set of r linearly independent row vectors

Column space: a set of r linearly independent column vectors

The essential vectors !

Row and Colomn space (cont.)

We use Gauss elimination to find the row and colomn space, since elementary row operations:

- do not change dependency between columns
- row equivalent matrices have the same rank

! can change row dependency (use A^T)

Inner product space

A vector space V where to each pair of vectors we define the inner product $(\bar{v}_1, \bar{v}_2)$

! different notations use $<\bar{v}_1, \bar{v}_2>$ or $\bar{v}_1 \cdot \bar{v}_2$ (dot)

so that we can measure "length"

$R^n : (\bar{v}_1, \bar{v}_2) = \bar{v}_1^T \bar{v}_2$ $\quad C^n : \bar{v}_1^{*T} \bar{v}_2$!

Inner product space (cont.)

Linearity

$$(c_1\bar{v}_1 + c_2\bar{v}_2, \bar{v}_3) = c_1(\bar{v}_1, \bar{v}_3) + c_2(\bar{v}_2, \bar{v}_3)$$

$$(\bar{v}_1, \bar{v}_2) = (\bar{v}_2, \bar{v}_1)$$

symmetry

$$(\bar{v}, \bar{v}) \geq 0$$

$$(\bar{v}, \bar{v}) = 0 \Leftrightarrow \bar{v} = \bar{0}$$

positive definiteness

$$(\bar{v}_1, \bar{v}_2) = 0 \Leftrightarrow \bar{v}_1 \perp \bar{v}_2$$

orthogonality

$$\|\bar{v}\|_2 = \sqrt{(\bar{v}, \bar{v})}$$

length