Adaptive and Array Signal Processing - Lecture Notes

Christian, Michael, Steffi

February 22, 2019

Contents

1 Motivation

We may classify filters as linear or non-linear. A filter is said to be *linear* if the filtered, smoothed, or predicted quantity at the output of the filter is a *linear function of the observations applied to the filter input.* Otherwise, the filter is nonlinear.

1.1 Linear Filter

C_2 in graphic 1 can be tuned $\rightarrow$ programmable filter (due to changeable parameters). Note that this is not an Adaptive Filter yet, just a programmable one.
E.g. Bandpass with changeable frequency but same bandwidth (you have to change all: C, R and L!).

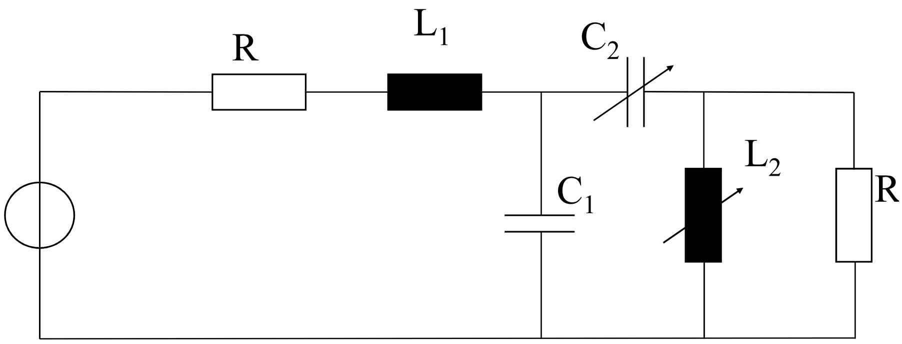

Figure 1: Linear Filter, fixed parameters

1.2 Discrete Filter

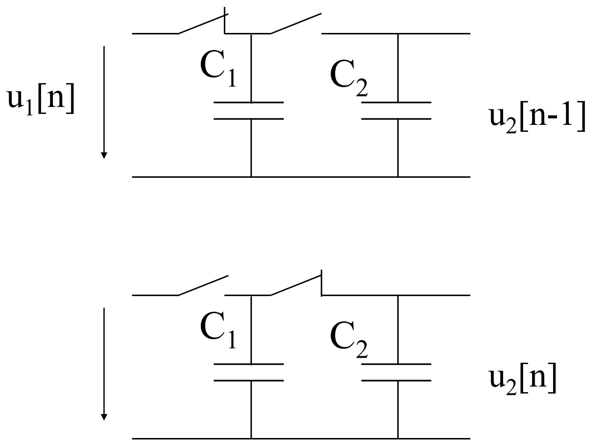

Figure 2: One open, one closed, discrete time filter

1.2.1 IIR-Filter

Infinite impulse response (IIR) is a property applying to many linear time-invariant systems. Common examples of linear time-invariant systems are most electronic and digital filters. Systems with this property are known as IIR systems or IIR filters, and are distinguished by having

Table 1: discrete filter - Impulse Response

n	≤ 0	0	1	2	...	k
$u_1[n]$	0	1	0	0	...	0
$u_2[n]$	0	$1-a$	$a(1-a)$	$a^2(1a)$	...	$a^k(1-a)$

an impulse response which does not become exactly zero past a certain point, but continues indefinitely. This is in contrast to a finite impulse response (FIR) in which the impulse response h(t) does become exactly zero at times $t > T$ for some finite T, thus being of finite duration. The discrete filter in picture 2 is defined by:

$$C_1 \cdot u_1[n] + C_2 \cdot u_2[n-1] = (C_1 + C_2)u_2[n]$$

$$u_2[n] = \underbrace{\frac{C_2}{C_1 + C_2}}_{a} \cdot u_2[n-1] + \underbrace{\frac{C_1}{C_1 + C_2}}_{1-a} \cdot u_1[n]$$

It's impulse response is given by table 1. As can be seen coefficient a allows us to change the filter's impulse response and therefore to program the filter. Figure 3 shows a block diagram for such an easy, programmable discrete filter. From this block diagram it can be seen immediately that this filter is an IIR-Filter and can therefor be unstable.

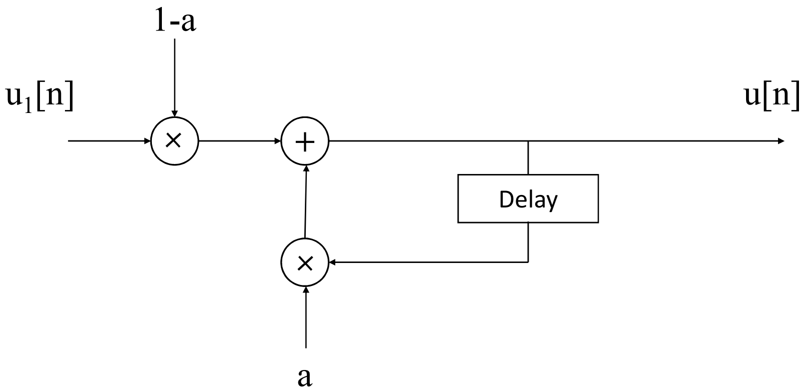

Figure 3: Discrete, easy to program filter

Caution! Proof of stability required!

$|a| \geq 1 \rightarrow$ instable $\quad \Rightarrow \quad |a| \leq 1 \rightarrow$ stable.

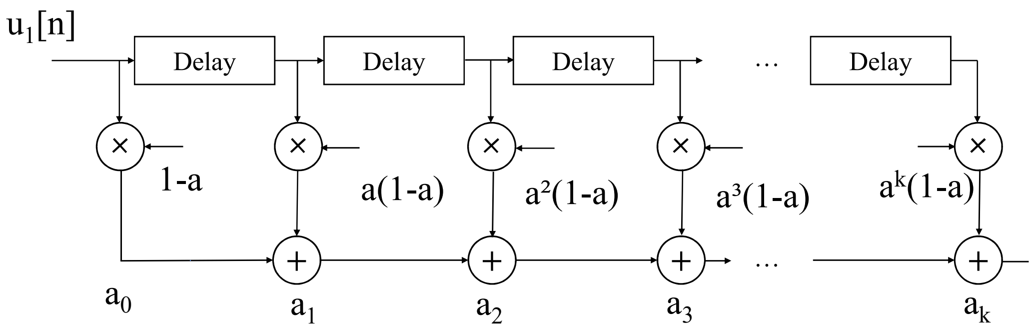

Figure 4: Linear adaptive filter (tapped delay line)

1.2.2 FIR-Filter

To overcome the stability problems an FIR-filter can be used instead. FIR-Filters are always stable. Suitable for adaptive filters but it's an approximation $\Rightarrow$ finite
Structure: tapped delay line
linear: neither delay only coefficient are dependent in input signal. $\rightarrow$ linear adaptive filter, see picture 4.

Delays of 1 clock cycle:
$D \cdot x[n] = x[n-1]$

Delay of 2 clock cycles:
$\underbrace{DD}_{D^2} \cdot x[n] = D \cdot x[n-1] = x[n-2] = D^2 \cdot x[n]$

Fractional delay:
$F \cdot x(n) = x(n - \frac{1}{2})$
$FF \cdot x(n) = F \cdot x(n - \frac{1}{2}) = x[n-1] = D \cdot x[n]$
$F^2 = FF \equiv D; F = \sqrt{D} = D^{\frac{1}{2}}$
$D^{\frac{1}{2}} \cdot D^{\frac{1}{2}} = D^{\frac{1}{2}+\frac{1}{2}} = D' = D$
Developed using a Taylor series: $D^{\frac{1}{2}} \approx 1 + \frac{1}{2}(D-1) + \frac{1}{8}(D-1)^2 = \frac{3}{8} + \frac{3}{4} \cdot D + \frac{1}{8}D^2$
The filter in graphic 5 shows the shifting of the signal by $\frac{1}{2}$.

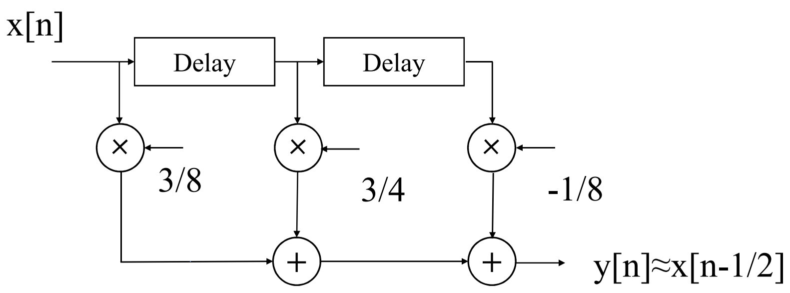

Figure 5: Used for synchronization of receiver and transmitter

A fractional delay $y[n] \approx u[n-a]$ for $0 \leq a \leq 1$
$y[n] = a_0 \cdot u[n] + a_1 \cdot u[n-1] + a_2 \cdot u[n-2]$ with $a_0 = \frac{1}{2}(a^2+3a+2)$ and $a_1 = (2-a)a$ and $a_2 = \frac{1}{2}(a-1)a$

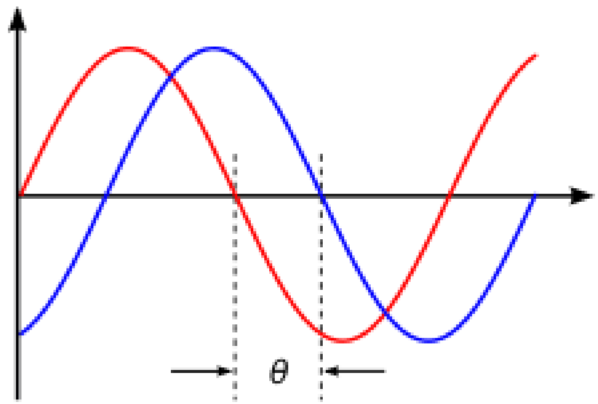

Figure 6: Phase-shift for θ

Example: $a = \frac{1}{2} \quad u[n] = sin(\frac{n}{2})$
A node is passing

$$S = \sum_{i=1}^{2} |a_i|^2 = \frac{1}{2}a^4 - 6a^3 + \frac{5}{2}a^2 - 3a + 1$$

remains $\leqq 1$ only for $0 \leq a \leq 2$. For $a < 0$ or $a > 2$ it grows unboundenly.
This is relevant, since uncorrelated noise at the input leads to an output which mean square is proportional to S. Thus, $a < 2$ or $a < 0$ are possible, but with a (huge) noise penalty. The mean squared output noise is minimum for $a = 1 - \frac{1}{\sqrt{2}} \approx 0,3$

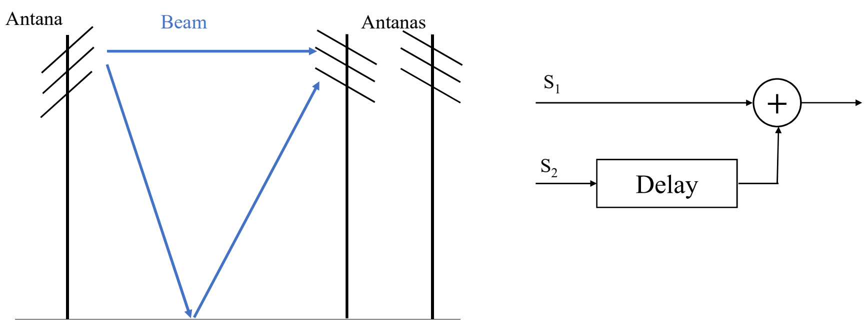

Figure 7: 2 signals, but delayed.

The sketch in figure 7 shows a signal which travels trough two paths from the transmitter to the receiver. One signal is faster (has a shorter way) than the other. To be able to add the two signals the faster one has to be delayed and since the delay between the two signals is (usually) not an integer the fractional delay is needed here.
⇒ After delaying the faster one, they're only different in their amplitude and can be added. (Otherwise we get fading effects.)

1.3 Principle Structure of an Adaptive Filter

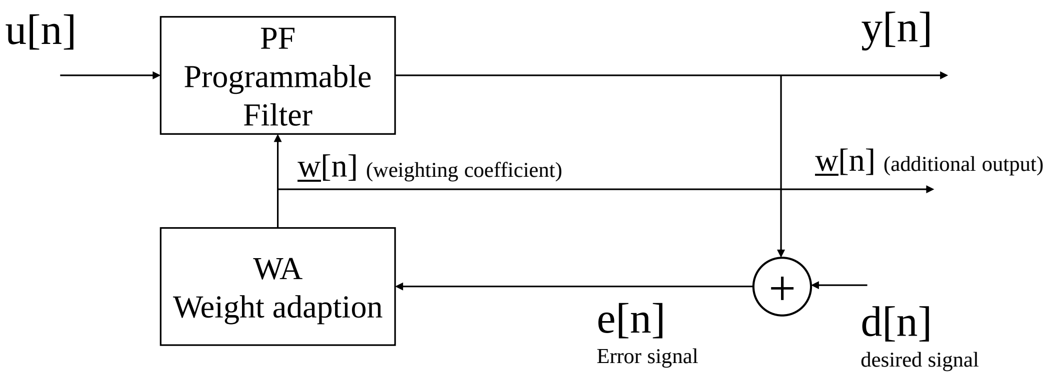

Figure 8: Principle structure of an Adaptive Filter

$e[n] = 0 \quad \Rightarrow \quad \underline{w}[n+1] = \underline{w}[n]$

1.3.1 The 4 Application Classes

System Identification

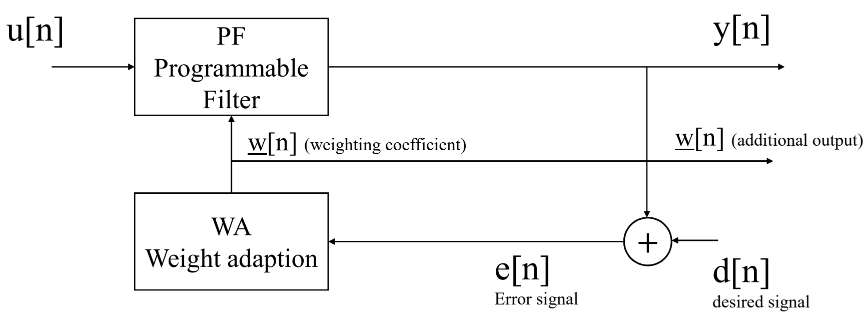

Figure 9: System identification

Channel Estimation

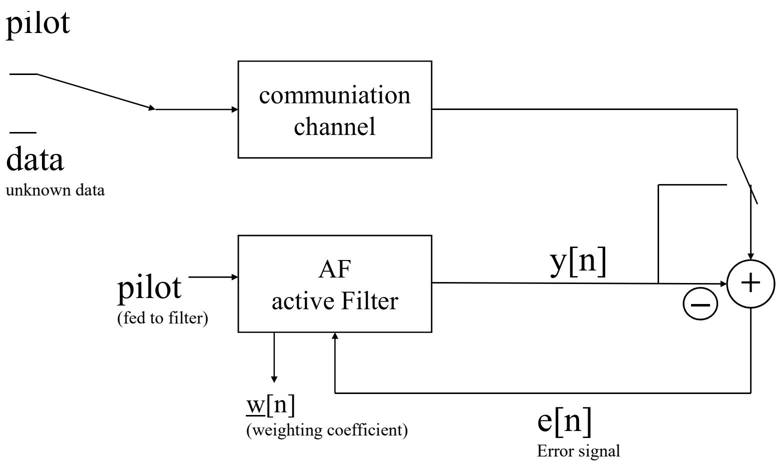

Figure 10: Channel Estimation

The Pilot is a known data string. Either the transmitter and the receiver know the pilot. The data is unknown.

Inverse Modelling

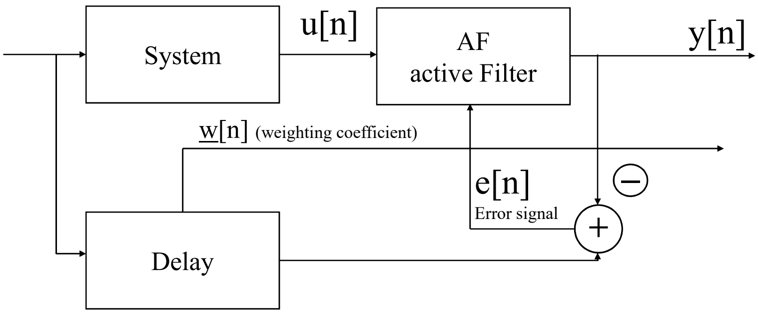

Figure 11: Inverse Modelling

Example to picture 11: Channel equilization (interested in data, not the channel).

Prediction

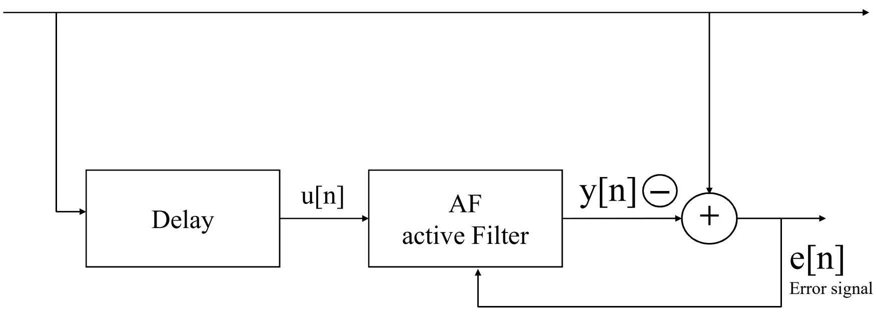

Figure 12: Prediction

Notes to 12: Blind to future and present, but knows the past.. Trouble: if quantized e[n] won't work!
Therefore:

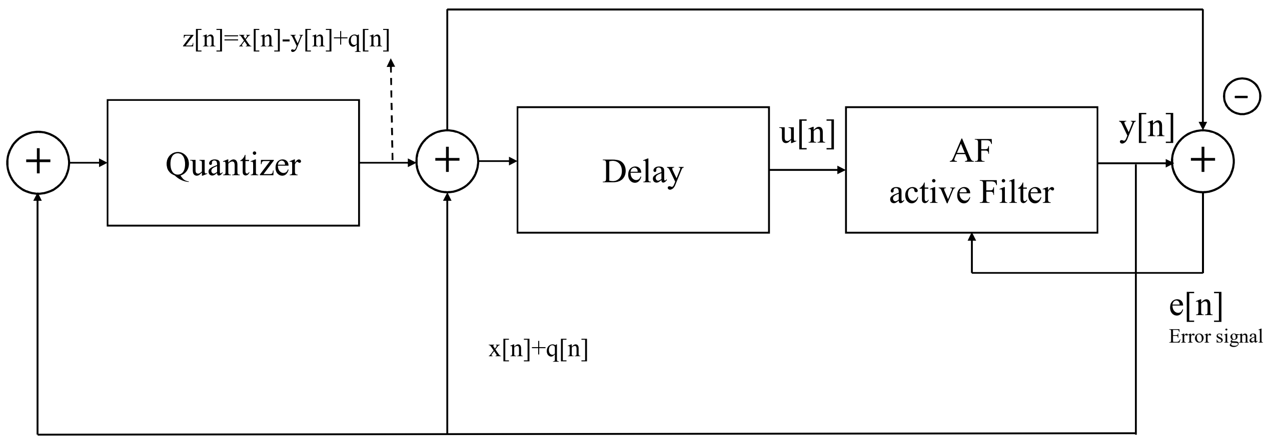

Figure 13: Compressor

Predictive filter works well $\Rightarrow$ if error is small

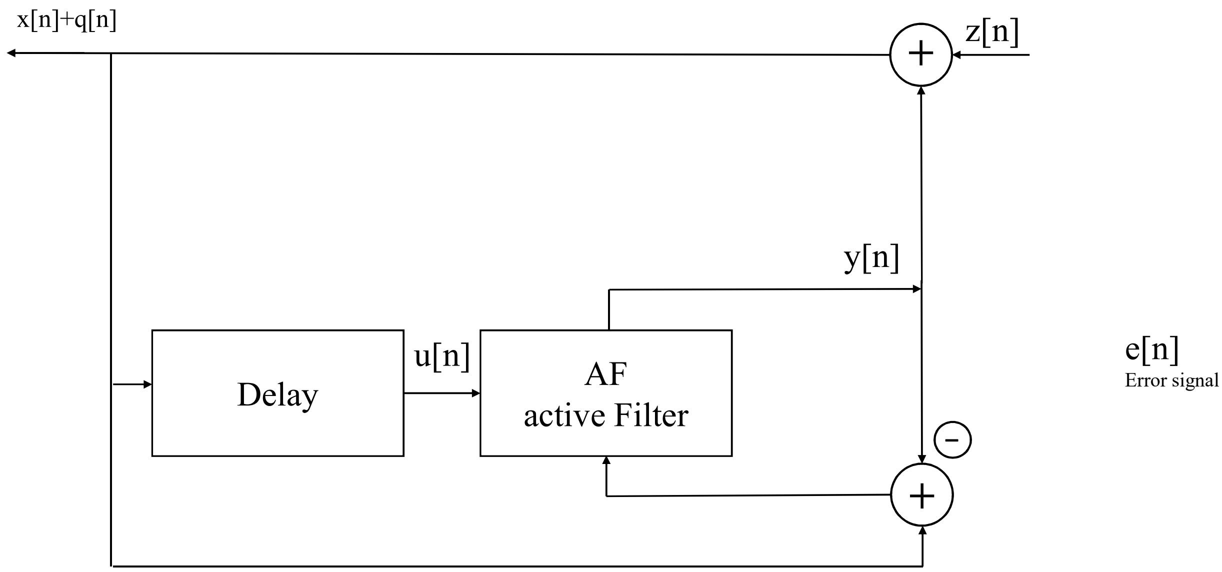

Figure 14: Decompressor

Interference Cancellation

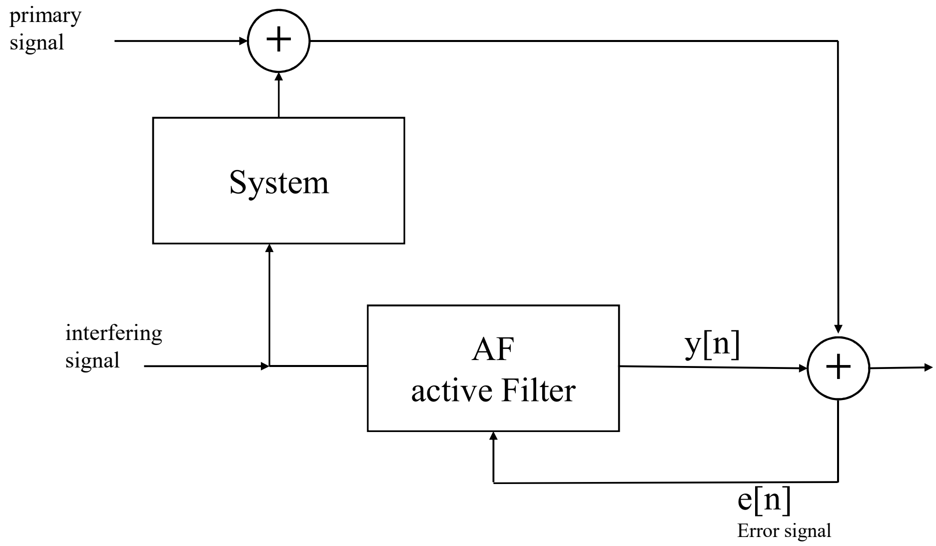

Figure 15: Interference Cancellation

Assumption: The desired (primary) signal is statistically independent of the interfering signal. For linear adaptive filters: uncorrelated is enough.

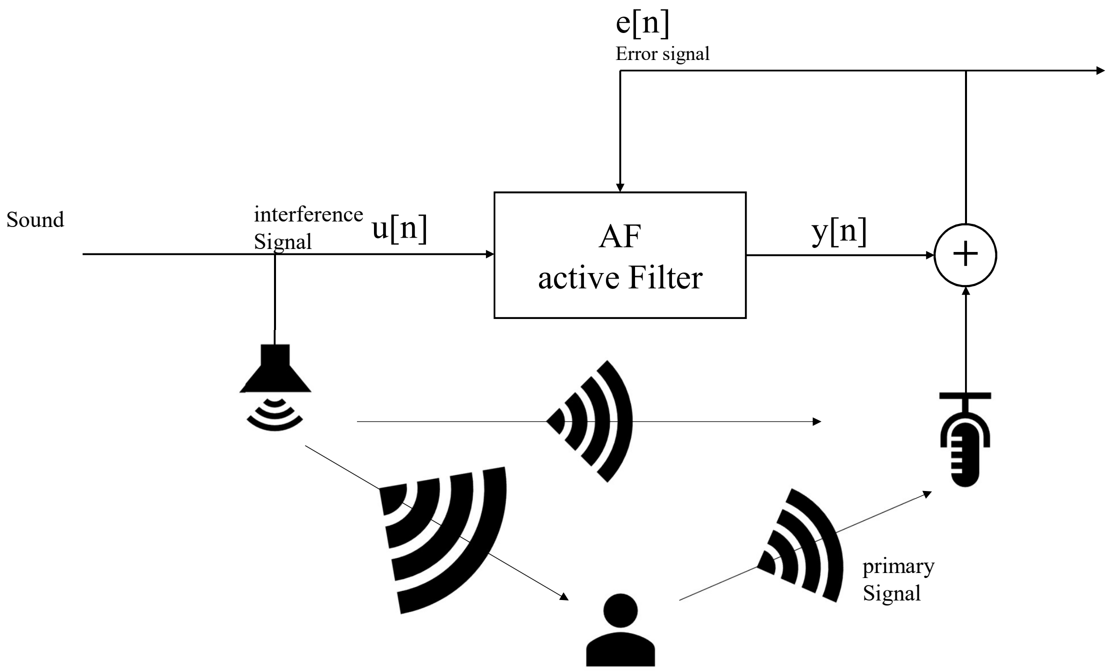

Figure 16: Example for an application for a an interference-cancelling-filter

For 16: Linear filters are better than nonlinear ones for this case. Filter tries to "'kill"' the interference because it knows, what has been played on the speaker.

1.3.2 Complex Envelope

Definition:
$x(t) = \mathrm{Re}\left\{u(t)e^{\mathrm{j}2\pi f_0 t}\right\}$
with

- $x(t) \in \mathbb{R}$: Real Signal
- $u(t) \in \mathbb{R}$: Complex Envelope of $x(t)$
- f_0: Center of Frequency; $[f_0] = 1Hz$

Notation with Magnitude and phase:
$u(t) = |u(t)| \cdot e^{j\varphi_u(t)}$
Complex Envelope $u(t)$ mapped on signal $x(t)$
$x(t) = \mathrm{Re}\left\{u(t)e^{\mathrm{j}2\pi f_0 t}\right\} = |u(t)| \cdot \cos(2\pi f_0 t + \varphi_u(t)) = \frac{1}{2}u(t) \cdot e^{\mathrm{j}2\pi f_0 t} + \frac{1}{2}u^*(t) \cdot e^{-\mathrm{j}2\pi f_0 t}$
with $\mathrm{Re}\{z\} = \frac{z+z^*}{2}$
approach to map signal $x(t)$ on complex envelope $u(t) \stackrel{?}{=} 2x(t) \cdot e^{-j2\pi ft}$
$2x(t) \cdot e^{-j2\pi ft} = 2\,\mathrm{Re}\left\{u(t)e^{\mathrm{j}2\pi f_0 t}\right\} \cdot e^{-j2\pi ft} = (u(t)e^{\mathrm{j}2\pi f_0 t} + u^*(t)e^{-\mathrm{j}2\pi f_0 t}) \cdot e^{-j2\pi ft}$
$2x(t) \cdot e^{-j2\pi ft} = u(t) + u^*(t) \cdot e^{-j4\pi f_0 t}$
Check approach with the Fourier transform:
$2x(t) \cdot e^{-j2\pi ft} \quad ○\!\!-\!\!\!-\!\!\bullet \quad U(f) + \int\limits_{-\infty}^{\infty} u^*(t) \cdot e^{-\mathrm{j}4\pi f_0 t} e^{\mathrm{j}2\pi ft} dt$

$$\text{with} \quad \int\limits_{-\infty}^{\infty} u^*(t) \cdot e^{-\mathrm{j}4\pi f_0 t} e^{\mathrm{j}2\pi ft} dt = \int\limits_{-\infty}^{\infty} u^*(t) \cdot e^{-\mathrm{j}2\pi(f+2f_0)t} dt = \left(\int\limits_{-\infty}^{\infty} u(t) \cdot e^{-\mathrm{j}2\pi(f+2f_0)t}\right)^*$$

$$= \left(\left.\underbrace{\int\limits_{-\infty}^{\infty} u(t) \cdot e^{-\mathrm{j}2\pi(f+2f_0)t}}_{U(f)}\right|_{-f-2f_0 \to f}\right)^* = U^*(-f - 2f_0)$$

$\Rightarrow \quad 2x(t) \cdot e^{-j2\pi ft} \quad ○\!\!-\!\!\!-\!\!\bullet \quad U(f) + U^*(-f - 2f_0)$

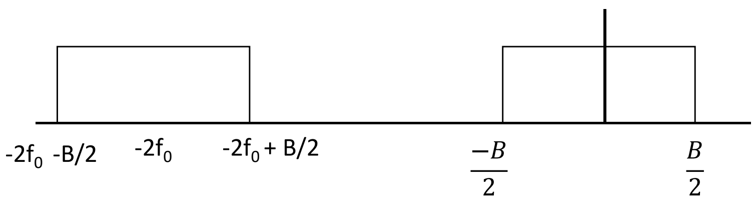

Figure 17: Result of the Fourier Transform

The result shows an additional shifted envelope $\Rightarrow$ therefore a low pass filter (LPF) is needed to suppress overlapping (aliasing). $\Rightarrow$ band limitation
$\Rightarrow \quad u(t) = LPF(x(t) \cdot 2 \cdot e^{-j2\pi ft})$
$x(t) \iff u(t)$ [It's not a Fourier transform, but related.]

Constraints for Bandwidth:
Out of Picture 17:
$-2f_0 + \frac{B}{2} \leq -\frac{B}{2}$

$$\boxed{\Rightarrow \quad B \leq 2f_0 \qquad \text{or} \qquad f_0 \geq \frac{B}{2}}$$

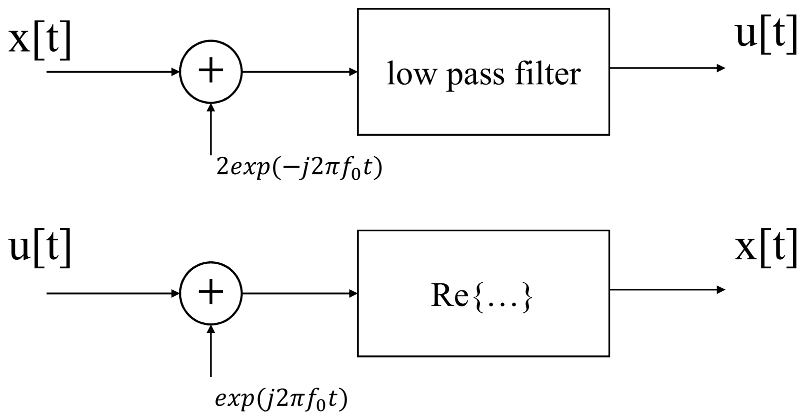

Figure 18: Transformation from complex envelope and signal as block diagram and vice versa

$x(t) \iff u(t)$ [It's not a fourier transform, but related.]

1.3.3 Whatever we did here - No. 264

1.3.3.1 Ivrlač the dimensions Conqueror

$$\sum_m \underbrace{a_m}_{\in\mathbb{R}} x(t-t_m) \iff \sum_m \underbrace{a_m \cdot e^{-i2\pi f_0 t_m}}_{b_m \in \mathbb{C}} u(t-t_m) = \sum_m |b_m| \cdot e^{j\varphi_m} u(t-t_m)$$

Under condition of:$|b_m| = a_m \quad \varphi_{b_m} = -2\pi f_0 t_m$
Note: Correspondence is linear; center frequency has to be known.

$u(t) \iff x(t) = \mathrm{Re}\left\{u(t)e^{j2\pi f_0 t}\right\}$

1.3.3.2 Approach - Complex Envelope shifted and scaled to fit a real signal

$w^* u(t-\tau)$ `corresponds to` $c \cdot x(t-\widetilde{t}) = x'(t)$ with $c \in \mathbb{R}$

$\underbrace{w^*}_{const.} \cdot u(t-\tau)$ `corresponds to` $x'(t) = \mathrm{Re}\left\{w^* \cdot u(t-\tau)e^{j2\pi f_0 t}\right\}$,
with $\quad u(t) = |u(t)|e^{j2\pi f_0 (t-\tau)} \quad w = |w|e^{j\varphi_w}$

$c \cdot x(t-\tau) = \mathrm{Re}\left\{w^* u(t - \tau e^{j2\pi f_0 t})\right\}$

$c \cdot \mathrm{Re}\left\{u(t-\widetilde{t}) \cdot e^{j2\pi f_0 (t-\widetilde{t})}\right\} = \mathrm{Re}\left\{w^* \cdot u(t-\tau) \cdot e^{j2\pi f_0 t}\right\}$

$c \cdot |u(t-\widetilde{t})| \cos(2\pi f_0 t - 2\pi f_0 \widetilde{t} + \varphi_u(t-\widetilde{t})) = |w||u(t-\tau)| \cos(2\pi f_0 t + 2\pi\varphi_u(t-\tau) - \varphi_w)$

$\Rightarrow$ cos(arg) should be equal!

$2\pi f_0 t - 2\pi f_0 \widetilde{t} + \varphi_u(t-\widetilde{t}) = 2\pi f_0 t + 2\pi\varphi_u(t-\tau) - \varphi_w$

$2\pi f_0 t - 2\pi f_0 \widetilde{t} + \varphi_u(t-\widetilde{t}) = 2\pi f_0 t + 2\pi\varphi_u(t-\tau) - \varphi_w + q \cdot 2\pi \quad \text{with} \quad q \in \mathbb{Z}$

$f_0 \tilde{t} = \frac{\varphi_u(t-\tilde{t}) - \varphi_u(t-\tau) + \varphi_w}{2\pi} - q$

Non-linear in $\tilde{t}$

Substitution: $\tilde{t} = \tau + \Delta t$

$f_0 \Delta t = \frac{\varphi_u(t-\widetilde{t}-\Delta t) - \varphi_u(t-\tau) + \varphi_w}{2\pi} - f_0 \tau - q$

Choose q:

$0 \leq f_0 \Delta t < 1$

Assumption: u(t) is **narrow band**

$u(t-t') \approx u(t) \quad \text{with} \quad |t'| \leq \frac{1}{f_0}$

$f_0 \Delta t = (\frac{\varphi_w}{2\pi} - f_0 \tau) mod(1)$

$\Delta t = (\frac{\varphi_w}{2\pi f_0} - \tau) mod(\frac{1}{f_0})$

$\tilde{t} = (\frac{\varphi_w}{2\pi f_0} - \tau) mod(\frac{1}{f_0}) + \tau$

$c = |w|$ Complex envelope only for narrow band!

1.3.3.3 Power computation

$x(t) = \frac{1}{2}u(t) \cdot e^{\mathrm{j}2\pi f_0 t} + \frac{1}{2}u^*(t) \cdot e^{-\mathrm{j}2\pi f_0 t} \quad |(\bullet)^2$

$x(t)^2 = (\frac{1}{2}u(t) \cdot e^{\mathrm{j}2\pi f_0 t} + \frac{1}{2}u^*(t) \cdot e^{-\mathrm{j}2\pi f_0 t})(\frac{1}{2}u(t) \cdot e^{\mathrm{j}2\pi f_0 t} + \frac{1}{2}u^*(t) \cdot e^{-\mathrm{j}2\pi f_0 t})$

$x(t)^2 = \frac{1}{4}u^2(t) \cdot e^{\mathrm{j}4\pi f_0 t} + \frac{1}{2}u(t) \cdot u^*(t) + \frac{1}{4}u^{2*}(t) \cdot e^{-\mathrm{j}4\pi f_0 t}$

with $\frac{1}{2}u(t) \cdot u^*(t) = \frac{1}{2}|u(t)|^2$

with $\frac{1}{4}u^2(t) \cdot e^{\mathrm{j}4\pi f_0 t} + \frac{1}{4}u^{2*}(t) \cdot e^{-\mathrm{j}4\pi f_0 t} = \frac{1}{4}u^2(t) \cdot e^{\mathrm{j}4\pi f_0 t} + (\frac{1}{4}u^2(t) \cdot e^{\mathrm{j}4\pi f_0 t})^* = \frac{1}{2}\mathrm{Re}\left\{u^2(t)e^{j4\pi f_0 t}\right\}$

Square of Magnitude

$|x(t)|^2 = \frac{1}{2}|u(t)|^2 + \frac{1}{2}\mathrm{Re}\left\{u^2(t)e^{j4\pi f_0 t}\right\}$

Get Time average with expectation:

Restrict u(t) to be proper:

$E[u^2(t)] = 0$

$E[|x(t)|^2] = \frac{1}{2}E[|u(t)|^2]$

$u(t) = a(t) + jb(t)$

$u^2(t) = a^2(t) + 2ja(t)b(t) - b^2(t)$

$E[u^2] = \underbrace{E[a^2(t) - b^2(t)]}_{=0} + \underbrace{2jE[a(t)b(t)]}_{=0} \stackrel{!}{=} 0$ (widely linear processing)

Mean Time square average for narrow band: (equals power)

$\overline{x^2(t)} = f_0 \cdot \int\limits_{t-\frac{1}{2f_0}}^{t+\frac{1}{2f_0}} x^2(\tau)d\tau \approx \frac{1}{2}|u(t)|^2$

time average x^2 is a good assumption. $\Rightarrow$ We will stick with narrow band assumption

$\Rightarrow$ both will work.

1.3.4 FIR-Filter

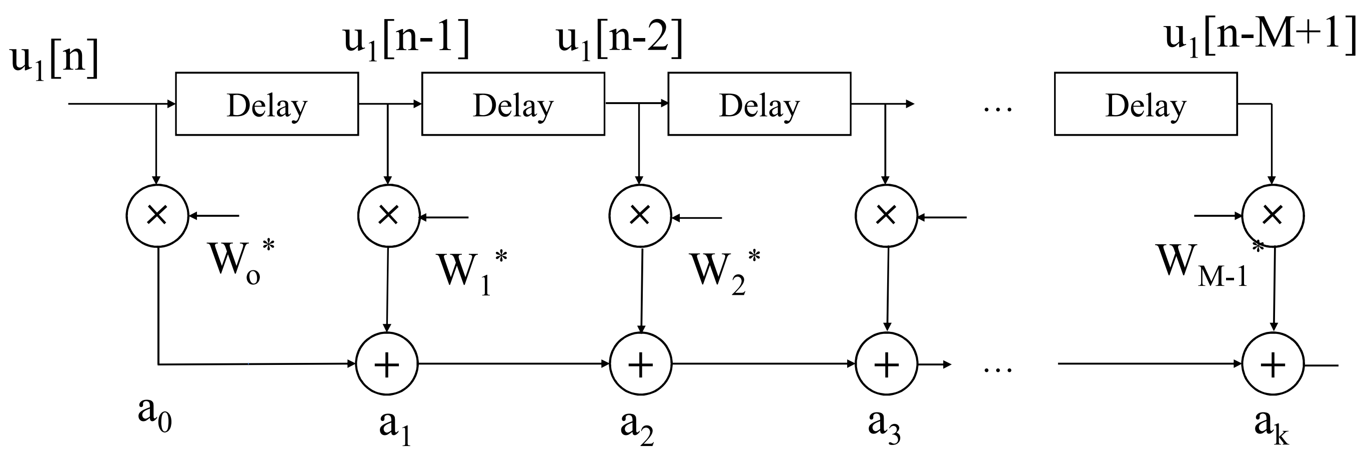

Figure 19: stable, finite response

$$\underline{\boldsymbol{w}} = \begin{bmatrix} w_0 \\ w_1 \\ . \\ . \\ . \\ w_{M-1} \end{bmatrix} \in \mathbb{C}^{M\times 1}; \qquad \underline{\boldsymbol{w}}^* = \begin{bmatrix} w_0^* \\ w_1^* \\ . \\ . \\ . \\ w_{M-1}^* \end{bmatrix}$$

$$\underline{\boldsymbol{w}}^T = \begin{bmatrix} w_0 & w_1 & ... & w_{M-1} \end{bmatrix}$$

$(\underline{\boldsymbol{w}}^T)^* = (\underline{\boldsymbol{w}}^*)^T = \underline{\boldsymbol{w}}^H$

$$\underline{\boldsymbol{u}}[n] = \begin{bmatrix} u[n] \\ u[n-1] \\ . \\ . \\ . \\ u[n-M+1] \end{bmatrix} \in \mathbb{C}^{M\times 1}$$

$$y[n] = \underline{\boldsymbol{w}}^H \cdot \underline{\boldsymbol{u}}[n] = u[n]w_0^* + u[n-1]w_1^* + ... + u[n-M+1]w_{M-1}^*$$

$$y^*[n] = (\underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n])^* = ((\underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n])^T)^* = (\underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n])^H = \underline{\boldsymbol{u}}^H[n]\underline{\boldsymbol{w}}$$

$$\underline{\boldsymbol{y}}[n] = \begin{bmatrix} \underline{\boldsymbol{y}}[n] \\ \underline{\boldsymbol{y}}[n+1] \\ . \\ . \\ . \\ \underline{\boldsymbol{y}}[n+N] \end{bmatrix} \in \mathbb{C}^{(N+1)\times 1}$$

$$\underline{\boldsymbol{d}}[n] = \begin{bmatrix} d[n] \\ d[n+1] \\ . \\ . \\ . \\ [n+N] \end{bmatrix} \in \mathbb{C}^{(N+1)\times 1} \qquad \underline{\boldsymbol{d}} \text{ is the desired signal}$$

error: $\underline{\boldsymbol{e}}[n] = \underline{\boldsymbol{d}}[n] - \underline{\boldsymbol{y}}[n]$

$$\underline{\boldsymbol{w}}_{MMSE} = arg \min_{\underline{\boldsymbol{w}}} \underbrace{\sum_{k=0}^{N} |e[n+k]|^2}_{\underline{\boldsymbol{e}}^H[n]\underline{\boldsymbol{e}}[n]=||\underline{\boldsymbol{e}}[n]||_2^2 \quad \text{Squared Euclidean Norm}}$$

Note: MMSE means Minimum Mean Square Error $\qquad \underline{\boldsymbol{w}}$ is the weighting coefficient

$$\underline{\boldsymbol{w}}_{LS} = arg \min_{\underline{\boldsymbol{w}}} E[||\underline{\boldsymbol{e}}[n]||_2^2] \qquad \text{Note: LS means Least Square Error}$$

$$\underline{\boldsymbol{w}}_{OPT} = arg \min_{\underline{\boldsymbol{w}}} E[||\underline{\boldsymbol{e}}[n]||_2^2], \quad \text{such that} \quad \begin{cases} \underline{\boldsymbol{w}}_1^H \underline{\boldsymbol{a}}_i = b_1 \\ i \in \{1, 2, ..., k\} \end{cases} \quad \text{Constraints}$$

OPT means Additional Constraints, Optimization.

$$\underline{\boldsymbol{e}}^*[n] = \underline{\boldsymbol{d}}^*[n] - \underline{\boldsymbol{y}}^*[n] = \underline{\boldsymbol{d}}^*[n] - \underbrace{\begin{bmatrix} \underline{\boldsymbol{u}}^H[n] \\ \underline{\boldsymbol{u}}^H[n+1] \\ : \\ \underline{\boldsymbol{u}}^H[n+N] \end{bmatrix}}_{\text{Signal Matrix } \underset{\sim}{\boldsymbol{U}}^H[n]} \underline{\boldsymbol{w}} = \underline{\boldsymbol{d}}^*[n] - \underset{\sim}{\boldsymbol{U}}^H[n]\underline{\boldsymbol{w}}$$

$$\begin{aligned}
||\underline{\boldsymbol{e}}[n]||_2^2 = \quad \underline{\boldsymbol{e}}^H[n]\underline{\boldsymbol{e}}[n] \quad &= (\underline{\boldsymbol{e}}^H[n]\underline{\boldsymbol{e}}[n])^* \ \big(\text{Note: It's a real number, so complex conjungation doesn't hurt!}\big) \\
&= (\underline{\boldsymbol{e}}^*[n])^H \underline{\boldsymbol{e}}^*[n] = (\underline{\boldsymbol{d}}^*[n] + \underset{\sim}{\boldsymbol{U}}^H \underline{\boldsymbol{w}})^H (\underline{\boldsymbol{d}}^*[n] - \underset{\sim}{\boldsymbol{U}}^H[n]\underline{\boldsymbol{w}}) \\
&= ((\underline{\boldsymbol{d}}^*[n]^H - \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{U}}[n])(\underline{\boldsymbol{d}}^*[n] - \underset{\sim}{\boldsymbol{U}}^H[n]\underline{\boldsymbol{w}})) \\
&= ||\underline{\boldsymbol{d}}^*[n]||_2^2 - \underbrace{(\underline{\boldsymbol{d}}^*[n])^H \underset{\sim}{\boldsymbol{U}}^H[n]}_{\underline{\boldsymbol{\rho^H}}} \underline{\boldsymbol{w}} - \underline{\boldsymbol{w}}^H \underbrace{\underset{\sim}{\boldsymbol{U}}[n]\underline{\boldsymbol{d}}^*[n]}_{\underline{\boldsymbol{\rho}}} + \underline{\boldsymbol{w}} \underbrace{\underset{\sim}{\boldsymbol{U}}[n]\underset{\sim}{\boldsymbol{U}}^H[n]}_{\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H} \underline{\boldsymbol{w}}
\end{aligned}$$

Minimization of quadratic forms

$$||\underline{\boldsymbol{e}}[n]||_2^2 = \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}} - \underline{\boldsymbol{\rho}}^H \underline{\boldsymbol{w}} - \underline{\boldsymbol{w}}^H \underline{\boldsymbol{\rho}} + ||\underline{\boldsymbol{d}}[n]||_2^2 \qquad \in \mathbb{R}$$

1.3.5 Spatial Filtering

1.3.5.1 Planar wave

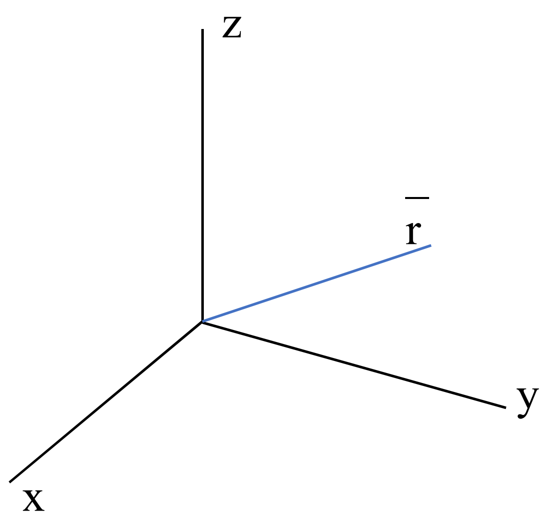

Figure 20: Direction of planar wave

$\underbrace{q(t, \overrightarrow{\boldsymbol{r}})}_{field} = F(t - \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}}{c})$

$\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{n}} = 1; \quad c = const > 0 \qquad \overrightarrow{\boldsymbol{n}} =$ direction of propagation

Property 1

$\forall \Delta \overrightarrow{\boldsymbol{r}} : \Delta \overrightarrow{\boldsymbol{r}} \cdot \overrightarrow{\boldsymbol{n}} = 0 \qquad \Rightarrow \qquad q(t, \overrightarrow{\boldsymbol{r}} + \Delta \overrightarrow{\boldsymbol{r}}) = q(t, \overrightarrow{\boldsymbol{r}})$

"Planar"

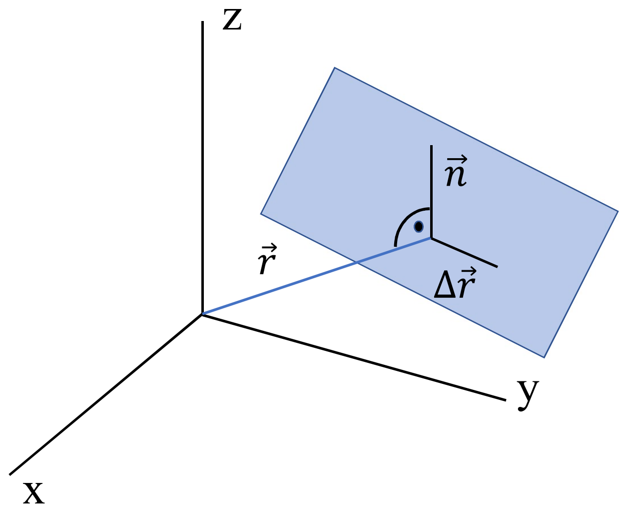

Figure 21: plane in space

Property 2

$q(t + \Delta t, \overrightarrow{\boldsymbol{r}} + \Delta t \cdot c \overrightarrow{\boldsymbol{n}}) \equiv q(t, \overrightarrow{\boldsymbol{r}})$

Moving of $q(t, \overrightarrow{\boldsymbol{r}})$ into direction of $\overrightarrow{\boldsymbol{n}}$ with speed c.

`Harmonic planar wave:`

$F(\tau) = A\cos(2\pi f_0 \tau + \varphi) \qquad A, f_0, \varphi = const.$

$q(t, \overrightarrow{\boldsymbol{r}}) = A\cos(2\pi f_0 (t - \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}}{c}) + \varphi) = \mathrm{Re}\left\{u(\overrightarrow{\boldsymbol{r}}) e^{(j 2\pi f_0 t)}\right\}$

$u(\overrightarrow{\boldsymbol{r}}) = \underbrace{Ae^{(j\varphi)}}_{S \in \mathbb{C}; s = const.} \quad e^{(-j 2\pi f_0 \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}}{c})}$

Property 3

$u(\overrightarrow{\boldsymbol{r}} + \frac{c}{f_0} \overrightarrow{\boldsymbol{n}}) \equiv u(\overrightarrow{\boldsymbol{r}}) \qquad \lambda$ wave length; $\lambda = \frac{c}{f_0}$

$u(\overrightarrow{\boldsymbol{r}}) = s \cdot e^{-j 2\pi \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}}{\lambda}}$

$u(\overrightarrow{\boldsymbol{r}} + l\underline{\boldsymbol{n}}) = u(\overrightarrow{\boldsymbol{r}}) \cdot e^{-j2\pi\frac{l}{\lambda}}$

Phase shift, if you go into direction of wave propagation

`Modulated harmonic planar wave`

s is no longer constant

$s = s(t, \overrightarrow{\boldsymbol{r}}) = s(t - \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}}{c})$ remain wave

$$\boxed{u(\overrightarrow{\boldsymbol{r}}, t) = s(t - \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}}{c}) e^{(-j2\pi\frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}}{\lambda})}}$$

1.3.6 Spatial Sampling

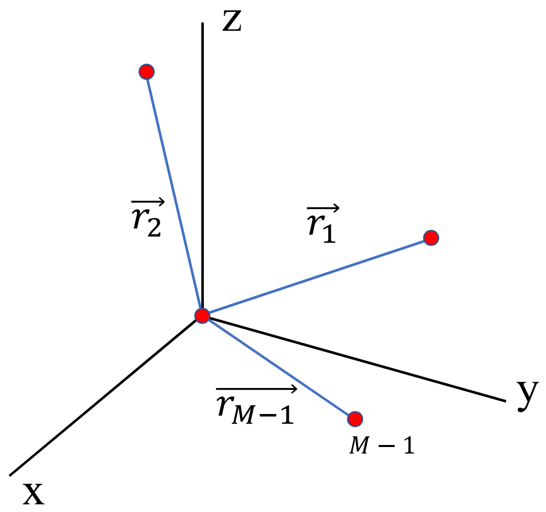

Figure 22: Spatial Sampling with M sensors

$$\underline{\boldsymbol{u}}[t] = \begin{bmatrix} s(t) \\ s(t - \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}_1}{c}) e^{-j2\pi\frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}_1}{\lambda}} \\ \vdots \\ s(t - \frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}_{M-1}}{c}) e^{-j2\pi\frac{\overrightarrow{\boldsymbol{n}} \cdot \overrightarrow{\boldsymbol{r}}_{M-1}}{\lambda}} \end{bmatrix} \Rightarrow \text{Sensor array (receive vector)}$$

Definition coherence: Coherence (physics), an ideal property of waves that enables stationary (i.e. temporally and spatially constant) interference.

- coherence distance, length l_{coh}

if $|l| \leq l_{coh} \iff s(t) \approx s(t-\frac{l}{c})$

- assume all sensors within coherence distance.

$$\underline{\boldsymbol{u}}(t) \approx s(t) \underbrace{\begin{bmatrix} 1 \\ e^{-j2\pi\frac{\vec{n}\cdot\vec{r}_1}{\lambda}} \\ \vdots \\ e^{-j2\pi\frac{\vec{n}\cdot\vec{r}_{M-1}}{\lambda}} \end{bmatrix}}_{\underline{\boldsymbol{a}}(\vec{\boldsymbol{n}}) \text{ Array steering Vector}} \quad \Rightarrow \quad \text{no time delay, only phase shifts}$$

$\underline{\boldsymbol{u}}(nT) = s[n] \cdot \underline{\boldsymbol{a}}(\vec{\boldsymbol{n}})$ with T= Sampling Time, n = Sampling Index $\neq \vec{\boldsymbol{n}}$

`Example:`

$f_0 = 2.6GHz \qquad B = 20MHz$

$\frac{l_{coh}}{c} << \frac{1}{B}, \quad \frac{l_{coh}}{c} \approx \frac{1}{30B}, \quad \lambda = \frac{c}{f_0}$

$l_{coh} \approx 0.5m \approx 4\lambda$

$u[n]$

1.3.7 Large Array

$s[n - \underbrace{\frac{\tau_i}{T}}_{\text{fractioned Delay}}] = s(nT - \tau_i) \qquad \tau_i = \frac{\vec{n}\cdot\vec{r_i}}{c}$

$$s[n - \frac{\tau_i}{T}] \approx \sum_{K=-L_1}^{L_2} a_{i,k} \quad s[n-k]$$

$$\underline{\boldsymbol{u}}[n] = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & e^{-j2\pi\frac{\vec{n}\cdot\vec{r}_1}{\lambda}} & & \\ \vdots & & \ddots & \\ 0 & \cdots & e^{-j2\pi\frac{\vec{n}\cdot\vec{r}_{M-1}}{c}} & \end{bmatrix} \begin{bmatrix} \overbrace{0,...,0,}^{L_1} & 1 & \overbrace{0,...,0,}^{L_2} \\ a_{1,-L_1} & ... & a_{1,-L_2} \\ \vdots & ... & \vdots \\ a_{M-1,-L_1} & ... & a_{M-1,-L_2} \end{bmatrix}}_{\underset{\sim}{\boldsymbol{A}}\vec{\boldsymbol{n}}} \underbrace{\begin{bmatrix} s[n+L_1] \\ s[n+L_1-1] \\ \vdots \\ s[n-L_2] \end{bmatrix}}_{\underline{\boldsymbol{s}}[n]}$$

$\underline{\boldsymbol{u}}[n] = \underset{\sim}{\boldsymbol{A}}(\underline{\boldsymbol{n}}) \cdot \underline{\boldsymbol{s}}[n] \quad \Rightarrow \quad$ long array

$u[n] = \underline{\boldsymbol{a}}(\vec{\boldsymbol{n}}) \cdot s[n] \quad \Rightarrow \quad$ small array

1.3.8 Uniform Linear Array, ULA

same distance in between sensors, all in one line

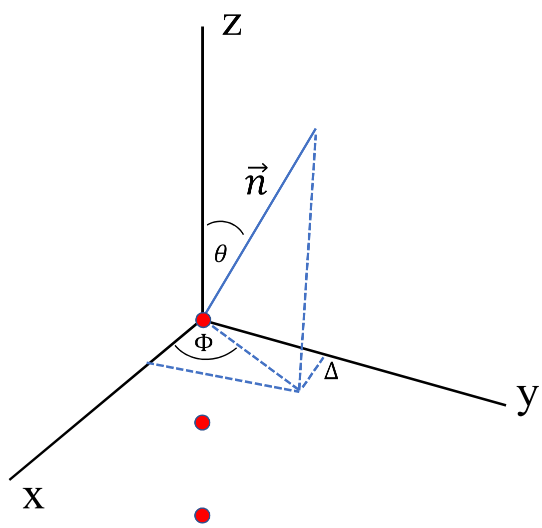

Figure 23: Uniform Linear Array

$$\vec{\boldsymbol{r}}_i = -\vec{\boldsymbol{e}}_z \Delta i \qquad i \in \{0, 1, ..., M-1\}$$

$$\vec{\boldsymbol{r}} = r \begin{bmatrix} cos\varphi sin\theta \\ sin\varphi sin\theta \\ cos\theta \end{bmatrix} \qquad \text{spherical coordinates in cartesian vector}$$

$$\vec{\boldsymbol{n}} = -\frac{\vec{r}}{r}, \qquad \vec{\boldsymbol{n}} \cdot \vec{\boldsymbol{n}} = 1$$

$$\vec{\boldsymbol{r}}_i \cdot \vec{\boldsymbol{n}} = -\Delta i(-\cos\Theta) = cos\Theta \cdot \Delta \cdot i$$

$$\underline{\boldsymbol{a}}(\vec{\boldsymbol{n}}) = \underline{\boldsymbol{a}}(\Theta) = \begin{bmatrix} 1 \\ e^{-j2\pi\frac{\Delta}{\lambda}cos\Theta} \\ e^{-j2\pi\frac{2\Delta}{\lambda}cos\Theta} \\ \vdots \\ e^{-j2\pi\frac{(M-1)\Delta}{\lambda}cos\Theta} \end{bmatrix} \qquad \Rightarrow \quad \text{Van Der Monde-Vector}$$

(ULA Steering Vector)

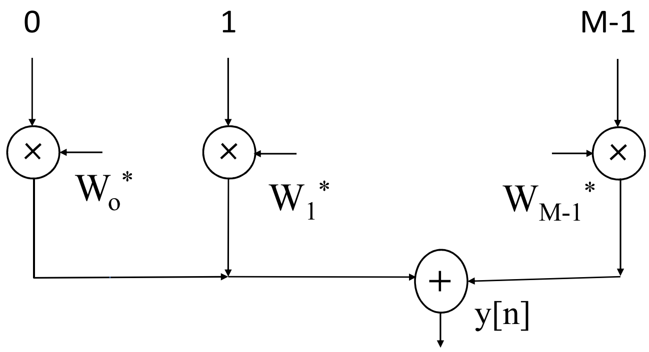

Figure 24: Sensors, figure 23

$$y[n] = \underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n]$$

$$y[n]^* = \underline{\boldsymbol{u}}^H[n] \underline{\boldsymbol{w}}$$

$$y^*[n] = \begin{bmatrix} y^*[n] \\ y^*[n-1] \\ : \\ y^*[n+N] \end{bmatrix} = \underbrace{\begin{bmatrix} \underline{\boldsymbol{u}}^H[n] \\ \underline{\boldsymbol{u}}^H[n+1] \\ : \\ \underline{\boldsymbol{u}}^H[n+N] \end{bmatrix}}_{\underset{\sim}{\boldsymbol{U}}^H[n]} \underline{\boldsymbol{w}} = \underset{\sim}{\boldsymbol{U}}^H[n] \underline{\boldsymbol{w}}$$

$$\underline{\boldsymbol{u}}[n] = \underbrace{\underline{\boldsymbol{a}}(\Theta)s[n]}_{desiredsignal} + \underbrace{\underline{\boldsymbol{\eta}}[n]}_{\text{interference + noise}}$$

Constraint: $\underline{\boldsymbol{w}}^H \underline{\boldsymbol{a}}(\Theta) = 1$

$$y[n] = \underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n] = \underline{\boldsymbol{w}}^H (\underline{\boldsymbol{a}}(\Theta)s[n] + \underline{\boldsymbol{\eta}}[n]) = \underbrace{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{a}}(\Theta)}_{1} s[n] + \underline{\boldsymbol{w}}^H \underline{\boldsymbol{\eta}}[n] = s[n] + \underline{\boldsymbol{w}}^H \underline{\boldsymbol{\eta}}[n]$$

$$\underset{\underline{\boldsymbol{w}}}{min} \quad ||y[n]||_F^2 \quad \text{such that } \underline{\boldsymbol{w}}^H \underline{\boldsymbol{a}}(\Theta) = 1$$

$$\begin{aligned} ||y[n]||_2^2 &= \underline{\boldsymbol{y}}^H[n] \underline{\boldsymbol{y}}[n] \\ &= (\underline{\boldsymbol{y}}^*[n])^H (\underline{\boldsymbol{y}}[n]) \in \mathbb{R} \\ &= (\underset{\sim}{\boldsymbol{U}}^H[n] \underline{\boldsymbol{w}})^H (\underset{\sim}{\boldsymbol{U}}^H[n] \underline{\boldsymbol{w}}) \\ &= \underline{\boldsymbol{w}} \underbrace{\underset{\sim}{\boldsymbol{U}}[n] \underset{\sim}{\boldsymbol{U}}^H[n]}_{\underset{\sim}{\boldsymbol{R}}} \underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}} \end{aligned}$$

$$\boxed{\underline{\boldsymbol{w}}_{OPT} = arg \min_{\underline{\boldsymbol{w}}} \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}}, \qquad \text{s.t.} \quad \underline{\boldsymbol{w}}^H \underline{\boldsymbol{a}}(\Theta) = 1}$$

2 Mathematical background

2.1 Gernral

2.1.1 Matrix inverse

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22}-a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

2.1.2 Hermitian matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{H} = \begin{bmatrix} a_{11}^* & a_{21}^* \\ a_{12}^* & a_{22}^* \end{bmatrix}$$

2.1.3 Quadratic equation

$ax^2 + bx + c = 0 \quad \Rightarrow \quad x_{1/2} = \frac{-b\pm\sqrt{b^2-4ac}}{2a}$
$x^2 + px + q = 0 \quad \Rightarrow \quad x_{1/2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$

2.1.4 Euklidian Norm

$$||\underline{\boldsymbol{v}}||_2 = \sqrt{v_1^2 + v_2^2 + ... + v_N} = \sqrt{\sum_{i=1}^{N} |v_i|^2}$$

2.1.5 Frobenius Norm

$$||\underset{\sim}{\boldsymbol{V}}||_F = \sqrt{|v_{1,1}|^2 + |v_{1,2}|^2 + ... + |v_{N,M}|^2} = \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{M} |v_{i,j}|^2}$$

2.1.6 Taylor series

$$Tf(x,a) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f'''(a)}{6}(x-a)^3 + \dots$$

2.1.7 Fourier series

$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos(kt) + b_k \sin(kt)\right)$
$a_k = c_k + c_{-k} \quad$ für $k \geq 0$
$b_k = \mathrm{i}\left(c_k - c_{-k}\right) \quad$ für $k \geq 1$
or as an alternative:
$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cdot \cos(kt)\,\mathrm{d}t \quad$ für $k \geq 0$
$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cdot \sin(kt)\,\mathrm{d}t \quad$ für $k \geq 1$

2.2 Complex derivative

2.2.1 Analytic function

$$h : \mathbb{C} \ni z \mapsto h(z) \in \mathbb{C} \qquad \text{with} \quad z = x + \mathrm{j}y \quad x, y \in \mathbb{R}$$

$\frac{dh}{dz} = \lim\limits_{\Delta z \to 0} \frac{h(z+\Delta z)-h(z)}{\Delta z}$

$h(z) = h(x + \mathrm{j}y) = f(x,y) \quad f : \mathbb{R}^2 \ni (x,y) \mapsto f(x,y) \in \mathbb{C}$

$dh = df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$

Substitution: $dx = a \cdot dt, \quad dy = b \cdot dt, \quad (a,b) \in \mathbb{R}^2 \backslash (0,0)$

$dz = dx + \mathrm{j}dy = (a + jb)dt \quad \Rightarrow \quad dx = \frac{a}{a+\mathrm{j}b}dz, \quad dy = \frac{b}{a+\mathrm{j}b}dz$

$$df = dh = \underbrace{\frac{\frac{\partial f}{\partial x}a + \frac{\partial f}{\partial y}b}{a+\mathrm{j}b}}_{H(z)} dz$$

- Definition: $h(z)$ is analytic if and only if (iff) $H(z)$ exist and is indepent of $(a,b) \in \mathbb{R}^2 \backslash (0,0)$
- if $h(z)$ is analytic then $\frac{dh}{dz} = H$

$$\frac{\partial H}{\partial a} = \frac{\mathrm{j}b}{(a+\mathrm{j}b)^2}\underbrace{\left(\frac{\partial f}{\partial x} + \mathrm{j}\frac{\partial f}{\partial y}\right)}_{=0}, \qquad \frac{\partial H}{\partial b} = \frac{-\mathrm{j}a}{(a+\mathrm{j}b)^2}\underbrace{\left(\frac{\partial f}{\partial x} + \mathrm{j}\frac{\partial f}{\partial y}\right)}_{=0}$$

- $h(z)$ is analytic if and only if $\frac{df}{dx}$ and $\frac{df}{dy}$ exist and $\frac{df}{dx} + \mathrm{j}\frac{df}{dy} = 0$.
- if $h(z)$ is analytic then $\frac{dh}{dz} = \frac{df}{dx} = -\mathrm{j}\frac{df}{dy}$

Example for analytic functions - Exponential function

$h(z) = z^k \quad k \in \mathbb{R}$

$f(x,y) = (x + \mathrm{j}y)^k$

$\frac{df}{dx} + \mathrm{j}\frac{df}{dy} = k(x + \mathrm{j}y)^{k-1} + \mathrm{j}k(x + \mathrm{j}y)^{k-1}\mathrm{j} = 0$, hence it is an analytic function

Result: $\frac{dh}{dz} = k(x + \mathrm{j}y)^{k-1} = k \cdot z^{k-1}$

other analytic functions

- $h(z) = const.$
- $h(z) = e^z$
- $h(z) = \ln(z)$, of $z = 0$
- $h(z) = \frac{\sum_{k=0}^{\infty} a_k \cdot z^k}{\sum_{k=0}^{\infty} b_k \cdot z^k}$, for the poles

example of non-analytic functions:

- $h(z) = z^* = x - \mathrm{j}y = f(x,y)$

  $\frac{\partial f}{\partial x} + \mathrm{j}\frac{\partial f}{\partial y} = 1 + \mathrm{j}(-\mathrm{j}) = 2 \neq 0 \Rightarrow$ non analytic
- $h(z) = \mathrm{Re}\{z\} = \frac{z+z^*}{2}$

  $h : \mathbb{C} \ni z \mapsto h(z) \in \mathbb{R}$ is non-analytic, it is a constant
- $\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}}$

2.2.2 Derivatives of non-analytic functions

$$\left.\begin{array}{l} z = x + jy, \\ z^* = x - jy \end{array}\right\} \quad x = \frac{z+z^*}{2}, y = \frac{z-z^*}{2j} \quad x, y \in \mathbb{R}$$

$h(z) = h(x + \mathrm{j}y) = f(x, y) = f\left(\frac{z+z^*}{2}, \frac{z-z^*}{\mathrm{j}2}\right) = g(z, z^*)$
$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$
$dg = \frac{\partial g}{\partial z}dz + \frac{\partial g}{\partial z^*}dz^* = \frac{\partial g}{\partial z}(dx + \mathrm{j}dy) + \frac{\partial g}{\partial z^*}(dx - \mathrm{j}dy) = \underbrace{\left(\frac{\partial g}{\partial z} + \frac{\partial g}{\partial z^*}\right)}_{\frac{\partial f}{\partial x}} dx + \mathrm{j}\underbrace{\left(\frac{\partial g}{\partial z} - \frac{\partial g}{\partial z^*}\right)}_{\frac{\partial f}{\partial y}} dy$

In matrix notation

$$\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \mathrm{j} & -\mathrm{j} \end{bmatrix} \begin{bmatrix} \frac{\partial g}{\partial z} \\ \frac{\partial g}{\partial z^*} \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial g}{\partial z} \\ \frac{\partial g}{\partial z^*} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -\mathrm{j} \\ 1 & \mathrm{j} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Note: $\frac{\partial g}{\partial z^*} = \frac{1}{2}\frac{\partial f}{\partial x} + \frac{1}{2}\mathrm{j}\frac{\partial f}{\partial y}$
$g(z, z^*) = h(z)$ is analytic iff $\frac{\partial g}{\partial z^*} = 0$
$(\frac{\partial g}{\partial z^*})^* = \frac{\partial g^*}{\partial z} \neq \frac{\partial g}{\partial z}$
when $g : \mathbb{C}^2 \mapsto \mathbb{R}$ then: $(\frac{\partial g}{\partial z^*})^* = \frac{\partial g}{\partial z}$
In the following, we assume that
$h : \mathbb{C} \mapsto \mathbb{R}, \quad g : \mathbb{C}^2 \mapsto \mathbb{R}$
$dg = \frac{\partial g}{\partial z}dz + \underbrace{\frac{\partial g}{\partial z^*}dz^*}_{\text{only if}\mapsto\mathbb{R}:(\frac{\partial g}{\partial z}dz)^*} = 2\,\mathrm{Re}\left\{\frac{\partial g}{\partial z}dz\right\}$
Local extremum $dg = 0 \quad \forall dz \qquad \Longleftrightarrow \qquad \frac{\partial g}{\partial z} = 0$
multiple variables
$g : \mathbb{C}^{2n} \mapsto \mathbb{R}; \qquad g(z_1, z_2, ..., z_n, z_1^*, z_2^*, ..., z_n^*) = g(\underline{\boldsymbol{z}}, \underline{\boldsymbol{z}}^*)$
$dg = \sum_{k=1}^n \frac{\partial g}{\partial z_k}dz_k + \sum_{k=1}^n \frac{\partial g}{\partial z_k^*}dz_k^* = 2\,\mathrm{Re}\left\{\sum_{k=1}^n \frac{\partial g}{\partial z_k}dz_k\right\}$

2.2.3 Gradient vector

$$\frac{\partial}{\partial \underline{z}} = \begin{bmatrix} \partial/\partial z_1 \\ \partial/\partial z_2 \\ \vdots \\ \partial/\partial z_n \end{bmatrix} \qquad \frac{\partial}{\partial \underline{z}}g = \frac{\partial g}{\partial \underline{z}} = \begin{bmatrix} \partial g/\partial z_1 \\ \partial g/\partial z_2 \\ \vdots \\ \partial g/\partial z_n \end{bmatrix} \quad d\underline{\boldsymbol{z}} = \begin{bmatrix} dz_1 \\ dz_2 \\ \vdots \\ dz_n \end{bmatrix}$$

Scalaproduct:

$$\sum_{k=1}^{n} = \frac{\partial g}{\partial z_k}dz_k = \left(\frac{\partial g}{\partial \underline{z}}\right)^T d\underline{\boldsymbol{z}}$$

$$dg = 2\,\mathrm{Re}\left\{\left(\frac{\partial g}{\partial \underline{z}}\right)^T d\underline{\boldsymbol{z}}\right\} = 2\,\mathrm{Re}\left\{\left(\left(\frac{\partial g}{\partial \underline{z}^*}\right)^*\right)^T d\underline{\boldsymbol{z}}\right\} = 2\,\mathrm{Re}\left\{\left(\frac{\partial g}{\partial \underline{z}^*}\right)^H d\underline{\boldsymbol{z}}\right\}$$

Local extremum $dg = 0 \quad \forall d\underline{\boldsymbol{z}} \qquad \Longleftrightarrow \qquad \frac{\partial g}{\partial \underline{\boldsymbol{z}}^*} = \underline{0}$

Steepest descent

$\underline{\boldsymbol{z}} \to \underline{\boldsymbol{z}} + \Delta\underline{\boldsymbol{z}}$

$dg = 2\,\mathrm{Re}\left\{\left(\frac{\partial g}{\partial \underline{z}^*}\right)^H \Delta\underline{\boldsymbol{z}}\right\} \leq 2\left|\left(\frac{\partial g}{\partial \underline{z}^*}\right)^H \Delta\underline{\boldsymbol{z}}\right|$

equality only if $\left(\frac{\partial g}{\partial \underline{z}^*}\right)^H \Delta\underline{\boldsymbol{z}} \in \mathbb{R}_0^+$

$\Delta\underline{\boldsymbol{z}} = \varepsilon\frac{\partial g}{\partial \underline{\boldsymbol{z}}^*} \Rightarrow \left(\frac{\partial g}{\partial \underline{z}^*}\right)^H \Delta\underline{\boldsymbol{z}} = \varepsilon\left(\frac{\partial g}{\partial \underline{z}^*}\right)^H\left(\frac{\partial g}{\partial \underline{z}^*}\right) \quad \in \mathbb{R}_0^+ \qquad \varepsilon > 0$

steepest descent: $\Delta\underline{\boldsymbol{z}} = -\frac{\partial g}{\partial \underline{\boldsymbol{z}}^*}$

Example 1

$g = \underline{\boldsymbol{z}}^H\underline{\boldsymbol{p}} + \underline{\boldsymbol{p}}^H\underline{\boldsymbol{z}}; \quad \underline{\boldsymbol{p}} = const$

g is real:

$g = g^* = g^H = (\underline{\boldsymbol{z}}^H\underline{\boldsymbol{p}} + \underline{\boldsymbol{p}}^H\underline{\boldsymbol{z}})^H = \underline{\boldsymbol{p}}^H\underline{\boldsymbol{z}} + \underline{\boldsymbol{z}}^H\underline{\boldsymbol{z}} = g$

$g = g^* \Rightarrow real$

$$\frac{\partial g}{\partial \underline{\boldsymbol{z}}^*} = \frac{\partial}{\partial \underline{\boldsymbol{z}}^*}(\underline{\boldsymbol{z}}^H\underline{\boldsymbol{p}} + \underline{\boldsymbol{p}}^H\underline{\boldsymbol{z}}) = \frac{\partial}{\partial \underline{\boldsymbol{z}}^*}(\underline{\boldsymbol{z}}^H\underline{\boldsymbol{p}}) = \frac{\partial}{\partial \underline{\boldsymbol{z}}^*}\sum_{k=1}^{n} z_k^* P_k = \begin{bmatrix} \frac{\partial}{\partial z_1^*}\sum_{k=1}^n z_k^* P_k \\ \frac{\partial}{\partial z_2^*}\sum_{k=1}^n z_k^* P_k \\ \vdots \\ \frac{\partial}{\partial z_n^*}\sum_{k=1}^n z_k^* P_k \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix} = \underline{\boldsymbol{p}}$$

result: $\frac{\partial}{\partial \underline{\boldsymbol{z}}^*}(\underline{\boldsymbol{z}}^H\underline{\boldsymbol{p}} + \underline{\boldsymbol{p}}^H\underline{\boldsymbol{z}}) = \underline{\boldsymbol{p}}$

alternativ: $\frac{\partial}{\partial \underline{\boldsymbol{z}}}(\underline{\boldsymbol{z}}^H\underline{\boldsymbol{p}} + \underline{\boldsymbol{p}}^H\underline{\boldsymbol{z}}) = \underline{\boldsymbol{p}}^*$

Example 2

$g = \underline{\boldsymbol{z}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{z}}; \qquad \underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H = const$

g is real:

$g^* = g^H = (\underline{\boldsymbol{z}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{z}})^H = \underline{\boldsymbol{z}}^H\underset{\sim}{\boldsymbol{R}}^H\underline{\boldsymbol{z}} \overset{\underset{\sim}{\boldsymbol{R}}=\underset{\sim}{\boldsymbol{R}}^H}{=} g \Rightarrow real$

$\frac{\partial g}{\partial \underline{\boldsymbol{z}}^*} = \frac{\partial}{\partial \underline{\boldsymbol{z}}^*}(\underline{\boldsymbol{z}}^H\underbrace{\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{z}}}_{\underline{\boldsymbol{p}}}) = \frac{\partial}{\partial \underline{\boldsymbol{z}}^*}(\underline{\boldsymbol{z}}^H\underline{\boldsymbol{p}}) = \underline{\boldsymbol{p}} = \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{z}}$

$\underline{\boldsymbol{p}}$ dosn't depent on $\underline{\boldsymbol{z}}^*$

result: $\frac{\partial g}{\partial \underline{\boldsymbol{z}}^*} = \frac{\partial}{\partial \underline{\boldsymbol{z}}^*}(\underline{\boldsymbol{z}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{z}}) = \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{z}}$

2.3 Minimization with linear equality constraints

2.3.1 Real-valued

$$\min_{\underline{\boldsymbol{x}}\in\mathbb{R}^n} \underline{\boldsymbol{x}}^T\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{x}}, \qquad s.t. \quad \underline{\boldsymbol{a}}^T\underline{\boldsymbol{x}} = b, \quad \underset{\sim}{\boldsymbol{R}} \in \mathbb{R}^{nxn}, \underline{\boldsymbol{a}} \in \mathbb{R}^n, b \in \mathbb{R}$$

Method of Lagrange

$\mathcal{L}(\underline{\boldsymbol{x}}, \lambda) := \underline{\boldsymbol{x}}^T\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{x}} + \lambda(\underline{\boldsymbol{a}}^T\underline{\boldsymbol{x}} - b) \quad \Rightarrow \quad \min_{\underline{\boldsymbol{x}}}\max_{\lambda}\mathcal{L}(\underline{\boldsymbol{x}}, \lambda)$

2.3.2 Complex-valued

$$\min_{\underline{\boldsymbol{z}}\in\mathbb{C}^n} \underline{\boldsymbol{z}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{z}}, \qquad s.t. \quad \underline{\boldsymbol{a}}^H\underline{\boldsymbol{z}} = b, \quad \underset{\sim}{\boldsymbol{R}} \in \mathbb{C}^{nxn}, \underline{\boldsymbol{a}} \in \mathbb{C}^n, b \in \mathbb{C}$$

$\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H \in \mathbb{C} \quad \text{and} \quad \underset{\sim}{\boldsymbol{R}} \geq \underline{\boldsymbol{0}}$

$$\mathcal{L}(\underline{\boldsymbol{z}}, \lambda_R, \lambda_I) := \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + \lambda_R \operatorname{Re}\left\{\underline{\boldsymbol{a}}^H \underline{\boldsymbol{z}} - b\right\} + \lambda_I \operatorname{Im}\left\{\underline{\boldsymbol{a}}^H \underline{\boldsymbol{z}} - b\right\}$$

$$\begin{aligned}\mathcal{L}(\underline{\boldsymbol{z}}, \lambda) &:= \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + 2\operatorname{Re}\left\{\lambda^*(\underline{\boldsymbol{a}}^H \underline{\boldsymbol{z}} - b)\right\} \\ &= \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + \lambda^*(\underline{\boldsymbol{a}}^H \underline{\boldsymbol{z}} - b) + (\underline{\boldsymbol{z}}^H \underline{\boldsymbol{a}} - b^H)\lambda \\ &= \mathcal{L}(\underline{\boldsymbol{z}}, \underline{\boldsymbol{z}}^*, \lambda, \lambda^*)\end{aligned}$$

2.3.3 Several complex constraints

$$\min_{\underline{\boldsymbol{z}} \in \mathbb{C}^n} \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}}, \qquad s.t. \quad \begin{cases} \underline{\boldsymbol{a}}_1^H \underline{\boldsymbol{z}} = b_1 \\ : \\ \underline{\boldsymbol{a}}_m^H \underline{\boldsymbol{z}} = b_m \end{cases}$$

$$\mathcal{L} = \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + \sum_{k=1}^{m} \lambda_k^*(\underline{\boldsymbol{a}}_k^H \underline{\boldsymbol{z}} - b_k) + \sum_{k=1}^{m} (\underline{\boldsymbol{z}}^H \underline{\boldsymbol{a}}_k - b_k^*)\lambda_k$$

with:

- $\underline{\boldsymbol{\lambda}} = \begin{bmatrix} \lambda_1 \\ : \\ \lambda_m \end{bmatrix} \in \mathbb{C}^m, \quad \underline{\boldsymbol{b}} = \begin{bmatrix} b_1 \\ : \\ b_m \end{bmatrix} \in \mathbb{C}^m$
- $\sum_{k=1}^{m} \lambda_k^* b_k = \underline{\boldsymbol{\lambda}}^H \underline{\boldsymbol{b}}$
- $\sum_{k=1}^{m} \lambda_k^* \underline{\boldsymbol{a}}_k^H \underline{\boldsymbol{z}} = \underline{\boldsymbol{\lambda}}^H \begin{bmatrix} \underline{\boldsymbol{a}}_1^H \underline{\boldsymbol{z}} \\ : \\ \underline{\boldsymbol{a}}_m^H \underline{\boldsymbol{z}} \end{bmatrix} = \underline{\boldsymbol{\lambda}}^H \begin{bmatrix} \underline{\boldsymbol{a}}_1^H \\ : \\ \underline{\boldsymbol{a}}_m^H \end{bmatrix} \underline{\boldsymbol{z}} = \underline{\boldsymbol{\lambda}}^H \underset{\sim}{\boldsymbol{A}}^H \underline{\boldsymbol{z}}$

follows:

$$\min_{\underline{\boldsymbol{z}} \in \mathbb{C}^n} \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}}, \qquad s.t. \quad \underset{\sim}{\boldsymbol{A}}^H \underline{\boldsymbol{z}} = \underline{\boldsymbol{b}}, \quad \underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H$$

$$\mathcal{L} = \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + \underline{\boldsymbol{\lambda}}^H (\underset{\sim}{\boldsymbol{A}}^H \underline{\boldsymbol{z}} - \underline{\boldsymbol{b}}) + (\underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{A}} - \underline{\boldsymbol{b}}^H)\underline{\boldsymbol{\lambda}}$$

$$\Rightarrow \quad \min_{\underline{\boldsymbol{z}}} \max_{\underline{\boldsymbol{\lambda}}} \mathcal{L}(\underline{\boldsymbol{x}}, \lambda)$$

Solution

$\underset{\sim}{\boldsymbol{R}}$ is positiv definit: $\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H \iff \underset{\sim}{\boldsymbol{R}}^{-1}$ exists and $\underset{\sim}{\boldsymbol{R}}^{-1} = (\underset{\sim}{\boldsymbol{R}}^{-1})^H$

$$\frac{\partial L}{\partial \underline{\boldsymbol{z}}^*} = \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{\lambda}} \overset{!}{=} \underline{\boldsymbol{0}} \quad \Rightarrow \underline{\boldsymbol{z}} = -\underset{\sim}{\boldsymbol{R}}^{-1} \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{\lambda}}$$

$$\underline{\boldsymbol{b}} = \underset{\sim}{\boldsymbol{A}}^H \underline{\boldsymbol{z}} = -\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}^{-1} \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{\lambda}} \quad \Rightarrow \underline{\boldsymbol{\lambda}} = -(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}^{-1} \underset{\sim}{\boldsymbol{A}})^{-1} \underline{\boldsymbol{b}}$$

$$\underline{\boldsymbol{z}}_{\mathbf{opt}} = \underset{\sim}{\boldsymbol{R}}^{\mathbf{-1}} \underset{\sim}{\boldsymbol{A}} (\underset{\sim}{\boldsymbol{A}}^{\mathbf{H}} \underset{\sim}{\boldsymbol{R}}^{\mathbf{-1}} \underset{\sim}{\boldsymbol{A}})^{\mathbf{-1}} \underline{\boldsymbol{b}}$$

check if local minimum

$\Delta \underline{\boldsymbol{z}}$ such that $\underset{\sim}{\boldsymbol{A}}^H (\underline{\boldsymbol{z}}_{opt} + \Delta \underline{\boldsymbol{z}}) = \underline{\boldsymbol{b}} \quad \iff \underset{\sim}{\boldsymbol{A}}^H \Delta \underline{\boldsymbol{z}} = \underline{\boldsymbol{0}} \qquad \Delta \underline{\boldsymbol{z}} \in \operatorname{null} \underset{\sim}{\boldsymbol{A}}$; see section 2.4.7

$$d\mathcal{L} = 2\operatorname{Re}\left\{\frac{\partial L}{\partial \underline{\boldsymbol{z}}^*} d\underline{\boldsymbol{z}}\right\}$$

$$\frac{\partial \mathcal{L}}{\partial \underline{\boldsymbol{z}}^*} = \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{\lambda}} = \underset{\sim}{\boldsymbol{R}} (\underline{\boldsymbol{z}}_{opt} + \Delta \underline{\boldsymbol{z}}) + \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{\lambda}} = \underbrace{\underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}}_{opt} + \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{\lambda}}}_{\underline{\boldsymbol{0}}} + \underset{\sim}{\boldsymbol{R}} \Delta \underline{\boldsymbol{z}} = \underset{\sim}{\boldsymbol{R}} \Delta \underline{\boldsymbol{z}}$$

$$d\mathcal{L} = 2\operatorname{Re}\left\{\Delta \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}}^H (-\Delta \underline{\boldsymbol{z}})\right\} = -2\operatorname{Re}\left\{\underbrace{\Delta \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}}^H \Delta \underline{\boldsymbol{z}}}_{>0}\right\} = -2\, \underbrace{\Delta \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \Delta \underline{\boldsymbol{z}}}_{>0} < 0 \qquad \text{local minimum}$$

$\mathcal{L} = \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}} + \underline{\boldsymbol{\lambda}}^H (\underset{\sim}{\boldsymbol{A}}^H \underline{\boldsymbol{z}} - \underline{\boldsymbol{b}}) + (\underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{A}} - \underline{\boldsymbol{b}}^H)\underline{\boldsymbol{\lambda}}$ if contraints met: $v = \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{z}}$

2.4 Linear Algebra

2.4.1 Vector spaces

Definition: A **vector space** $\mathbb{V}$ over a set $\mathbb{K}$ is a set such that:

1. $\forall \underline{\boldsymbol{a}}, \underline{\boldsymbol{b}} \in \mathbb{V}: \quad \underline{\boldsymbol{a}} + \underline{\boldsymbol{b}} \in \mathbb{V}$
   The vectors $\underline{\boldsymbol{a}}$ and $\underline{\boldsymbol{b}}$ in vectorspace $\mathbb{V}$ can be added to a new vector in the vector space.

2. $\forall \underline{\boldsymbol{a}}, \underline{\boldsymbol{b}} \in \mathbb{V}: \quad \underline{\boldsymbol{a}} + \underline{\boldsymbol{b}} = \underline{\boldsymbol{b}} + \underline{\boldsymbol{a}}$
   Addition in commutative.

3. $\forall \underline{\boldsymbol{a}}, \underline{\boldsymbol{b}}, \underline{\boldsymbol{c}} \in \mathbb{V}: \quad (\underline{\boldsymbol{a}} + \underline{\boldsymbol{b}}) + \underline{\boldsymbol{c}} = \underline{\boldsymbol{a}} + (\underline{\boldsymbol{b}} + \underline{\boldsymbol{c}})$
   The order to operate addition ist not relevant.

4. $\forall \underline{\boldsymbol{a}} \in \mathbb{V} : \exists \underline{\mathbf{0}} \in \mathbb{V} : \underline{\boldsymbol{a}} + \underline{\mathbf{0}} = \underline{\boldsymbol{a}}$ with $\underline{\mathbf{0}}$ as the identity element.
   The sum of a vector with the zero-vector equals the original vector.

5. $\forall \underline{\boldsymbol{a}} \in \mathbb{V} : \exists - \underline{\boldsymbol{a}} \in \mathbb{V} : \underline{\boldsymbol{a}} + (-\underline{\boldsymbol{a}}) = \underline{\mathbf{0}}$
   The sum of a vector with it's negative opponent equals the zero-vector.

6. $\forall \underline{\boldsymbol{a}} \in \mathbb{V} : 1\underline{\boldsymbol{a}} = \underline{\boldsymbol{a}}; 1 \in \mathbb{K}$
   Scaling a vector with 1 results in the vector.

7. $\forall \underline{\boldsymbol{a}} \in \mathbb{V}, \forall \lambda, \mu \in \mathbb{K} : \lambda(\mu\underline{\boldsymbol{a}}) = (\lambda\mu)\underline{\boldsymbol{a}}$
   Scaling a vector with two scalars is commutative.

8. $\forall \underline{\boldsymbol{a}}, \underline{\boldsymbol{b}} \in \mathbb{V}, \forall \lambda \in \mathbb{K} : \lambda(\underline{\boldsymbol{a}} + \underline{\boldsymbol{b}}) = \lambda\underline{\boldsymbol{a}} + \lambda\underline{\boldsymbol{b}}$
   Scaling two vectors with the same scalar equals scaling two vectors with the scalar seperately.

9. $\forall \underline{\boldsymbol{a}} \in \mathbb{V}, \forall \lambda, \mu \in \mathbb{K} : (\mu + \lambda)\underline{\boldsymbol{a}} = \mu\underline{\boldsymbol{a}} + \lambda\underline{\boldsymbol{a}}$
   It's irrelevant if two scalars are summed up first or multiplied by a vector individually.

Conclusion:

1. $\underline{\boldsymbol{a}} + \underline{\boldsymbol{a}} = 2\underline{\boldsymbol{a}} \quad \mathbb{K} = \mathbb{C}$
   proof: $\overbrace{\underline{\boldsymbol{a}} + \underline{\boldsymbol{a}}}^{1} \overset{6}{=} 1\underline{\boldsymbol{a}} + 1\underline{\boldsymbol{a}} \overset{9}{=} \overbrace{(1+1)}^{2 in \mathbb{K}} \underline{\boldsymbol{a}} = 2\underline{\boldsymbol{a}}$

2. $\underline{\mathbf{0}}$ is unique
   proof: Assume there is another $\underline{\mathbf{0}}$
   $$\left.\begin{array}{ll} 4 & \underline{\mathbf{0}}' + \underline{\mathbf{0}} = \underline{\mathbf{0}}' \\ 4 & \underline{\mathbf{0}} + \underline{\mathbf{0}}' = \underline{\mathbf{0}} \\ 2 & \underline{\mathbf{0}}' + \underline{\mathbf{0}} = \underline{\mathbf{0}} + \underline{\mathbf{0}}' \end{array}\right\} \underline{\mathbf{0}}' = \underline{\mathbf{0}}$$

Example of vector spaces

- $\{\underline{\mathbf{0}}\}$

- $\mathbb{K}=\mathbb{Z}, \mathbb{V}=\mathbb{Z}$ integer numbers
- $\mathbb{K}=\mathbb{C}, \mathbb{V}=\mathbb{C}$ complex numbers
- $\mathbb{K}=\mathbb{C}$,
  $\mathbb{V}=\{a_0+a_1z+a_2z^2+\cdots+a_nz^n|z,a_0,\cdots,a_n\in\mathbb{C}\}$ polynoms
- $\mathbb{K}=\mathbb{C}, \mathbb{V}=\mathbb{C}^n=\left\{\begin{bmatrix}a_1\\ :\\ a_n\end{bmatrix}\right\}, a_1,a_2,\cdots,a_n\in\mathbb{C}$ vectors

counter example

- $\{\}$ empty set

Definition: A set $\mathbb{S}$ is called **subspace of a vector space** $\mathbb{V}$ over $\mathbb{K}$, if and only if:

1. $\mathbb{S}\subseteq\mathbb{V}$ Must be a subset of $\mathbb{V}$
2. $\forall\underline{\boldsymbol{a}},\underline{\boldsymbol{b}}\in\mathbb{S}:\underline{\boldsymbol{a}}+\underline{\boldsymbol{b}}\in\mathbb{S}$
3. $\forall\underline{\boldsymbol{a}}\in\mathbb{S},\forall\lambda\in\mathbb{K}:\lambda\underline{\boldsymbol{a}}\in\mathbb{S}$

Note: $\underline{\mathbf{0}}\in\mathbb{S};\mathbb{S}\neq\{\}$
Linear combination: $\underline{\boldsymbol{v}}_1,...,\underline{\boldsymbol{v}}_n\in\mathbb{S}\Rightarrow\sum\limits_{k=1}^{n}a_k\underline{\boldsymbol{v}}_k\in\mathbb{S};a_k\in\mathbb{K}$

2.4.2 Vectors

In the following

$\mathbb{K}=\mathbb{C},\mathbb{V}=\mathbb{C}^m$

$a_1,...,a_m\in\mathbb{C}$

colum vector: $\underline{\boldsymbol{a}}:=\begin{bmatrix}a_1\\ :\\ a_m\end{bmatrix}$

row vector: $\underline{\boldsymbol{a}}^T:=\begin{bmatrix}a_1 & \cdots & a_m\end{bmatrix}$ $(\underline{\boldsymbol{a}}^T)^T=\underline{\boldsymbol{a}}$

complex conjugate: $\underline{\boldsymbol{a}}^*=\begin{bmatrix}a_1^*\\ :\\ a_m^*\end{bmatrix}$ $(\underline{\boldsymbol{a}}^*)^*=\underline{\boldsymbol{a}}$

Hermite-ian vector: $(\underline{\boldsymbol{a}}^*)^T=(\underline{\boldsymbol{a}}^T)^*:=\underline{\boldsymbol{a}}^H$ $(\underline{\boldsymbol{a}}^H)^H=\underline{\boldsymbol{a}}$

complex scalar product: $\underline{\boldsymbol{a}}^H\underline{\boldsymbol{b}}=a_1^*b_1+...+a_m^*b_m=(\underline{\boldsymbol{b}}^H\underline{\boldsymbol{a}})^*$

$\underline{\boldsymbol{a}},\underline{\boldsymbol{b}}$ are orthogonal iff $\underline{\boldsymbol{a}}^H\underline{\boldsymbol{b}}=\underline{\boldsymbol{b}}^H\underline{\boldsymbol{a}}=0$

euclidian norm: $||\underline{\boldsymbol{a}}||_2:=\sqrt{\underline{\boldsymbol{a}}^H\underline{\boldsymbol{a}}}=\sqrt{a_1^*a_1+...+a_m^*a_m}=\sqrt{|a_1|^2+...+|a_m|^2}$

Definition: Span

$\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n) := \left\{ \sum_{k=1}^{n} a_k \underline{\boldsymbol{v}}_k | a_1, ..., a_n \in \mathbb{C} \right\}$

$\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n \in \mathbb{C}^m$

Note: $\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ... \underline{\boldsymbol{v}}_n)$ is a subspace of $\mathbb{C}^m$

Definition:

The vectors $\underline{\boldsymbol{v}}_1, ... \underline{\boldsymbol{v}}_n$ are linearly independent (LID) if and only if: $a_1 \underline{\boldsymbol{v}}_1 + a_2 \underline{\boldsymbol{v}}_2 ... + a_n \underline{\boldsymbol{v}}_n = \underline{\mathbf{0}}$ requires that $a_1 = ... = a_n = 0$

Theorem:

if $\underline{\boldsymbol{v}}_1, ... \underline{\boldsymbol{v}}_n$ are not LID, that is if they are linearly dependent (LD), then $\exists_i : \underline{\boldsymbol{v}}_i = \sum_{k=1, k \neq i}^{n} b_k \underline{\boldsymbol{v}}_k$ and vice versa (und umgekehrt)

Proof:

$a_1 \underline{\boldsymbol{v}}_1 + ... + a_n \underline{\boldsymbol{v}}_n = \underline{\mathbf{0}}$ for $\underline{\boldsymbol{v}}_i$ with $a_i \neq 0$

$$a_k = \begin{cases} b_k \text{ for } k \neq i \\ -1 \text{ for } k = 1 \end{cases} \qquad \square$$

Definition:

Dimension, $dim\mathbb{S}$ of a subset $\mathbb{S}$ is the maximum number of LID vectors in $\mathbb{S}$

Theorem:

Every subset $\mathbb{S}$ with $dim\mathbb{S} = n$ is a span $\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$ of n LID vectors $\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n$ in $\mathbb{S}$

Proof:

- $\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n) \subseteq \mathbb{S}$, from definition
- show $\mathbb{S} \subseteq \mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$

suppose $\underline{\boldsymbol{s}} \in \mathbb{S} : \underline{\boldsymbol{s}} \in \mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$, then $(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n, \underline{\boldsymbol{s}})$ are LID

$\underbrace{(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n, \underline{\boldsymbol{s}})}_{dim=n+1} \subseteq \underbrace{\mathbb{S}}_{dim=n}$ because $\underline{\boldsymbol{s}} \in \mathbb{S}$

$\Rightarrow \forall \underline{\boldsymbol{s}} \in \mathbb{S} : \underline{\boldsymbol{s}} \in \mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n))$

$\Rightarrow \mathbb{S} = \mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$ $\square$

Note: $dim(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n) = n$ if $\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n$ are LID

Definition: the n LID vectors $\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n$ are called **base vectors** of $\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$

Note: base vectors are not unique:

eg: $\mathbb{S}_1 = \mathrm{Sp}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right), \mathbb{S}_2 = \mathrm{Sp}\left(\begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \right)$

$$\begin{bmatrix}0\\2\\1\end{bmatrix}=\begin{bmatrix}0\\1\\0\end{bmatrix}+\begin{bmatrix}0\\1\\1\end{bmatrix},\begin{bmatrix}0\\0\\-1\end{bmatrix}=\begin{bmatrix}0\\1\\0\end{bmatrix}-\begin{bmatrix}0\\1\\1\end{bmatrix}\Rightarrow \mathbb{S}_1\subseteq\mathbb{S}_2①$$

$$\begin{bmatrix}0\\1\\0\end{bmatrix}=\frac{1}{2}\begin{bmatrix}0\\2\\1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}0\\0\\-1\end{bmatrix},\begin{bmatrix}0\\1\\1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}0\\2\\1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}0\\0\\-1\end{bmatrix}\Rightarrow \mathbb{S}_2\subseteq\mathbb{S}_1②$$

①&② $\Rightarrow \mathbb{S}_1=\mathbb{S}_2$

Definition: Orthonormal base vectors $\underline{\boldsymbol{w}}_1, ..., \underline{\boldsymbol{2}}_n$ have the property:

$$\underline{\boldsymbol{w}}_j^H\underline{\boldsymbol{w}}_k=\begin{cases}1, & i=k\\ 0, & else\end{cases}$$

Theroem: orthonormal vectors are LID

Proof: $a_1\underline{\boldsymbol{w}}_1+a_2\underline{\boldsymbol{w}}_2+...+a_n\underline{\boldsymbol{w}}_n=\underline{\boldsymbol{0}} \quad |\underline{\boldsymbol{w}}_1^H\cdot$

$$a_1\underbrace{\underline{\boldsymbol{w}}_1^H\underline{\boldsymbol{w}}_1}_{1}+a_2\underbrace{\underline{\boldsymbol{w}}_1^H\underline{\boldsymbol{w}}_2}_{0}+...+a_n\underbrace{\underline{\boldsymbol{w}}_1^H\underline{\boldsymbol{w}}_n}_{0}=\underbrace{\underline{\boldsymbol{w}}_1^H\underline{\boldsymbol{0}}}_{0}\Rightarrow a_1=0$$

Repeat: $\underline{\boldsymbol{w}}_2^H : a_2=0 \quad ... \quad \underline{\boldsymbol{w}}_n^H : a_n=0$ □

Theorem: Let $\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n \in \mathbb{C}^n$ be n LID vectors of $\mathbb{C}^n$, then $\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)=\mathrm{Sp}(\underline{\boldsymbol{e}}_1, ..., \underline{\boldsymbol{e}}_n)\in\mathbb{C}^n$

with: $\underline{\boldsymbol{e}}_1=\begin{bmatrix}1\\0\\0\\ \vdots\end{bmatrix},\underline{\boldsymbol{e}}_2=\begin{bmatrix}0\\1\\0\\ \vdots\end{bmatrix},...,\underline{\boldsymbol{e}}_n=\begin{bmatrix}0\\ \vdots\\0\\1\end{bmatrix}\in\mathbb{C}^n$

Proof: $\underline{\boldsymbol{v}}_i=\sum_{j=1}^{n}a_{ji}\underline{\boldsymbol{e}}_j\Rightarrow \mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)\subseteq\mathbb{C}^n$

if $\exists\underline{\boldsymbol{s}}\in\mathbb{C}^n:\underline{\boldsymbol{s}}\in\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$ then $(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n,\underline{\boldsymbol{s}})$ is LID $\underbrace{\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n,\underline{\boldsymbol{s}})}_{dim=n+1}\subseteq\underbrace{\mathbb{C}^n}_{dim=n(wrong)}$ □

Theorem: Every subspace has got orthonormal base-vectors

Proof: $\mathbb{S}=\mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$ with $\underline{\boldsymbol{v}}_i$ is LID

Orthogonal vectors:

$\underline{\boldsymbol{u}}_1=\underline{\boldsymbol{v}}_1$

$\underline{\boldsymbol{u}}_2=\underline{\boldsymbol{v}}_2-\frac{\underline{\boldsymbol{u}}_1^H\underline{\boldsymbol{v}}_2}{\underline{\boldsymbol{u}}_1^H\underline{\boldsymbol{u}}_1}\underline{\boldsymbol{u}}_1$

$\underline{\boldsymbol{u}}_1^H\underline{\boldsymbol{u}}_2=\underline{\boldsymbol{u}}_1^H\underline{\boldsymbol{v}}_2-\underline{\boldsymbol{u}}_1^H\underline{\boldsymbol{v}}_2=0 \quad \underline{\boldsymbol{u}}_2^H\underline{\boldsymbol{u}}_1=0$

$\underline{\boldsymbol{u}}_3=\underline{\boldsymbol{v}}_3-\frac{\underline{\boldsymbol{u}}_1^H\underline{\boldsymbol{v}}_3}{\underline{\boldsymbol{u}}_1^H\underline{\boldsymbol{u}}_1}\underline{\boldsymbol{u}}_1-\frac{\underline{\boldsymbol{u}}_2^H\underline{\boldsymbol{v}}_3}{\underline{\boldsymbol{u}}_2^H\underline{\boldsymbol{u}}_2}\underline{\boldsymbol{u}}_2$

$\underline{\boldsymbol{u}}_k = \underline{\boldsymbol{v}}_k - \sum\limits_{m=1}^{k-1} \frac{\underline{\boldsymbol{u}}_m^H \underline{\boldsymbol{v}}_k}{\underline{\boldsymbol{u}}_m^H \underline{\boldsymbol{u}}_m} \underline{\boldsymbol{u}_m}, \quad \text{with} \quad 2 \leq k \leq n$

normalized vectors $\rightarrow$ orthonormal vectors:

$\underline{\boldsymbol{w}}_i = \frac{\underline{\boldsymbol{u}}_i}{\sqrt{\underline{\boldsymbol{u}}_i^H \underline{\boldsymbol{u}}_i}}$

$\Rightarrow \underline{\boldsymbol{w}}_i^H \underline{\boldsymbol{w}}_k = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}$

$\mathrm{Sp}(\underline{\boldsymbol{w}}_1, ..., \underline{\boldsymbol{w}}_n) = \mathrm{Sp}(\underline{\boldsymbol{v}}_1, ..., \underline{\boldsymbol{v}}_n)$ □

2.4.3 Matrices

$\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \in \mathbb{C}^{m \times n} \quad \text{with} \quad \text{m rows and n columns}$

$A_{ij} = (\underset{\sim}{\boldsymbol{A}})_{ij}$ element of $\underset{\sim}{\boldsymbol{A}}$ at the i-th row and j-th column

$\underset{\sim}{\boldsymbol{A}}^* = \begin{bmatrix} A_{11}^* & \cdots & A_{1n}^* \\ \vdots & & \vdots \\ A_{m1}^* & \cdots & A_{mn}^* \end{bmatrix} \in \mathbb{C}^{m \times n}$

$\underset{\sim}{\boldsymbol{A}}^T = \begin{bmatrix} A_{11} & \cdots & A_{m1} \\ \vdots & & \vdots \\ A_{1n} & \cdots & A_{mn} \end{bmatrix} \in \mathbb{C}^{n \times m}$

Hermite-ian matrix: $(\underset{\sim}{\boldsymbol{A}}^*)^T = (\underset{\sim}{\boldsymbol{A}}^T)^* = \underset{\sim}{\boldsymbol{A}}^H$

Frobenius norm: $\left\|\underset{\sim}{\boldsymbol{A}}\right\|_F := \sqrt{\sum\limits_{k=1}^{m} \sum\limits_{p=1}^{n} |A_{kp}|^2}$

Zero matrix: $\underset{\sim m \times n}{\boldsymbol{0}} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in \mathbb{C}^{m \times n}, \quad \underset{\sim m}{\boldsymbol{0}} \stackrel{!}{=} \underset{\sim m \times m}{\boldsymbol{0}}$

Identity matrix: $\underset{\sim m}{\boldsymbol{I}} = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 1 & \\ \vdots & & \ddots \end{bmatrix}$

Matrix multiplication: $\underset{\sim}{\boldsymbol{C}} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}} \iff C_{ik} = \sum\limits_t A_{it} \cdot B_{tk}$

$\underbrace{\begin{bmatrix} \underline{\boldsymbol{a}}_1 & \cdots & \underline{\boldsymbol{a}}_p \end{bmatrix}}_{\underset{\sim}{\boldsymbol{A}}} \underbrace{\begin{bmatrix} \underline{\boldsymbol{\beta}}_1^H \\ \vdots \\ \underline{\boldsymbol{\beta}}_p^H \end{bmatrix}}_{\underset{\sim}{\boldsymbol{B}}} = \underline{\boldsymbol{a}}_1 \underline{\boldsymbol{\beta}}_1^H + ... + \underline{\boldsymbol{a}}_p \underline{\boldsymbol{\beta}}_p^H = \underset{\sim}{\boldsymbol{C}}$

$\underbrace{\begin{bmatrix} \underline{\boldsymbol{\alpha}}_1^H \\ \vdots \\ \underline{\boldsymbol{\alpha}}_m^H \end{bmatrix}}_{\underset{\sim}{\boldsymbol{A}}} \underbrace{\begin{bmatrix} \underline{\boldsymbol{b}}_1 & \cdots & \underline{\boldsymbol{b}}_n \end{bmatrix}}_{\underset{\sim}{\boldsymbol{B}}} = \begin{bmatrix} \underline{\boldsymbol{\alpha}}_1^H \underline{\boldsymbol{b}}_1 & \cdots & \underline{\boldsymbol{\alpha}}_1^H \underline{\boldsymbol{b}}_n \\ \vdots & & \vdots \\ \underline{\boldsymbol{\alpha}}_m^H \underline{\boldsymbol{b}}_1 & \cdots & \underline{\boldsymbol{\alpha}}_m^H \underline{\boldsymbol{b}}_n \end{bmatrix} = \underset{\sim}{\boldsymbol{C}}$

Multiplication with sub matrices:

$$\begin{bmatrix}\utilde{\boldsymbol{A}} & \utilde{\boldsymbol{C}}\\ \utilde{\boldsymbol{B}} & \utilde{\boldsymbol{D}}\end{bmatrix}\begin{bmatrix}\utilde{\boldsymbol{E}} & \utilde{\boldsymbol{G}}\\ \utilde{\boldsymbol{F}} & \utilde{\boldsymbol{H}}\end{bmatrix}=\begin{bmatrix}\utilde{\boldsymbol{A}}\utilde{\boldsymbol{E}}+\utilde{\boldsymbol{C}}\utilde{\boldsymbol{F}} & \utilde{\boldsymbol{A}}\utilde{\boldsymbol{G}}+\utilde{\boldsymbol{C}}\utilde{\boldsymbol{H}}\\ \utilde{\boldsymbol{B}}\utilde{\boldsymbol{E}}+\utilde{\boldsymbol{D}}\utilde{\boldsymbol{F}} & \utilde{\boldsymbol{B}}\utilde{\boldsymbol{G}}+\utilde{\boldsymbol{D}}\utilde{\boldsymbol{H}}\end{bmatrix}$$

example:

$\utilde{\boldsymbol{M}}=\begin{bmatrix}\utilde{\boldsymbol{I}}_m & \utilde{\boldsymbol{F}}\end{bmatrix};\utilde{\boldsymbol{Q}}=\begin{bmatrix}-\utilde{\boldsymbol{F}}^H & \utilde{\boldsymbol{I}}_m\end{bmatrix}$

$\utilde{\boldsymbol{M}}\utilde{\boldsymbol{Q}}^H=\begin{bmatrix}\utilde{\boldsymbol{I}}_m & \utilde{\boldsymbol{F}}\end{bmatrix}\begin{bmatrix}-\utilde{\boldsymbol{F}}\\ \utilde{\boldsymbol{I}}_m\end{bmatrix}=\utilde{\boldsymbol{I}}_m(-\utilde{\boldsymbol{F}})+\utilde{\boldsymbol{F}}\utilde{\boldsymbol{I}}_m=\utilde{\boldsymbol{F}}-\utilde{\boldsymbol{F}}=\utilde{\boldsymbol{0}}$

Identity matrix

$\utilde{\boldsymbol{I}}\utilde{\boldsymbol{A}}=\utilde{\boldsymbol{A}}\utilde{\boldsymbol{I}}=\utilde{\boldsymbol{A}}$

$\utilde{\boldsymbol{A}}^H(\utilde{\boldsymbol{A}}^H)^{-1}=\utilde{\boldsymbol{I}}$

If $\utilde{\boldsymbol{B}}\in\mathbb{C}^{m\times n}$ has n orthonormal columns then $\utilde{\boldsymbol{B}}^H\utilde{\boldsymbol{B}}=\utilde{\boldsymbol{I}}$

Multiplication with a vector:

$$\utilde{\boldsymbol{A}}\underline{\boldsymbol{x}}=\begin{bmatrix}\underline{\boldsymbol{a}}_1 & \dots & \underline{\boldsymbol{a}}_n\end{bmatrix}\begin{bmatrix}x_1\\ \dots\\ x_n\end{bmatrix}=\underline{\boldsymbol{a}}_1x_1+\dots+\underline{\boldsymbol{a}}_nx_n$$

In general: $\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}}\neq\utilde{\boldsymbol{B}}\utilde{\boldsymbol{A}}$

- $(\utilde{\boldsymbol{A}}+\utilde{\boldsymbol{B}})^2=(\utilde{\boldsymbol{A}}+\utilde{\boldsymbol{B}})(\utilde{\boldsymbol{A}}+\utilde{\boldsymbol{B}})=\utilde{\boldsymbol{A}}^2+\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}}+\utilde{\boldsymbol{B}}\utilde{\boldsymbol{A}}+\utilde{\boldsymbol{B}}^2$
- $(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}})\utilde{\boldsymbol{C}}=\utilde{\boldsymbol{A}}(\utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}})$
- $(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}})^T=\utilde{\boldsymbol{B}}^T\utilde{\boldsymbol{A}}^T$
- $(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}})^H=\utilde{\boldsymbol{B}}^H\utilde{\boldsymbol{A}}^H$

Definition: Trace: $\operatorname{tr}\utilde{\boldsymbol{A}}=\sum\limits_k A_{kk}$

- $\operatorname{tr}(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}})=\operatorname{tr}(\utilde{\boldsymbol{B}}\utilde{\boldsymbol{A}})$
- $\operatorname{tr}(\utilde{\boldsymbol{A}}\underbrace{\utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}}}_{\utilde{\boldsymbol{M}}})=\operatorname{tr}(\utilde{\boldsymbol{M}}\utilde{\boldsymbol{A}})=\operatorname{tr}(\utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}}\utilde{\boldsymbol{A}})$
- $\operatorname{tr}(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}})=\operatorname{tr}(\utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}}\utilde{\boldsymbol{A}})=\operatorname{tr}(\utilde{\boldsymbol{C}}\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}})$
- $\operatorname{tr}(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{A}}^H)=\operatorname{tr}(\utilde{\boldsymbol{A}}^H\utilde{\boldsymbol{A}})$

$$\operatorname{tr}(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{A}}^H)=\operatorname{tr}(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{A}}^H)=\operatorname{tr}\begin{bmatrix}\underline{\boldsymbol{a}}_1^H\\ \dots\\ \underline{\boldsymbol{a}}_n^H\end{bmatrix}\begin{bmatrix}\underline{\boldsymbol{a}}_1 & \dots & \underline{\boldsymbol{a}}_n\end{bmatrix}=\operatorname{tr}\begin{bmatrix}\underline{\boldsymbol{a}}_1^H\underline{\boldsymbol{a}}_1 & & X\\ & \ddots & \\ X & & \underline{\boldsymbol{a}}_n^H\underline{\boldsymbol{a}}_n\end{bmatrix}=\sum_{k=1}^n\underline{\boldsymbol{a}}_k^H\underline{\boldsymbol{a}}_k=\sum_{k=1}^n||\underline{\boldsymbol{a}}_k||_2^2=$$
$$\sum_{k=1}^n\sum_{p=1}^n|a_{kp}|^2=||\utilde{\boldsymbol{A}}||_F^2$$

Frobenius norm: $\operatorname{tr}(\utilde{\boldsymbol{A}}^H\utilde{\boldsymbol{A}})=\operatorname{tr}(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{A}}^H)=||\utilde{\boldsymbol{A}}||_F^2$

Definition: $\utilde{\boldsymbol{A}}\in\mathbb{C}^{n\times n}$ is invertable if and only if:
$\exists\utilde{\boldsymbol{B}}:\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}}=\utilde{\boldsymbol{B}}\utilde{\boldsymbol{A}}=\utilde{\boldsymbol{I}}_n$, then $\utilde{\boldsymbol{B}}=\utilde{\boldsymbol{A}}^{-1}$ is called inverse of $\utilde{\boldsymbol{A}}$

Theorem: If $\utilde{\boldsymbol{A}}$ is invertable then $\utilde{\boldsymbol{A}}^{-1}$ is unique

Proof: $\utilde{\boldsymbol{C}}$ is another inverse of $\utilde{\boldsymbol{A}}$

$\utilde{\boldsymbol{C}} = \utilde{\boldsymbol{C}}\utilde{\boldsymbol{I}} = \utilde{\boldsymbol{C}}\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}} = (\utilde{\boldsymbol{C}}\utilde{\boldsymbol{A}})\utilde{\boldsymbol{B}} = \utilde{\boldsymbol{B}} \Rightarrow \utilde{\boldsymbol{I}}\utilde{\boldsymbol{B}} = \utilde{\boldsymbol{B}} = \utilde{\boldsymbol{C}}$ □

Theorem: $(\utilde{\boldsymbol{A}}^H)^{-1} = (\utilde{\boldsymbol{A}}^{-1})^H = \utilde{\boldsymbol{A}}^{-H}$

Proof: $\utilde{\boldsymbol{A}}\utilde{\boldsymbol{A}}^{-1} = \utilde{\boldsymbol{I}} \quad |(\cdot)^H$

$(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{A}}^{-1})^H = (\utilde{\boldsymbol{A}}^{-1})^H\utilde{\boldsymbol{A}}^H = \utilde{\boldsymbol{I}} = (\utilde{\boldsymbol{A}}^H)^{-1}\utilde{\boldsymbol{A}}^H \quad | \cdot (\utilde{\boldsymbol{A}}^H)^{-1}$

$$(\utilde{\boldsymbol{A}}^{-1})^H \underbrace{\utilde{\boldsymbol{A}}^H(\utilde{\boldsymbol{A}}^H)^{-1}}_{\utilde{\boldsymbol{I}}} = (\utilde{\boldsymbol{A}}^H)^{-1} \underbrace{\utilde{\boldsymbol{A}}^H(\utilde{\boldsymbol{A}}^H)^{-1}}_{\utilde{\boldsymbol{I}}}$$

$\Rightarrow (\utilde{\boldsymbol{A}}^{-1})^H = (\utilde{\boldsymbol{A}}^H)^{-1}$ □

Theorem: if $\utilde{\boldsymbol{A}}^{-1}$ exist, then columns of $\utilde{\boldsymbol{A}}$ are LID

Proof: $\underbrace{\utilde{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}}_{\underline{\boldsymbol{a}}_1 x_1 + \ldots \underline{\boldsymbol{a}}_n x_n = \underline{\boldsymbol{0}}} \quad \Rightarrow \underbrace{\underline{\boldsymbol{x}} = \utilde{\boldsymbol{A}}^{-1}\underline{\boldsymbol{0}} = \underline{\boldsymbol{0}}}_{x_1 = x_2 = \ldots = x_n = 0}$

$\Rightarrow \mathrm{Sp}(\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_n) = \mathbb{C}^n$

$\exists \underline{\boldsymbol{b}}_i : \utilde{\boldsymbol{A}}\underline{\boldsymbol{b}}_i = \underline{\boldsymbol{e}}_i \in \mathbb{C}^n$

$\Rightarrow \utilde{\boldsymbol{A}} \underbrace{\left[\underline{\boldsymbol{b}}_1, ..., \underline{\boldsymbol{b}}_n\right]}_{\utilde{\boldsymbol{B}}} = \left[\underline{\boldsymbol{e}}_1, ..., \underline{\boldsymbol{e}}_n\right] = \utilde{\boldsymbol{I}}$

$\exists \utilde{\boldsymbol{B}} : \utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}} = \utilde{\boldsymbol{I}}$ □

2.4.4 Determinate

$\det \utilde{\boldsymbol{A}} : \mathbb{C}^{n \times n} \mapsto \mathbb{C}$

Property 1: $n = 1 : \det A = A$

Property 2: If $\utilde{\boldsymbol{A}}$ has LD columns, then $\det A = 0$

Property 3: $\det \utilde{\boldsymbol{A}}$ is linear in colums of $\utilde{\boldsymbol{A}}$

1: $\det[\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_{i-1}, \lambda\underline{\boldsymbol{a}}_i, \underline{\boldsymbol{a}}_{i+1}, ..., \underline{\boldsymbol{a}}_n] = \lambda \det \utilde{\boldsymbol{A}}$

2: $\det[\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_{i-1}, \underline{\boldsymbol{a}}_i, \underline{\boldsymbol{a}}_{i+1}, ..., \underline{\boldsymbol{a}}_n]$
$+ \det[\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_{i-1}, \underline{\boldsymbol{b}}_i, \underline{\boldsymbol{a}}_{i+1}, ..., \underline{\boldsymbol{a}}_n]$
$= \det[\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_{i-1}, \underline{\boldsymbol{a}}_i + \underline{\boldsymbol{b}}_i, \underline{\boldsymbol{a}}_{i+1}, ..., \underline{\boldsymbol{a}}_n]$

Theorems:

1. $\det \utilde{\boldsymbol{A}}$ is unique

2. $\det \utilde{\boldsymbol{A}} = \sum_{i=1}^{n} (-1)^{i+j} A_{ij} \det \utilde{\boldsymbol{A}}^{(i,j)}$; for any j

   $\utilde{\boldsymbol{A}}^{(i,j)} \in \mathbb{C}^{(n-1)\times(n-1)}$: $\utilde{\boldsymbol{A}}$ with i-th row and j-th columns removed

3. $\det \utilde{\boldsymbol{I}}_n = 1$

4. $\det(\utilde{\boldsymbol{A}}\utilde{\boldsymbol{B}}) = \det(\utilde{\boldsymbol{A}}) \cdot \det(\utilde{\boldsymbol{B}})$

5. $\det(\utilde{\boldsymbol{B}}^{-1}) = (\det \utilde{\boldsymbol{B}})^{-1}$

6. $\det \utilde{\boldsymbol{A}} = \det \utilde{\boldsymbol{A}}^T$

- without proof

example:

$\det a = a, a \in \mathbb{C}$

$\det \begin{bmatrix} a & c \\ b & d \end{bmatrix} = a\det(d) + (-1)b\det(c) = ad - bc$

2.4.5 Eigenvalues

Definition: $\lambda \in \mathbb{C}$ is an eigenvalue (Eval) of $\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{n\times n}$ if $\exists \underline{\boldsymbol{b}} \neq \underline{\boldsymbol{0}}$ such that $(\underset{\sim}{\boldsymbol{A}} - \lambda \underset{\sim}{\boldsymbol{I}})\underline{\boldsymbol{b}} = \underline{\boldsymbol{0}}$, then $\underline{\boldsymbol{b}}$ is a eigenvecotr (Evec) of $\underset{\sim}{\boldsymbol{A}}$ for the eigenvalue λ

Properties:

- **Eigenvalue** $\det(\underset{\sim}{\boldsymbol{A}} - \lambda\underset{\sim}{\boldsymbol{I}}) = 0$
- $\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{b}} = \lambda\underline{\boldsymbol{b}}$
- given n LID eigenvectors $(\underline{\boldsymbol{b}}_1, ..., \underline{\boldsymbol{b}}_n)$ of $\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{n\times n}$ of the eigenvalues $(\lambda_1, ...\lambda_n)$ then $\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1}$
- $\underset{\sim}{\boldsymbol{B}} = \begin{bmatrix} \underline{\boldsymbol{b}}_1 & ... & \underline{\boldsymbol{b}}_n \end{bmatrix} \in \mathbb{C}^{n\times n}$
- $\underset{\sim}{\boldsymbol{\Lambda}} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \in \mathbb{C}^{n\times n}$
- $\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}} = \begin{bmatrix} \underline{\boldsymbol{b}}_1 & ... & \underline{\boldsymbol{b}}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

  $= \begin{bmatrix} \underbrace{\underline{\boldsymbol{b}}_1\lambda_1}_{\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{b}}_1} & ... & \underbrace{\underline{\boldsymbol{b}}_n\lambda_n}_{\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{b}}_n} \end{bmatrix} = \underset{\sim}{\boldsymbol{A}} \begin{bmatrix} \underline{\boldsymbol{b}}_1 & ... & \underline{\boldsymbol{b}}_n \end{bmatrix} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}}$
- **Eigenvalue Decomposition, EVD**: $\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1}$

Note:

- $\mathrm{tr}(\underset{\sim}{\boldsymbol{A}}) = \mathrm{tr}(\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1}) = \mathrm{tr}(\underbrace{\underset{\sim}{\boldsymbol{B}}^{-1}\underset{\sim}{\boldsymbol{B}}}_{\underset{\sim}{\boldsymbol{I}}}\underset{\sim}{\boldsymbol{\Lambda}}) = \mathrm{tr}(\underset{\sim}{\boldsymbol{\Lambda}}) = \sum_{k=1}^{n} \lambda_k$
- $\det\underset{\sim}{\boldsymbol{A}} = \det\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1} = \det\underset{\sim}{\boldsymbol{B}}\det\underset{\sim}{\boldsymbol{\Lambda}}\det\underset{\sim}{\boldsymbol{B}}^{-1}$

  $= det\underset{\sim}{\boldsymbol{B}}\det\underset{\sim}{\boldsymbol{\Lambda}}\frac{1}{\det\underset{\sim}{\boldsymbol{B}}} = \det\underset{\sim}{\boldsymbol{\Lambda}} = \prod_{k=1}^{n} \lambda_k$
- $\underset{\sim}{\boldsymbol{A}}^k = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}...\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underbrace{\underset{\sim}{\boldsymbol{B}}^{-1}\underset{\sim}{\boldsymbol{B}}}_{\underset{\sim}{\boldsymbol{I}}}\underset{\sim}{\boldsymbol{\Lambda}}\underbrace{\underset{\sim}{\boldsymbol{B}}^{-1}\underset{\sim}{\boldsymbol{B}}}_{\underset{\sim}{\boldsymbol{I}}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1}...\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}^k\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{B}}\begin{bmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{bmatrix}\underset{\sim}{\boldsymbol{B}}^{-1}$

Function of matrices

Exponential function:

$$e^{\underset{\sim}{\boldsymbol{A}}} \overset{\text{Taylor}}{=} \underset{\sim}{\boldsymbol{I}} + \underset{\sim}{\boldsymbol{A}} + \frac{1}{!2}\underset{\sim}{\boldsymbol{A}}^2 + \frac{1}{!3}\underset{\sim}{\boldsymbol{A}}^3 + \frac{1}{!4}\underset{\sim}{\boldsymbol{A}}^4 + ... = \underbrace{\underset{\sim}{\boldsymbol{I}}}_{\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{I}}\underset{\sim}{\boldsymbol{B}}^{-1}} + \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1} + \frac{1}{!2}\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}^2\underset{\sim}{\boldsymbol{B}}^{-1} + \frac{1}{!3}\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}^3\underset{\sim}{\boldsymbol{B}}^{-1} + ...$$

$$= \underset{\sim}{\boldsymbol{B}}(\underset{\sim}{\boldsymbol{I}} + \underset{\sim}{\boldsymbol{\Lambda}} + \frac{1}{!2}\underset{\sim}{\boldsymbol{\Lambda}}^2 + \frac{1}{!3}\underset{\sim}{\boldsymbol{\Lambda}}^3 + ...)\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{B}} \begin{bmatrix} 1+\lambda_1+\frac{1}{2}\lambda_1^2... & & \\ & \ddots & \\ & & 1+\lambda_k+\frac{1}{2}\lambda_k^2... \end{bmatrix} \underset{\sim}{\boldsymbol{B}}^{-1}$$

$$= \underset{\sim}{\boldsymbol{B}} \begin{bmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_k} \end{bmatrix} \underset{\sim}{\boldsymbol{B}}^{-1}$$

Exponential Cosine:

$$\cos(\underset{\sim}{\boldsymbol{A}}) \overset{\text{Taylor}}{=} \underset{\sim}{\boldsymbol{I}} - \frac{1}{2}\underset{\sim}{\boldsymbol{A}}^2 + \frac{1}{24}\underset{\sim}{\boldsymbol{A}}^4 - \frac{1}{6}\underset{\sim}{\boldsymbol{A}}^6 + ...$$

$$= \underset{\sim}{\boldsymbol{B}}(\underset{\sim}{\boldsymbol{I}} - \frac{1}{2}\underset{\sim}{\boldsymbol{\Lambda}}^2 + \frac{1}{24}\underset{\sim}{\boldsymbol{\Lambda}}^4 - \frac{1}{6}\underset{\sim}{\boldsymbol{\Lambda}}^6 + ...)\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{B}} \begin{bmatrix} \cos\lambda_1 & \\ & \\ & \cos\lambda_n \end{bmatrix} \underset{\sim}{\boldsymbol{B}}^{-1}$$

Exponential Sine:

$$\sin(\underset{\sim}{\boldsymbol{A}}) \overset{\text{Taylor}}{=} \underset{\sim}{\boldsymbol{A}} - \frac{1}{6}\underset{\sim}{\boldsymbol{A}}^3 + \frac{1}{120}\underset{\sim}{\boldsymbol{A}}^5 - ...$$

$$= \underset{\sim}{\boldsymbol{B}}(\underset{\sim}{\boldsymbol{\Lambda}} - \frac{1}{6}\underset{\sim}{\boldsymbol{\Lambda}}^3 + \frac{1}{120}\underset{\sim}{\boldsymbol{\Lambda}}^5 - ...)\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{B}} \begin{bmatrix} \sin\lambda_1 & \\ & \\ & \sin\lambda_n \end{bmatrix} \underset{\sim}{\boldsymbol{B}}^{-1}$$

Homework 1:

$$\cos\left(\begin{bmatrix} \frac{\pi}{2} & \frac{\pi}{2} \\ \frac{\pi}{2} & \frac{\pi}{2} \end{bmatrix}\right) = \cos\underset{\sim}{\boldsymbol{A}}$$

$$\text{with:}\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \pi \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\cos\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \cos 0 & 0 \\ 0 & \cos\pi \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

Homework 2:

$$\text{with:}\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \frac{\pi}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\cos^2(\underset{\sim}{\boldsymbol{A}}) + \sin^2(\underset{\sim}{\boldsymbol{A}}) = \left(\cos\begin{bmatrix} \frac{\pi}{6} & \frac{\pi}{6} \\ \frac{\pi}{6} & \frac{\pi}{6} \end{bmatrix}\right)^2 + \left(\sin\begin{bmatrix} \frac{\pi}{6} & \frac{\pi}{6} \\ \frac{\pi}{6} & \frac{\pi}{6} \end{bmatrix}\right)^2$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \left(\begin{bmatrix} \cos^2(0) & 0 \\ 0 & \cos^2(\frac{\pi}{3}) \end{bmatrix} + \begin{bmatrix} \sin^2(0) & 0 \\ 0 & \sin^2(\frac{\pi}{3}) \end{bmatrix}\right) \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{I}}\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{I}}$$

Warning: EVD does not always exist

Example: $\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$; $\det\left[\underset{\sim}{\boldsymbol{A}} - \lambda\underset{\sim}{\boldsymbol{I}}\right] = 0$

Eval:

$$\det \begin{bmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{bmatrix} = (1-\lambda)^2 = 0; \quad \lambda_1 = \lambda_2 = 1$$

Evec:

$$(\underset{\sim}{\boldsymbol{A}} - \lambda_{1/2}\underset{\sim}{\boldsymbol{I}})\underline{\boldsymbol{b}} = \underline{\boldsymbol{0}} \Rightarrow \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \underline{\boldsymbol{b}}_{1/2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

only one LID eigenvector of $\underset{\sim}{\boldsymbol{A}}$

However:

$\boxed{\textbf{Theorem: } \text{Let } \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{A}}^H \in \mathbb{C}^{n\times n} \text{ with k-fold eigenvalue } \mu, \text{ then there are k LID eigenvectors for } \mu.}$

Proof:

- $p(\lambda) := \det(\underset{\sim}{\boldsymbol{A}} - \lambda\underset{\sim}{\boldsymbol{I}}) = (\mu-\lambda)^k \cdot q(\lambda)$
- $q(\mu) \neq 0 : \mu$ is no root of $q(\lambda)$
- $\mathbb{E}_\mu := \left\{\underline{\boldsymbol{v}} \middle| \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}} = \mu\underline{\boldsymbol{v}}\right\}$ Eigenspace (all eigenvectors of $\mu + \underline{\boldsymbol{0}}$)

  $\mathbb{E}_\mu = \mathrm{Sp}(\underbrace{\underline{\boldsymbol{b}}_1, ..., \underline{\boldsymbol{b}}_l}_{\text{orthonormal}}) \quad l$ LID eigenvectors

  $\dim \mathbb{E}_\mu = l$

Aim: want to show $l = k$

$$\underset{\sim}{\boldsymbol{B}} \left[\underbrace{\underline{\boldsymbol{b}}_1, ..., \underline{\boldsymbol{b}}_l, \overbrace{\underline{\boldsymbol{b}}_{l+1}, ..., \underline{\boldsymbol{b}}_n}^{\text{invented}}}_{\text{orthonormal}}\right] \in \mathbb{C}^{n\times n}$$

$$\underset{\sim}{\boldsymbol{B}}^H\underset{\sim}{\boldsymbol{B}} = \begin{bmatrix} \underline{\boldsymbol{b}}_1^H \\ . \\ \underline{\boldsymbol{b}}_n^H \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{b}}_1 & ... & \underline{\boldsymbol{b}}_n \end{bmatrix} = \underset{\sim}{\boldsymbol{I}} \quad \text{length of each vector is 1}$$

$$\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{B}}^H$$

$$\left.\begin{aligned} \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}} &= \mu\underline{\boldsymbol{v}} \\ \underline{\boldsymbol{v}} &= \underset{\sim}{\boldsymbol{B}}\underline{\boldsymbol{w}} \\ \underline{\boldsymbol{w}} &= \underset{\sim}{\boldsymbol{B}}^H\underline{\boldsymbol{v}} \end{aligned}\right\} \iff \underset{\sim}{\boldsymbol{B}}^H\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}}\underline{\boldsymbol{w}} = \mu\underline{\boldsymbol{w}}$$

$$\underset{\sim}{\boldsymbol{B}}^H\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}} = \begin{bmatrix} \underline{\boldsymbol{b}}_1^H \\ . \\ \underline{\boldsymbol{b}}_n^H \end{bmatrix} \begin{bmatrix} \mu\underline{\boldsymbol{b}}_1 & ... & \mu\underline{\boldsymbol{b}}_l & \underbrace{* \quad * \quad *}_{Junk} \end{bmatrix}$$

$$= \left[\begin{array}{c|c} \overbrace{\begin{matrix} \mu & 0 & ... & 0 \\ 0 & \mu & ... & 0 \\ . & . & ... & . \\ 0 & 0 & ... & \mu \end{matrix}}^{l} & \underset{\sim}{\boldsymbol{0}} \\ \hline \underset{\sim}{\boldsymbol{0}} & \underset{\sim}{\boldsymbol{A}}_1 \end{array}\right]$$

because:

$(\underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}})^H = \underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{B}} = \underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}}, \quad \underset{\sim}{\boldsymbol{A}}^H = \underset{\sim}{\boldsymbol{A}}$

$\det(\underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}} - \lambda \underset{\sim}{\boldsymbol{I}}) \underset{\underset{\sim}{\boldsymbol{B}}^H = \underset{\sim}{\boldsymbol{B}}^{-1}}{=} \det(\underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}} - \lambda \underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{I}}\underset{\sim}{\boldsymbol{B}})$

$= \det(\underset{\sim}{\boldsymbol{B}}^H (\underset{\sim}{\boldsymbol{A}} - \lambda \underset{\sim}{\boldsymbol{I}})\underset{\sim}{\boldsymbol{B}}) = \det(\underset{\sim}{\boldsymbol{B}}^{-1} (\underset{\sim}{\boldsymbol{A}} - \lambda \underset{\sim}{\boldsymbol{I}})\underset{\sim}{\boldsymbol{B}})$

$= \det(\underset{\sim}{\boldsymbol{A}} - \lambda \underset{\sim}{\boldsymbol{I}}) = p(\lambda) = (\mu - \lambda)^k \cdot q(\lambda) = (\mu - \lambda)^l \cdot \underbrace{det(\underset{\sim}{\boldsymbol{A}} - \lambda \underset{\sim}{\boldsymbol{I}})}_{\varphi(\lambda)}$

Assumption: that μ is an eigenvalue of $\underset{\sim}{\boldsymbol{A}}_1$:

$$\underset{\sim}{\boldsymbol{A}}_1 \underline{\boldsymbol{c}} = \mu \underline{\boldsymbol{c}}, \quad \underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}} \begin{bmatrix} \mathbf{0}\} \, l \\ \underline{\boldsymbol{c}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \underbrace{\underset{\sim}{\boldsymbol{A}}_1 \underline{\boldsymbol{c}}}_{\mu \underline{\boldsymbol{c}}} \end{bmatrix} = \mu \begin{bmatrix} \mathbf{0} \\ \underline{\boldsymbol{c}} \end{bmatrix}$$

$\mu \begin{bmatrix} \mathbf{0} \\ \underline{\boldsymbol{c}} \end{bmatrix}$ is an eigenvector of $\underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{B}}$

$\underline{\boldsymbol{v}} = \underset{\sim}{\boldsymbol{B}} \begin{bmatrix} \mathbf{0} \\ \underline{\boldsymbol{c}} \end{bmatrix}$ is an eigenvector of $\underset{\sim}{\boldsymbol{A}}$

$\underline{\boldsymbol{v}} = \begin{bmatrix} \underline{\boldsymbol{b}}_1, ..., \underline{\boldsymbol{b}}_l, \underline{\boldsymbol{b}}_{l+1}, ... \underline{\boldsymbol{b}}_n \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \underline{\boldsymbol{c}} \end{bmatrix} = \begin{bmatrix} \underline{\boldsymbol{b}}_{l+1}, ... \underline{\boldsymbol{b}}_n \end{bmatrix} \underline{\boldsymbol{c}} \in \mathrm{Sp}(\underline{\boldsymbol{b}}_{l+1}, ... \underline{\boldsymbol{b}}_n)$

$\underline{\boldsymbol{v}}$ is eigenvector of $\underset{\sim}{\boldsymbol{A}}$: $\quad \underline{\boldsymbol{v}} \in \mathrm{Sp}(\underline{\boldsymbol{b}}_1, ... \underline{\boldsymbol{b}}_l)$

$\Rightarrow \underline{\boldsymbol{v}}^H \underline{\boldsymbol{v}} = \underline{\mathbf{0}}$, orthogonal

$\iff \underline{\boldsymbol{v}} = \underline{\mathbf{0}} \quad$ but $\underline{\mathbf{0}}$ is no eigenvector!

actualy μ is not a eigenvalue of $\underset{\sim}{\boldsymbol{A}}_1$

$p(\lambda) = (\mu - \lambda)^k \cdot \underbrace{q(\lambda)}_{q(\mu) \neq 0} = (\mu - \lambda)^l \cdot \underbrace{q(\lambda)}_{q(\mu) \neq 0} \quad l = k \quad \dim \mathbb{E}_\mu = k$ $\square$

Conclusion:

$\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{A}}^H \Rightarrow$ EVD of $\underset{\sim}{\boldsymbol{A}}$ exist

Theorem:

For $\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{A}}^H \in \mathbb{C}^{n \times n}$ then:

- all eigenvalues $\in \mathbb{R}$
- all eigenvectors are orthonormal

Proof:

$\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{A}}^H = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^{-1} = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^H = (\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^H)^H = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}^H\underset{\sim}{\boldsymbol{B}}^H$

$\Rightarrow \underset{\sim}{\boldsymbol{\Lambda}} = \underset{\sim}{\boldsymbol{\Lambda}}^H = (\underset{\sim}{\boldsymbol{\Lambda}}^T)^* = (\underset{\sim}{\boldsymbol{\Lambda}})^*; \quad \forall \lambda_i : \lambda_i = \lambda_i^*$

$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_1 = \lambda_1 \underline{\boldsymbol{v}}_1 \overset{\underline{\boldsymbol{v}}_2^H \cdot}{\iff} \underline{\boldsymbol{v}}_2^H \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_1 = \lambda_1 \underline{\boldsymbol{v}}_2^H \underline{\boldsymbol{v}}_1 \quad (1)$

$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_2 = \lambda_2 \underline{\boldsymbol{v}}_2 \overset{\underline{\boldsymbol{v}}_1^H \cdot}{\iff} \underline{\boldsymbol{v}}_1^H \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_2 = \lambda_2 \underline{\boldsymbol{v}}_1^H \underline{\boldsymbol{v}}_2 \overset{(\cdot)^H}{\iff} \underline{\boldsymbol{v}}_2^H \underbrace{\underset{\sim}{\boldsymbol{A}}}_{\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{A}}^H} \underline{\boldsymbol{v}}_1 = \lambda_2 \underline{\boldsymbol{v}}_2^H \underline{\boldsymbol{v}}_1 \quad (2)$

$(1) - (2) \Rightarrow \quad \underline{\boldsymbol{v}}_2^H \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}} - \underline{\boldsymbol{v}}_2^H \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}} = (\lambda_1 - \lambda_2)\underline{\boldsymbol{v}}_2^H \underline{\boldsymbol{v}}_1 \overset{!}{=} 0$

if $\lambda_1 \neq \lambda_2 \quad \Rightarrow \quad \underline{\boldsymbol{v}}_2^H \underline{\boldsymbol{v}}_1 = 0 \quad$ orthogonal

else $\lambda_1 = \lambda_2 = \lambda \quad$ 2-dimensional solution

$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_1 = \lambda \underline{\boldsymbol{v}}_1 \quad \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_2 = \lambda \underline{\boldsymbol{v}}_2$

Define: $\underline{\boldsymbol{w}}_1 = \underline{\boldsymbol{v}}_1; \quad \underline{\boldsymbol{w}}_2 = \underline{\boldsymbol{v}}_1 + a \cdot \underline{\boldsymbol{v}}_2 \quad with \quad a \neq 0$

$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{w}}_1 = \lambda \underline{\boldsymbol{w}}_1; \quad \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_1 + a\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}_2 = \lambda \underline{\boldsymbol{v}}_1 + a\lambda \underline{\boldsymbol{v}}_2 \iff \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{w}}_2 = \lambda \underline{\boldsymbol{w}}_2$
$\mathrm{Sp}(\underline{\boldsymbol{v}}_1, \underline{\boldsymbol{v}}_2) = \mathrm{Sp}(\underline{\boldsymbol{w}}_1, \underline{\boldsymbol{w}}_2)$
on condition on: $\underline{\boldsymbol{w}}_1^H \underline{\boldsymbol{w}}_2 \overset{!}{=} 0$
$\underline{\boldsymbol{w}}_1^H \underline{\boldsymbol{w}}_2 = \underline{\boldsymbol{v}}_1^H (\underline{\boldsymbol{v}}_1 + a \cdot \underline{\boldsymbol{v}}_2) \overset{!}{=} 0 \quad \Rightarrow \quad a = -\frac{\underline{\boldsymbol{v}}_1^H \underline{\boldsymbol{v}}_1}{\underline{\boldsymbol{v}}_1^H \underline{\boldsymbol{v}}_2}$
$\underline{\boldsymbol{v}}_1, \underline{\boldsymbol{v}}_2$ are orthogonal if a is definde as above $\square$

2.4.6 Gram-ian Matrix

Theorem:
$\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{C}}^H \in \mathbb{C}^{n \times n} \quad \Rightarrow$ all eigenvalues of $\underset{\sim}{\boldsymbol{A}}$ are bigger or equal to 0

Proof:
$\underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{C}}^H \underline{\boldsymbol{v}} = \mu \underline{\boldsymbol{v}} \quad |\underline{\boldsymbol{v}}^H \cdot$
$\underbrace{\underline{\boldsymbol{v}}^H \underset{\sim}{\boldsymbol{C}}}_{\underline{\boldsymbol{a}}^H} \underbrace{\underset{\sim}{\boldsymbol{C}}^H \underline{\boldsymbol{v}}}_{\underline{\boldsymbol{a}}} = \mu \underline{\boldsymbol{v}}^H \underline{\boldsymbol{v}}$
$\mu = \frac{\underline{\boldsymbol{a}}^H \underline{\boldsymbol{a}}}{\underline{\boldsymbol{v}}^H \underline{\boldsymbol{v}}} = \frac{||\underline{\boldsymbol{a}}||_2^2}{||\underline{\boldsymbol{v}}||_2^2} \geq 0$ $\square$
$\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{C}}^H = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^H$ because $\underset{\sim}{\boldsymbol{A}}$ is hermitian
$\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{B}}^H = \underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{B}} = \underset{\sim}{\boldsymbol{I}}, \quad \underset{\sim}{\boldsymbol{B}}$ is unitary and orthonormal

2.4.7 Singular Value Decomposition

Sigular value decomposition exists for every matrix $\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{m \times n}$
with:

- $\underset{\sim}{\boldsymbol{U}} \in \mathbb{C}^{m \times m}, \quad \underset{\sim}{\boldsymbol{U}}^{-1} = \underset{\sim}{\boldsymbol{U}}^H$, Unitary matrix
- $\underset{\sim}{\boldsymbol{V}} \in \mathbb{C}^{n \times n}, \quad \underset{\sim}{\boldsymbol{V}}^{-1} = \underset{\sim}{\boldsymbol{V}}^H$, Unitary matrix
- $\underset{\sim}{\boldsymbol{\Sigma}} \in \mathbb{R}_{0+}^{m \times n} \quad (\underset{\sim}{\boldsymbol{\Sigma}})_{i,j} = \begin{cases} 0 & i \neq j \\ s_i & i = j \end{cases}$, Diagonal Matrix
  all $s_i > 0$ are Singular values
- Note: Since $\underset{\sim}{\boldsymbol{U}}$ and $\underset{\sim}{\boldsymbol{V}}$ are unitary matrices we get the following properties:
  $\underset{\sim}{\boldsymbol{U}}^H \underset{\sim}{\boldsymbol{U}} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{U}}^H = \underset{\sim}{\boldsymbol{I}}$
  $\underset{\sim}{\boldsymbol{V}}^H \underset{\sim}{\boldsymbol{V}} = \underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{V}}^H = \underset{\sim}{\boldsymbol{I}}$

Singular Value Decomposition: $\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H$

Connection to EVD: - Calculation of SVD

Calculate $\underset{\sim}{\boldsymbol{\Sigma}}$ and $\underset{\sim}{\boldsymbol{U}}$

$$\left.\begin{array}{l} \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H = \underbrace{\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H}_{\underset{\sim}{\boldsymbol{A}}}\underbrace{\underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma}}^T\underset{\sim}{\boldsymbol{U}}^H}_{\underset{\sim}{\boldsymbol{A}}^H} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{\Sigma}}^T\underset{\sim}{\boldsymbol{U}}^H \\ \text{EVD: } \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H = \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{B}}^H \quad \text{with} \quad \underset{\sim}{\boldsymbol{B}}^H = \underset{\sim}{\boldsymbol{B}}^{-1}, \underset{\sim}{\boldsymbol{\Lambda}} \geq 0 \end{array}\right\} \begin{array}{c} \underset{\sim}{\boldsymbol{U}} = \underset{\sim}{\boldsymbol{B}} \\ \underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{\Sigma}}^T = \underset{\sim}{\boldsymbol{\Lambda}} \end{array}$$

Calculate $\underset{\sim}{\boldsymbol{\Sigma}}$ and $\underset{\sim}{\boldsymbol{V}}$

$$\left.\begin{array}{l} \underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}} = \underbrace{\underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma}}^T\underset{\sim}{\boldsymbol{U}}^H}_{\underset{\sim}{\boldsymbol{A}}^H}\underbrace{\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H}_{\underset{\sim}{\boldsymbol{A}}} = \underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma}}^T\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H \\ \text{EVD: } \underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{B}}'\underset{\sim}{\boldsymbol{\Lambda}}'\underset{\sim}{\boldsymbol{B}}'^H \quad \text{with} \quad \underset{\sim}{\boldsymbol{B}}'^H = \underset{\sim}{\boldsymbol{B}}'^{-1}, \underset{\sim}{\boldsymbol{\Lambda}}' \geq 0 \end{array}\right\} \begin{array}{c} \underset{\sim}{\boldsymbol{V}} = \underset{\sim}{\boldsymbol{B}}' \\ \underset{\sim}{\boldsymbol{\Sigma}}^T\underset{\sim}{\boldsymbol{\Sigma}} = \underset{\sim}{\boldsymbol{\Lambda}}' \end{array}$$

Zeors in EVD: (for $\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H$)

① $n \geq m$:

$$\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{\Sigma}}^T = \left[\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & \underset{\sim}{\boldsymbol{0}} \\ & & s_m & \end{array}\right] \left[\begin{array}{ccc} s_1 & & \\ & \ddots & \\ & & s_m \\ \hline & \underset{\sim}{\boldsymbol{0}} & \end{array}\right] = \begin{bmatrix} s_1^2 & & \\ & \ddots & \\ & & s_m^2 \end{bmatrix} \overset{!}{=} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{bmatrix}$$

② $n < m$:

$$\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{\Sigma}}^T = \left[\begin{array}{ccc} s_1 & & \\ & \ddots & \\ & & s_n \\ \hline & \underset{\sim}{\boldsymbol{0}} & \end{array}\right] \left[\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & \underset{\sim}{\boldsymbol{0}} \\ & & s_n & \end{array}\right] = \left[\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & \underset{\sim}{\boldsymbol{0}} \\ & & s_n & \\ \hline & \underset{\sim}{\boldsymbol{0}} & & \underset{\sim}{\boldsymbol{0}} \end{array}\right]$$

Number of Singular Values:

$\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H\underline{\boldsymbol{x}} \in \mathrm{Sp}(\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_n)$

$$\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{\Lambda}}\underbrace{\underset{\sim}{\boldsymbol{B}}^{-1}\underline{\boldsymbol{x}}}_{\underline{\boldsymbol{y}}} = \begin{bmatrix} \underline{\boldsymbol{b}}_1\lambda_1 & \cdots & \underline{\boldsymbol{b}}_m\lambda_m \end{bmatrix} \underline{\boldsymbol{y}} \quad \begin{cases} \in \mathrm{Sp}(\underline{\boldsymbol{b}}_1, ..., \underline{\boldsymbol{b}}_m) \\ \in \mathrm{Sp}(\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_n) \end{cases}$$

- $m > n$
- $\underline{\boldsymbol{a}}_1, ..., \underline{\boldsymbol{a}}_n$ are LID
- at most n eigenvalues can be $\neq 0$

$\Rightarrow$ at least $m - n$ eigenvalues are $= 0$

$s_i = \sqrt{\lambda_i} \quad i \in \{1, 2, ..., n\} \quad n < m$

General:

$s_i = \sqrt{\lambda_i} \quad i \in \{1, 2, ..., \min(n, m)\}$

$$\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}} = \underbrace{\underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma}}^T\underset{\sim}{\boldsymbol{U}}^H}_{\underset{\sim}{\boldsymbol{A}}^H}\underbrace{\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H}_{\underset{\sim}{\boldsymbol{A}}} = \underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma}}^T\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H$$

EVD:

$$\left.\begin{aligned} &\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{B}}' \underset{\sim}{\boldsymbol{\Lambda}}' \underset{\sim}{\boldsymbol{B}}'^H \\ &(\underset{\sim}{\boldsymbol{B}}')^{-1} = (\underset{\sim}{\boldsymbol{B}}')^H \\ &\underset{\sim}{\boldsymbol{\Lambda}}' > \underset{\sim}{\boldsymbol{0}} \end{aligned}\right\} \qquad \underset{\sim}{\boldsymbol{V}} = \underset{\sim}{\boldsymbol{B}}', \quad \underset{\sim}{\boldsymbol{\Sigma}}'\underset{\sim}{\boldsymbol{\Sigma}} = \underset{\sim}{\boldsymbol{\Lambda}}$$

$$\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H \underline{\boldsymbol{v}} = \lambda \underline{\boldsymbol{v}}$$

$$\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} \underbrace{\underset{\sim}{\boldsymbol{A}}^H \underline{\boldsymbol{v}}}_{\underline{\boldsymbol{w}}} = \lambda \underbrace{\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}}}_{\underline{\boldsymbol{w}}}$$

$$\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}} = \lambda \underline{\boldsymbol{w}} = \lambda' \underline{\boldsymbol{w}} \iff \lambda = \lambda'$$

Summary:

$\forall \underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{m\times n} \quad \exists$unitary$\underset{\sim}{\boldsymbol{U}}, \underset{\sim}{\boldsymbol{V}}$ that: $\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H$

$$\underset{\sim}{\boldsymbol{\Sigma}} = \left[\begin{array}{ccc|c} s_1 & & & \\ & \ddots & & \underset{\sim}{\boldsymbol{0}} \\ & & s_r & \\ \hline & \underset{\sim}{\boldsymbol{0}} & & \underset{\sim}{\boldsymbol{0}} \end{array}\right] \qquad \begin{aligned} &r \le min(m,n) \\ &r = 0 \iff \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{0}} \\ &r > 0 \quad \forall \underset{\sim}{\boldsymbol{A}} \ne \underset{\sim}{\boldsymbol{0}} \end{aligned}$$

Undermatrices of SVD:

r: number of positive Sigular Values:

$s_1 \ge s_2 \ge ... \ge s_r > 0 \qquad s_{r+1} = s_{r+2} = ...s_{\min(m,n)} = 0$

$$\underset{\sim}{\boldsymbol{A}} = \underbrace{\begin{bmatrix} \underset{\sim 1}{\boldsymbol{U}} & \underset{\sim 2}{\boldsymbol{U}} \end{bmatrix}}_{m\times m} \underbrace{\begin{bmatrix} \underset{\sim 1}{\boldsymbol{\Sigma}} & \underset{\sim}{\boldsymbol{0}} \\ \underset{\sim}{\boldsymbol{0}} & \underset{\sim}{\boldsymbol{0}} \end{bmatrix}}_{m\times n} \underbrace{\begin{bmatrix} \underset{\sim 1}{\boldsymbol{V}}^H \\ \underset{\sim 2}{\boldsymbol{V}}^H \end{bmatrix}}_{n\times n} = \underbrace{\underset{\sim 1}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim 1}{\boldsymbol{V}}^H}_{\text{foreshortend SVD}}$$

- $\underset{\sim 1}{\boldsymbol{U}} \in \mathbb{C}^{m\times r}$
- $\underset{\sim 2}{\boldsymbol{U}} \in \mathbb{C}^{m\times (m-r)}$
- $\underset{\sim 1}{\boldsymbol{\Sigma}} \in \mathbb{R}^{r\times r}$
- $\underset{\sim 1}{\boldsymbol{V}}^H \in \mathbb{C}^{r\times n}$
- $\underset{\sim 2}{\boldsymbol{V}}^H \in \mathbb{C}^{(n-r)\times n}$

Inverse:

$\underset{\sim 1}{\boldsymbol{\Sigma}}^{-1}$ exists, provided that $\underset{\sim 1}{\boldsymbol{\Sigma}}$ exist

$$\underset{\sim}{\boldsymbol{U}}^H\underset{\sim}{\boldsymbol{U}} = \underset{\sim}{\boldsymbol{I}} = \begin{bmatrix} \underset{\sim 1}{\boldsymbol{U}}^H \\ \underset{\sim 2}{\boldsymbol{U}}^H \end{bmatrix} \begin{bmatrix} \underset{\sim 1}{\boldsymbol{U}} & \underset{\sim 2}{\boldsymbol{U}} \end{bmatrix} = \left[\begin{array}{c|c} \underset{\sim 1}{\boldsymbol{U}}^H\underset{\sim 1}{\boldsymbol{U}} & \underset{\sim 1}{\boldsymbol{U}}^H\underset{\sim 2}{\boldsymbol{U}} \\ \hline \underset{\sim 2}{\boldsymbol{U}}^H\underset{\sim 1}{\boldsymbol{U}} & \underset{\sim 2}{\boldsymbol{U}}^H\underset{\sim 2}{\boldsymbol{U}} \end{array}\right] = \begin{bmatrix} \underset{\sim}{\boldsymbol{I}} & \underset{\sim}{\boldsymbol{0}} \\ \underset{\sim}{\boldsymbol{0}} & \underset{\sim}{\boldsymbol{I}} \end{bmatrix}$$

- $\underset{\sim 1}{\boldsymbol{U}}^H\underset{\sim 1}{\boldsymbol{U}} = \underset{\sim}{\boldsymbol{I}}$
- $\underset{\sim 1}{\boldsymbol{U}}^H\underset{\sim 2}{\boldsymbol{U}} = \underset{\sim}{\boldsymbol{0}}$

- $\underset{\sim}{\boldsymbol{U}}_2^H \underset{\sim}{\boldsymbol{U}}_1 = \underset{\sim}{\boldsymbol{0}}$
- $\underset{\sim}{\boldsymbol{U}}_2^H \underset{\sim}{\boldsymbol{U}}_2 = \underset{\sim}{\boldsymbol{I}}$

But $\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{U}}^H$ leads to:

$$\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{U}}^H = \begin{bmatrix} \underset{\sim}{\boldsymbol{U}}_1 & \underset{\sim}{\boldsymbol{U}}_2 \end{bmatrix} \begin{bmatrix} \underset{\sim}{\boldsymbol{U}}_1^H \\ \underset{\sim}{\boldsymbol{U}}_2^H \end{bmatrix} = \underset{\sim}{\boldsymbol{I}} \quad \Rightarrow \quad \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{U}}_1^H + \underset{\sim}{\boldsymbol{U}}_2 \underset{\sim}{\boldsymbol{U}}_2^H = \underset{\sim}{\boldsymbol{I}}$$

Definition:
The Image of a Matirx $im\underset{\sim}{\boldsymbol{A}}$ is the span of the colums of $\underset{\sim}{\boldsymbol{A}}$

$$im\underset{\sim}{\boldsymbol{A}} = \left\{ \underline{\boldsymbol{x}} \middle| \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{y}}, \quad \underline{\boldsymbol{y}} \in \mathbb{C}^{n\times 1} \right\}$$

Theorem:
$\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H$ then $im\underset{\sim}{\boldsymbol{A}} = im\underset{\sim}{\boldsymbol{U}}_1$

Proof:

$\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{m\times n}$

$$im\underset{\sim}{\boldsymbol{A}} = \left\{ \underline{\boldsymbol{x}} \middle| \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{y}}, \quad \underline{\boldsymbol{y}} \in \mathbb{C}^{n\times 1} \right\} = \left\{ \underline{\boldsymbol{x}} \middle| \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underbrace{\underset{\sim}{\boldsymbol{V}}^H\underline{\boldsymbol{y}}}_{\underline{\boldsymbol{z}}}, \quad \underline{\boldsymbol{y}} \in \mathbb{C}^{n\times 1} \right\}$$

$\Rightarrow \underline{\boldsymbol{y}} = \underset{\sim}{\boldsymbol{V}}\underline{\boldsymbol{z}}$

$$im\underset{\sim}{\boldsymbol{A}} = \left\{ \underline{\boldsymbol{x}} \middle| \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{U}} \begin{bmatrix} \underset{\sim}{\boldsymbol{\Sigma}}_1 & \underset{\sim}{\boldsymbol{0}} \\ \underset{\sim}{\boldsymbol{0}} & \underset{\sim}{\boldsymbol{0}} \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{z}}_1 \\ \underline{\boldsymbol{z}}_2 \end{bmatrix}, \quad \begin{array}{l} \underline{\boldsymbol{z}}_1 \in \mathbb{C}^{r\times 1} \\ \underline{\boldsymbol{z}}_2 \in \mathbb{C}^{(n-r)\times 1} \end{array} \right\} = \left\{ \underline{\boldsymbol{x}} \middle| \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{U}}_1 \underbrace{\underset{\sim}{\boldsymbol{\Sigma}}_1 \underline{\boldsymbol{z}}_1}_{\underline{\boldsymbol{W}}}, \quad \underline{\boldsymbol{z}}_1 \in \mathbb{C}^{r\times 1} \right\}$$

$$= \left\{ \underline{\boldsymbol{x}} \middle| \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{U}}_1 \underline{\boldsymbol{w}}, \quad \underline{\boldsymbol{w}}_1 \in \mathbb{C}^{r\times 1} \right\} = im\underset{\sim}{\boldsymbol{U}}_1 \qquad \square$$

Note: $\dim im\underset{\sim}{\boldsymbol{A}} = r$

Definition:
The null space: $null\underset{\sim}{\boldsymbol{A}} = \left\{ \underline{\boldsymbol{x}} \middle| \quad \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}} \right\}$

Theorem:

$$\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} \underset{\sim}{\boldsymbol{U}}_1 & \underset{\sim}{\boldsymbol{U}}_2 \end{bmatrix} \begin{bmatrix} \underset{\sim}{\boldsymbol{\Sigma}}_1 & \underset{\sim}{\boldsymbol{0}} \\ \underset{\sim}{\boldsymbol{0}} & \underset{\sim}{\boldsymbol{0}} \end{bmatrix} \begin{bmatrix} \underset{\sim}{\boldsymbol{V}}_1^H \\ \underset{\sim}{\boldsymbol{V}}_2^H \end{bmatrix} \quad \Rightarrow \quad null\underset{\sim}{\boldsymbol{A}} = im\underset{\sim}{\boldsymbol{V}}_2$$

Proof:

$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}, \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{y}}, \quad \underline{\boldsymbol{x}}$ represent the image of $\underset{\sim}{\boldsymbol{V}}_2$

$$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underbrace{\underset{\sim}{\boldsymbol{V}}_1^H \underset{\sim}{\boldsymbol{V}}_2}_{\underset{\sim}{\boldsymbol{0}}} \underline{\boldsymbol{y}} = \underline{\boldsymbol{0}} : \text{(TRUE)} \quad \Rightarrow \quad im\underset{\sim}{\boldsymbol{V}}_2 \subseteq null\underset{\sim}{\boldsymbol{A}}$$

Suppose: $\exists \underline{\boldsymbol{v}} \in null\underset{\sim}{\boldsymbol{A}}, \quad \underline{\boldsymbol{v}} \notin im\underset{\sim}{\boldsymbol{V}}_2$

$\underline{\boldsymbol{v}} \notin im\underset{\sim}{\boldsymbol{V}}_2 \Rightarrow \exists \underline{\boldsymbol{z}} : \underline{\boldsymbol{v}} = \underset{\sim}{\boldsymbol{V}}_1 \underline{\boldsymbol{z}}, \quad \underline{\boldsymbol{z}} \neq \underline{\boldsymbol{0}} \quad (1)$

$$\underline{\boldsymbol{0}} = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{v}} = \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underbrace{\underset{\sim}{\boldsymbol{V}}_1^H \underset{\sim}{\boldsymbol{V}}_1}_{\underset{\sim}{\boldsymbol{I}}} \underline{\boldsymbol{z}} = \underset{\sim}{\boldsymbol{U}}_1 \underbrace{\underset{\sim}{\boldsymbol{\Sigma}}_1 \underline{\boldsymbol{z}}}_{\underline{\boldsymbol{w}}} \quad \Rightarrow \quad \underline{\boldsymbol{0}} = \underset{\sim}{\boldsymbol{U}}_1 \underline{\boldsymbol{w}} \quad \textit{with} \quad \underline{\boldsymbol{w}} = \underset{\sim}{\boldsymbol{\Sigma}}_1 \underline{\boldsymbol{z}} = \underline{\boldsymbol{0}}$$

$$\Rightarrow \quad \underline{\boldsymbol{z}} = \underset{\sim}{\boldsymbol{\Sigma}}_1^{-1} \underline{\boldsymbol{0}} = \underline{\boldsymbol{0}} \quad (2)$$

$(1),(2) \quad \Rightarrow$ supposition is wrong: $\forall \underline{\boldsymbol{v}} \in null\underset{\sim}{\boldsymbol{A}} : \quad \underline{\boldsymbol{v}} \in im\underset{\sim}{\boldsymbol{V}}_2 \quad \Rightarrow \quad null\underset{\sim}{\boldsymbol{A}} = im\underset{\sim}{\boldsymbol{V}}_2 \qquad \square$

Note: $dim\, null\, \underset{\sim}{\boldsymbol{A}} = n - r \qquad dim\, im\, \underset{\sim}{\boldsymbol{A}} + dim\, null\, \underset{\sim}{\boldsymbol{A}} = n$

Definition:

The $rank\underset{\sim}{\boldsymbol{A}}$ is the number of LID columns of $\underset{\sim}{\boldsymbol{A}}$

Theorem:

$rank\underset{\sim}{\boldsymbol{A}} = r$

Proof:

$$\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H \quad \Rightarrow \quad im\underset{\sim}{\boldsymbol{A}} = im\underset{\sim}{\boldsymbol{U}}_1 = im\begin{bmatrix}\underline{\boldsymbol{u}}_1 & ... & \underline{\boldsymbol{u}}_r\end{bmatrix} \quad \Rightarrow \quad dim\underset{\sim}{\boldsymbol{A}} = dim\underset{\sim}{\boldsymbol{U}}_1 = r \qquad \square$$

Theorem:

$rank\underset{\sim}{\boldsymbol{A}} = rank\underset{\sim}{\boldsymbol{A}}^T = rank\underset{\sim}{\boldsymbol{A}}^H = rank\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H = rank\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}} = rank\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^T = rank\underset{\sim}{\boldsymbol{A}}^T\underset{\sim}{\boldsymbol{A}}$

Proof:

$\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H$

$\underset{\sim}{\boldsymbol{A}}^H = \underset{\sim}{\boldsymbol{V}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{U}}_1^H$

$\underset{\sim}{\boldsymbol{A}}$ and $\underset{\sim}{\boldsymbol{A}}^H$ have the same singular values.

$\Rightarrow$ The singular values for $\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H$ and $\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}}$ and $\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^T$ ... are $\underset{\sim}{\boldsymbol{\Sigma}}_1\underset{\sim}{\boldsymbol{\Sigma}}_1$, hence the rank is the same as the rank of $\underset{\sim}{\boldsymbol{A}}$ $\square$

2.4.8 Projectors

A Projector is a matrix $\underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{y}}$ that:

1. $im\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} = \mathbb{S}$ (image: Linear combination of the columns: set of vectors)
2. $\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} = \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}^H$
3. $\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} = \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}$

Theorem:

$\forall\underline{\boldsymbol{x}} \in \mathbb{S}: \quad \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{x}}$

$\forall\underline{\boldsymbol{x}} \notin \mathbb{S}: \quad \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{x}} \neq \underline{\boldsymbol{x}}$

Proof:

$\underline{\boldsymbol{x}} \in \mathbb{S} \quad \Rightarrow \quad \exists\underline{\boldsymbol{u}}: \quad \underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{y}}$

$$\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{x}} \overset{\underline{\boldsymbol{x}}=\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{y}}}{=} \underbrace{\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}}_{\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}}\underline{\boldsymbol{y}} = \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{y}} = \underline{\boldsymbol{x}}$$

if $\underline{\boldsymbol{x}} \notin \mathbb{S}$ but $\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{x}} \in \mathbb{S}$ then: $\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{x}} \neq \underline{\boldsymbol{x}}$ $\square$

Theorem:

$\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}$ is unique to $\mathbb{S}$

Proof:

$\underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}$ is a alternative Projector

$||(\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})\underline{\boldsymbol{z}}||_2^2$

$= ((\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})\underline{\boldsymbol{z}})^H(\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})\underline{\boldsymbol{z}} = \underline{\boldsymbol{z}}^H(\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}^H - \underset{\sim}{\boldsymbol{P}}'^H_{\mathbb{S}})(\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})\underline{\boldsymbol{z}}$

$$= \underline{\boldsymbol{z}}^H(\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})(\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})\underline{\boldsymbol{z}} = \underline{\boldsymbol{z}}^H(\underbrace{\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}}_{\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}} - \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} + \underbrace{\underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}\underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}}_{\underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}})\underline{\boldsymbol{z}}$$

$$= \underline{\boldsymbol{z}}^H(\underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})\underbrace{\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{z}}}_{\underline{\boldsymbol{y}}\in\mathbb{S}} + \underline{\boldsymbol{z}}^H(\underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}})\underbrace{\underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}\underline{\boldsymbol{z}}}_{\underline{\boldsymbol{w}}\in\mathbb{S}} = \underline{\boldsymbol{z}}^H(\underline{\boldsymbol{y}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}\underline{\boldsymbol{y}}) + \underline{\boldsymbol{z}}^H(\underline{\boldsymbol{w}} - \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{w}})$$

with $\underline{\boldsymbol{y}} = \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}\underline{\boldsymbol{y}}$, $\quad \underline{\boldsymbol{w}} = \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}}\underline{\boldsymbol{w}} \quad \Rightarrow \quad ||(\underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} - \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}})\underline{\boldsymbol{z}}||_2^2 = \underline{\boldsymbol{z}}^H(\underline{\boldsymbol{y}} - \underline{\boldsymbol{y}}) + \underline{\boldsymbol{z}}^H(\underline{\boldsymbol{w}} - \underline{\boldsymbol{w}}) = 0 \quad \forall \underline{\boldsymbol{z}}$

$\Rightarrow \quad \underset{\sim}{\boldsymbol{P}}_{\mathbb{S}} = \underset{\sim}{\boldsymbol{P}}'_{\mathbb{S}}$ $\square$

Theorem:

1. $\underset{\sim}{\boldsymbol{P}}_{im\underset{\sim}{\boldsymbol{A}}} = \underset{\sim}{\boldsymbol{U}}_1\underset{\sim}{\boldsymbol{U}}_1^H; \qquad \underset{\sim}{\boldsymbol{P}}_{im\underset{\sim}{\boldsymbol{A}}^H} = \underset{\sim}{\boldsymbol{V}}_1\underset{\sim}{\boldsymbol{V}}_1^H$

2. $\underset{\sim}{\boldsymbol{P}}_{null\underset{\sim}{\boldsymbol{A}}} = \underset{\sim}{\boldsymbol{V}}_2\underset{\sim}{\boldsymbol{V}}_2^H = \underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{P}}_{im\underset{\sim}{\boldsymbol{A}}^H}; \qquad \underset{\sim}{\boldsymbol{P}}_{null\underset{\sim}{\boldsymbol{A}}^H} = \underset{\sim}{\boldsymbol{U}}_2\underset{\sim}{\boldsymbol{U}}_2^H = \underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{P}}_{im\underset{\sim}{\boldsymbol{A}}}$

with $\underset{\sim}{\boldsymbol{I}} = \underset{\sim}{\boldsymbol{V}}_1\underset{\sim}{\boldsymbol{V}}_1^H + \underset{\sim}{\boldsymbol{V}}_2\underset{\sim}{\boldsymbol{V}}_2^H$ and $\underset{\sim}{\boldsymbol{I}} = \underset{\sim}{\boldsymbol{U}}_1\underset{\sim}{\boldsymbol{U}}_1^H + \underset{\sim}{\boldsymbol{U}}_2\underset{\sim}{\boldsymbol{U}}_2^H$, but $\underset{\sim}{\boldsymbol{I}} = \underset{\sim}{\boldsymbol{V}}_1^H\underset{\sim}{\boldsymbol{V}}_1$

-without proof

2.4.9 System of Linear Equation

System of Linear Equation: $\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$; $\qquad \underbrace{\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{m\times n}}_{\text{given}}; \qquad \underbrace{\underline{\boldsymbol{b}} \in \mathbb{C}^{m\times 1}}_{\text{given}}; \qquad \underbrace{\underline{\boldsymbol{x}} \in \mathbb{C}^{n\times 1}}_{\text{find x}}$

$\Rightarrow$ m equations and n unknowns

Solution set: $\mathbb{S} = \left\{\underline{\boldsymbol{x}} | \quad \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}\right\}$

①: $\underline{\boldsymbol{b}} \in im\underset{\sim}{\boldsymbol{A}}$: solution exist $\exists \underline{\boldsymbol{x}}_p \in \mathbb{C}^{n\times 1}$: $\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}}_p = \underline{\boldsymbol{b}}$ (p means patirticular solution)

The whole Solution set is $\mathbb{S} = \left\{\underline{\boldsymbol{x}}_p + \underline{\boldsymbol{y}} | \quad \underline{\boldsymbol{y}} \in null\underset{\sim}{\boldsymbol{A}}\right\}$ because:

1. $\underset{\sim}{\boldsymbol{A}}(\underline{\boldsymbol{x}}_p + \underline{\boldsymbol{y}}) = \underbrace{\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}}_p}_{\underline{\boldsymbol{b}}} + \underbrace{\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{y}}}_{\underline{\boldsymbol{0}}} = \underline{\boldsymbol{b}}$

2. $\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{z}} = \underline{\boldsymbol{b}}$; ($\underline{\boldsymbol{z}}$ is another particular solution)

   $\underset{\sim}{\boldsymbol{A}}(\underline{\boldsymbol{z}} - \underline{\boldsymbol{x}}_p) = \underline{\boldsymbol{0}} \quad \Rightarrow \quad \underline{\boldsymbol{z}} - \underline{\boldsymbol{x}}_p \in null\underset{\sim}{\boldsymbol{A}}$

   (difference between two particular solutions is part of the null-space of $\underset{\sim}{\boldsymbol{A}}$)

   $\underline{\boldsymbol{z}} = \underline{\boldsymbol{x}}_p + \underline{\boldsymbol{y}}, \qquad \underline{\boldsymbol{y}} \in null\underset{\sim}{\boldsymbol{A}}$

   (the sum of a particular solution and a vector of the null-space is another particluar solution)

with: $\quad null\underset{\sim}{\boldsymbol{A}} = im\underset{\sim}{\boldsymbol{V}}_2; \quad \underset{\sim}{\boldsymbol{V}} \in \mathbb{C}^{n\times(n-r)}$

$\Rightarrow \quad \mathbb{S} = \left\{\underline{\boldsymbol{x}}_p + \underset{\sim}{\boldsymbol{V}}_2 \cdot \underline{\boldsymbol{w}} | \quad \underline{\boldsymbol{w}} \in \mathbb{C}^{n-r}\right\}$ with $r = rank\underset{\sim}{\boldsymbol{A}}$

Special case: $rank\underset{\sim}{\boldsymbol{A}} = n$ full column rank matrix

unique solution $\mathbb{S} = \left\{\underline{\boldsymbol{x}}_p\right\}$

Particular Solution:

$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}}_p = \underline{\boldsymbol{b}}$ with $\underline{\boldsymbol{x}}_p = \underset{\sim}{\boldsymbol{A}}^{-1}\underline{\boldsymbol{b}} = \underset{\sim}{\boldsymbol{V}}_1\underset{\sim}{\boldsymbol{\Sigma}}_1^{-1}\underset{\sim}{\boldsymbol{U}}_1^H\underline{\boldsymbol{b}}$

$\underbrace{\underset{\sim}{\boldsymbol{U}}_1\underset{\sim}{\boldsymbol{\Sigma}}_1\underset{\sim}{\boldsymbol{V}}_1^H}_{\underset{\sim}{\boldsymbol{A}}}\underbrace{\underset{\sim}{\boldsymbol{V}}_1\underset{\sim}{\boldsymbol{\Sigma}}_1^{-1}\underset{\sim}{\boldsymbol{U}}_1^H\underline{\boldsymbol{b}}}_{\underline{\boldsymbol{x}}_p} = \underbrace{\underset{\sim}{\boldsymbol{U}}_1\underset{\sim}{\boldsymbol{U}}_1^H}_{\underset{\sim}{\boldsymbol{P}}_{im\underset{\sim}{\boldsymbol{A}}}}\underline{\boldsymbol{b}} \underset{b\in im\underset{\sim}{\boldsymbol{A}}}{=} \underline{\boldsymbol{b}}$ with $\underset{\sim}{\boldsymbol{V}}_1^H\underset{\sim}{\boldsymbol{V}}_1 = \underset{\sim}{\boldsymbol{I}}$, and $\underset{\sim}{\boldsymbol{\Sigma}}_1\underset{\sim}{\boldsymbol{\Sigma}}_1^{-1} = \underset{\sim}{\boldsymbol{I}}$

Particular Solution: $\underline{\boldsymbol{x}}_p = \underset{\sim 1}{\boldsymbol{V}}\underset{\sim 1}{\boldsymbol{\Sigma}}^{-1}\underset{\sim 1}{\boldsymbol{U}}^H\underline{\boldsymbol{b}}$

The whole Solution set with the SVD is: $\mathbb{S} = \left\{\underset{\sim 1}{\boldsymbol{V}}\underset{\sim 1}{\boldsymbol{\Sigma}}^{-1}\underset{\sim 1}{\boldsymbol{U}}^H\underline{\boldsymbol{b}} + \underset{\sim 2}{\boldsymbol{V}}\underline{\boldsymbol{w}} \middle| \quad \underline{\boldsymbol{w}} \in \mathbb{C}^{(n-r)\times 1}\right\}$

if $rank\underset{\sim}{\boldsymbol{A}} < n \quad \Rightarrow$ mulitple solution with $(n-r)$ dimensions

Minimum Norm Solution:

$\underline{\boldsymbol{x}}_{MN} = arg\min\limits_{\underline{\boldsymbol{x}}} ||\underline{\boldsymbol{x}}||_2^2 \quad \text{s.t. } \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}}_{MN} = \underline{\boldsymbol{b}}$

$\mathcal{L} = \underbrace{\underline{\boldsymbol{x}}^H\underline{\boldsymbol{x}}}_{||\underline{\boldsymbol{x}}||_2^2} + \underline{\boldsymbol{\lambda}}^H(\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} - \underline{\boldsymbol{b}}) + (\underline{\boldsymbol{x}}^H\underset{\sim}{\boldsymbol{A}}^H - \underline{\boldsymbol{b}}^H)\underline{\boldsymbol{\lambda}} \quad \Rightarrow \quad \frac{\partial\mathcal{L}}{\partial\underline{\boldsymbol{x}}^*}\underline{\boldsymbol{x}} + \underset{\sim}{\boldsymbol{A}}^H\lambda \overset{!}{=} \underline{\boldsymbol{0}} \quad \Rightarrow \quad \underline{\boldsymbol{x}} = -\underset{\sim}{\boldsymbol{A}}^H\underline{\boldsymbol{\lambda}}$

$\underline{\boldsymbol{b}} = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = -\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H\underline{\boldsymbol{\lambda}} \quad \Rightarrow \quad \underline{\boldsymbol{\lambda}} = -(\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H)^{-1}\underline{\boldsymbol{b}}$

Minimum Norm Solution: $\quad \Rightarrow \quad \underline{\boldsymbol{x}}_{MN} = \underset{\sim}{\boldsymbol{A}}^H(\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H)^{-1}\underline{\boldsymbol{b}}$

Asume: $\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H$ is inverteble $\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{m\times n} \quad \Rightarrow \quad \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H \in \mathbb{C}^{m\times m}$

$rank\underset{\sim}{\boldsymbol{A}} = m = r \quad$ full row rank

$$\underset{\sim}{\boldsymbol{U}} = \left.\begin{bmatrix}\underbrace{\underset{\sim 1}{\boldsymbol{U}}}_{r} & \underbrace{\underset{\sim 2}{\boldsymbol{U}}}_{n-r}\end{bmatrix}\right\} m \overset{m=r}{=} \underset{\sim 1}{\boldsymbol{U}}$$

SVD: $\underset{\sim}{\boldsymbol{A}} = \underset{\sim 1}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim 1}{\boldsymbol{V}}^H = \underset{\sim}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim 1}{\boldsymbol{V}}^H \quad$ with $\underset{\sim 1}{\boldsymbol{U}} = \underset{\sim}{\boldsymbol{U}} \qquad \underset{\sim}{\boldsymbol{U}}^{-1} = \underset{\sim}{\boldsymbol{U}}^H$

$$\underline{\boldsymbol{x}}_{MN} = \underset{\sim}{\boldsymbol{A}}^H(\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H)^{-1}\underline{\boldsymbol{b}} = \underset{\sim 1}{\boldsymbol{V}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{U}}^H(\underset{\sim}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underbrace{\underset{\sim 1}{\boldsymbol{V}}^H\underset{\sim 1}{\boldsymbol{V}}}_{\underset{\sim}{\boldsymbol{I}}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{U}}^H)^{-1}\underline{\boldsymbol{b}} = \underset{\sim 1}{\boldsymbol{V}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{U}}^H(\underset{\sim}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}^2\underset{\sim}{\boldsymbol{U}}^H)^{-1}\underline{\boldsymbol{b}}$$

$$= \underset{\sim 1}{\boldsymbol{V}}\underbrace{\underset{\sim 1}{\boldsymbol{\Sigma}}\underbrace{\underset{\sim}{\boldsymbol{U}}^H\underset{\sim}{\boldsymbol{U}}}_{\underset{\sim}{\boldsymbol{I}}}\underset{\sim 1}{\boldsymbol{\Sigma}}^{-2}}_{\Sigma_1^{-1}}\underset{\sim}{\boldsymbol{U}}^H\underline{\boldsymbol{b}} = \underset{\sim 1}{\boldsymbol{V}}\Sigma_1^{-1}\underset{\sim}{\boldsymbol{U}}^H\underline{\boldsymbol{b}}$$

$\Rightarrow \quad \underline{\boldsymbol{x}}_{MN} = \underset{\sim 1}{\boldsymbol{V}}\Sigma_1^{-1}\underset{\sim 1}{\boldsymbol{U}}^H\underline{\boldsymbol{b}} \quad = \underline{\boldsymbol{x}}_p \quad$ (out of section "particular solution")

②: $\underline{\boldsymbol{b}} \notin im\underset{\sim}{\boldsymbol{A}} \quad \Rightarrow \quad \mathbb{S} = \emptyset = \{\}$

Approximate Solution: eg. Least-Squeare Solution

$\underline{\boldsymbol{x}}_{LS} = arg\min\limits_{\underline{\boldsymbol{x}}} ||\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} - \underline{\boldsymbol{b}}||_2^2$

$\mathcal{L} = (\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} - \underline{\boldsymbol{b}})^H(\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} - \underline{\boldsymbol{b}}) = (\underline{\boldsymbol{x}}^H\underset{\sim}{\boldsymbol{A}}^H - \underline{\boldsymbol{b}}^H)(\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} - \underline{\boldsymbol{b}}) = \underline{\boldsymbol{x}}^H\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} - \underline{\boldsymbol{x}}^H\underset{\sim}{\boldsymbol{A}}^H\underline{\boldsymbol{b}} - \underline{\boldsymbol{b}}^H\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} + \underline{\boldsymbol{b}}^H\underline{\boldsymbol{b}}$

$\frac{\partial\mathcal{L}}{\partial\underline{\boldsymbol{z}}^*} = \underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} - \underset{\sim}{\boldsymbol{A}}^H\underline{\boldsymbol{b}} \overset{!}{=} 0$

$\Rightarrow \quad \underline{\boldsymbol{x}}_{LS} = (\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}})^{-1}\underset{\sim}{\boldsymbol{A}}^H\underline{\boldsymbol{b}}$

asume $rank\underset{\sim}{\boldsymbol{A}} = n \quad \Rightarrow \quad$ full column rank $\quad \Rightarrow \quad \underset{\sim 1}{\boldsymbol{V}} = \underset{\sim}{\boldsymbol{V}}$

SVD: $\underset{\sim}{\boldsymbol{A}} = \underset{\sim 1}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim 1}{\boldsymbol{V}}^H = \underset{\sim 1}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H \quad with \quad \underset{\sim 1}{\boldsymbol{V}} = \underset{\sim}{\boldsymbol{V}}$

$\underline{\boldsymbol{x}}_{LS} = (\underset{\sim}{\boldsymbol{V}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim 1}{\boldsymbol{U}}^H\underset{\sim 1}{\boldsymbol{U}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H)^{-1}\underset{\sim}{\boldsymbol{V}}\underset{\sim 1}{\boldsymbol{\Sigma}}\underset{\sim 1}{\boldsymbol{U}}^H\underline{\boldsymbol{b}}$

$$\Rightarrow \quad \underline{\boldsymbol{x}}_{LS} = \underbrace{\underset{\sim 1}{\boldsymbol{V}}\Sigma_1^{-1}\underset{\sim 1}{\boldsymbol{U}}^H}_{\underset{\sim}{\boldsymbol{A}}^+}\underline{\boldsymbol{b}} \quad = \underline{\boldsymbol{x}}_p \quad = \underset{\sim}{\boldsymbol{A}}^+\underline{\boldsymbol{b}}$$

Moore-Penrose-Pseudo-Inverse: $\underset{\sim}{\boldsymbol{A}}^+ = \underset{\sim 1}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma_1}}^{-1}\underset{\sim 1}{\boldsymbol{U}}^H$

$$\underset{\sim}{\boldsymbol{A}}^+ := \begin{cases}\underset{\sim 1}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma_1}}^{-1}\underset{\sim 1}{\boldsymbol{U}}^H & \underset{\sim}{\boldsymbol{A}} \neq \underset{\sim}{\boldsymbol{0}} \\ \underset{\sim}{\boldsymbol{A}}^T & \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{0}}\end{cases}$$

$$\underset{\sim}{\boldsymbol{A}}^+\underline{\boldsymbol{b}} = \begin{cases}\text{Minimum-Norm-Solution if } \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}} \in im\underset{\sim}{\boldsymbol{A}} \\ \text{Least-Square-Solution if } \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}} \notin im\underset{\sim}{\boldsymbol{A}}\end{cases}$$

Compute Pseudo Inverse:

- $\underset{\sim}{\boldsymbol{A}}^+ = \underset{\sim}{\boldsymbol{A}}^H(\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^H)^{-1} \quad \underset{\sim}{\boldsymbol{A}}$ has full row rank

- $\underset{\sim}{\boldsymbol{A}}^{+} = (\underset{\sim}{\boldsymbol{A}}^{H}\underset{\sim}{\boldsymbol{A}})^{-1}\underset{\sim}{\boldsymbol{A}}^{H} \quad \underset{\sim}{\boldsymbol{A}}$ has full column rank

- $\underset{\sim}{\boldsymbol{A}}^{+} = \lim\limits_{\mu\to 0}\underset{\sim}{\boldsymbol{A}}^{H}(\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^{H} + \mu\underset{\sim}{\boldsymbol{I}})^{-1} = \lim\limits_{\mu\to 0}(\underset{\sim}{\boldsymbol{A}}^{H}\underset{\sim}{\boldsymbol{A}} + \mu\underset{\sim}{\boldsymbol{I}})^{-1}\underset{\sim}{\boldsymbol{A}}^{H}$

Theorem:

- $\underset{\sim}{\boldsymbol{P}}_{im\underset{\sim}{\boldsymbol{A}}} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^{+}$

- $\underset{\sim}{\boldsymbol{P}}_{null\underset{\sim}{\boldsymbol{A}}} = \underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^{+}$

- without proof

Solution set with Pseudo Inverse:
$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}} \quad \Rightarrow \quad \mathbb{S} = \left\{\underset{\sim}{\boldsymbol{A}}^{+}\underline{\boldsymbol{b}} + (\underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{A}}^{+})\underline{\boldsymbol{w}} | \quad \underline{\boldsymbol{w}} \in \mathbb{C}^{(n-r)\times 1}\right\}$

2.4.10 The Eckart-Young Theorem

Low-rank approximation
given: $\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{m\times n}; \qquad rank\underset{\sim}{\boldsymbol{A}} = r$
looking for $\underset{\sim}{\boldsymbol{B}} \in \mathbb{C}^{m\times n}; \quad rank\underset{\sim}{\boldsymbol{A}} = k < r$
$\min\limits_{\underset{\sim}{\boldsymbol{B}}} ||\underset{\sim}{\boldsymbol{A}} - \underset{\sim}{\boldsymbol{B}}||_F^2$ s.t. $rank\underset{\sim}{\boldsymbol{B}} = k < r$
Remember: Frobenius norm:
$||\underset{\sim}{\boldsymbol{X}}||_F^2 = tr(\underset{\sim}{\boldsymbol{X}}^{H}\underset{\sim}{\boldsymbol{X}}) = tr(\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^{H})$
Frobenius norm for Matrices:
$||\underset{\sim}{\boldsymbol{C}}||_F^2 = ||\underset{\sim}{\boldsymbol{U}}^{H}\underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{V}}||_F^2$ with $\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^{H}; \quad \underset{\sim}{\boldsymbol{U}}, \underset{\sim}{\boldsymbol{V}}$ unitary
Explanation:

$$||\underset{\sim}{\boldsymbol{U}}^{H}\underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{V}}||_F^2 = tr(\underset{\sim}{\boldsymbol{V}}^{H}\underset{\sim}{\boldsymbol{C}}^{H}\underbrace{\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{U}}^{H}}_{\underset{\sim}{\boldsymbol{I}}}\underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{V}}) = tr(\underset{\sim}{\boldsymbol{V}}^{H}\underset{\sim}{\boldsymbol{C}}^{H}\underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{V}}) = tr(\underset{\sim}{\boldsymbol{C}}\underbrace{\underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{V}}^{H}}_{\underset{\sim}{\boldsymbol{I}}}\underset{\sim}{\boldsymbol{C}}^{H})$$

$= tr(\underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{C}}^{H}) = ||\underset{\sim}{\boldsymbol{C}}||_F^2$
Now: $\underset{\sim}{\boldsymbol{C}} = \underset{\sim}{\boldsymbol{A}} - \underset{\sim}{\boldsymbol{B}}$

$$||\underset{\sim}{\boldsymbol{A}} - \underset{\sim}{\boldsymbol{B}}||_F^2 = ||\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^{H} - \underset{\sim}{\boldsymbol{B}}||_F^2 = ||\underset{\sim}{\boldsymbol{\Sigma}} - \underbrace{\underset{\sim}{\boldsymbol{U}}^{H}\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{V}}}_{\underset{\sim}{\boldsymbol{M}}}||_F^2 \quad \Rightarrow \quad \underset{\sim}{\boldsymbol{M}} = \underset{\sim}{\boldsymbol{U}}^{H}\underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{V}}$$

$||\underset{\sim}{\boldsymbol{A}} - \underset{\sim}{\boldsymbol{B}}||_F^2 = ||\underset{\sim}{\boldsymbol{\Sigma}} - \underset{\sim}{\boldsymbol{M}}||_F^2 = \sum\limits_{i=1}^{r}|s_i - M_{ii}|^2 + \sum\limits_{i>r}|M_{ii}|^2 + \sum\limits_{i,j\neq i}|M_{ij}|^2$
Get Minimum if:

- $M_{i,j\neq i} = 0$

- $M_{i,i} = 0, i > r$

- $M_{i,i} = s_i; \quad for \quad i \in \{1, 2, ..., k\}$ because $rank\underset{\sim}{\boldsymbol{B}} = k < r$
  $s_1 \geq s_2 \geq ... \geq s_k > s_{r+1} = s_{r+2} = ... = 0$

Approximation of $\underset{\sim}{\boldsymbol{A}}$ with lower Rank: $\underset{\sim}{\boldsymbol{B}} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{M}}\underset{\sim}{\boldsymbol{V}}^H$

$$\text{with} \quad \underset{\sim}{\boldsymbol{M}} = \begin{bmatrix} s_1 & & & & \\ & \ddots & & & \\ & & s_k & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix} = \underset{\sim}{\boldsymbol{U}}^H \underset{\sim}{\boldsymbol{B}} \underset{\sim}{\boldsymbol{V}} \qquad \underset{\sim}{\boldsymbol{B}} = \begin{bmatrix} \underline{\boldsymbol{u}}_1 & \cdots & \underline{\boldsymbol{u}}_k \end{bmatrix} \begin{bmatrix} s_1 & & \\ & \ddots & \\ & & s_k \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{v}}_1^H \\ \vdots \\ \underline{\boldsymbol{v}}_k^H \end{bmatrix}$$

$$\boxed{\underset{\sim}{\boldsymbol{B}} = \sum_{i=1}^{k} s_i \underline{\boldsymbol{u}}_i \underline{\boldsymbol{v}}_i^H}$$

2.5 2^{nd}-Order Statistic

2.5.1 reale valued discret time random processes

Expected Value:

$\mu_x[n] = E[x[n]] := \lim_{m \to M} \frac{1}{m} \sum_{i=1}^{m} (x[n])^i$

Auto-Correlation:

$r_x[n,k] := E[x[n]x[n-k]]$

Auto-Covariance

$C_x[n,k] := E[(x[n] - \mu_x[n])(x[n-k] - \mu_x[n-k])]$ (in this lecture $\mu_x = 0 \to C_x[n,k] = r_x[n,k]$)

$C_x[n,k] = r_x[n,k] - \mu_x[n]\mu_x[n-k]$

Cross-Correlation

$r_{xy}[n,k] := E[x[n]y[n-k]]$

$r_{xx}[n,k] = r_x[n,k]$

Cross-Covariance

$C_{xy}[n,k] := E[(x[n] - \mu_x[n])(y[n-k] - \mu_y[n-k])]$

$C_{xy}[n,k] = r_{xy}[n,k] - \mu_x[n]\mu_y[n-k]$

Time Average

$\hat{\mu}_x^{(N)} := \frac{1}{N} \sum_{n=0}^{N-1} x[n]$ sample mean

$\hat{r}_n(N) := \frac{1}{N} \sum_{n=0}^{N-1} x[n]x[n-k]$

Stationarity $\iff$ all statistic properties are independent of time

wide sense stationarity $\iff$ 1^{st} and 2^{nd} order statistic do not depend on time

$\Rightarrow$ $r_{xy}[n,k] = r_{xy}[k]$ (Cross-Correlation is no function of time)

ergodicity $\iff$ $\mu_x[x] = \lim_{N \to \infty} \hat{\mu}_x^{(N)}, \quad \forall n \quad \Rightarrow$ sample mean = expectation

you see one, you see them all

$\Rightarrow$ **In the follwoing:**

- ergodicity
- zero mean, $\mu_x[n] = 0$

Consequence

1. $r_x[n,k] = r_x[k]$
2. $r_x[-k] = E[x[n]x[n+k]] \overset{n-k \to n}{=} E[x[n-k]x[n]] = r_x[k]$, even function
3. $r_{xy}[-k] = r_{yx}[k]$

2.5.2 complex random proccesses

complex random variable

$u[n] := x[n] + \mathrm{j}y[n]; \quad x[n] = \mathrm{Re}\{u[n]\}, \quad y[n] = \mathrm{Im}\{u[n]\}$

Definition: Auto-Correlations - two different functions

- $\tilde{r}_u[k] := E[u[n] \cdot u[n-k]]$
- $r_u[k] := E[u[n] \cdot u^*[n-k]]$

$\tilde{r}_u[k] = r_x[k] - r_y[k] + \mathrm{j}(r_{xy}[-k] + r_{xy}[k])$
$r_u[k] = r_x[k] + r_y[k] + \mathrm{j}(r_{xy}[-k] - r_{xy}[k])$
with:

- $r_x[k] = \frac{1}{2}\,\mathrm{Re}\,\{r_u[k] + \tilde{r}_u[k]\}$
- $r_y[k] = \frac{1}{2}\,\mathrm{Re}\,\{r_u[k] - \tilde{r}_u[k]\}$
- $r_{xy}[k] = -\frac{1}{2}\,\mathrm{Im}\,\{r_u[k] - \tilde{r}_u[k]\}$

Definiton:
$u[n]$ is **proper** if and only if $\tilde{r}_u[k] \equiv 0$
Consequence of properness:

- $r_x[k] = r_y[k]$
- $r_{xy}[k] = -r_{xy}[-k]$

  $\rightarrow\ r_{xy} = 0 = E[x[0]y[0]]$

  $\rightarrow\ E[u^2[n]] = 0 = E[u[n]u[n]]$

Note: $E[u[n]u^*[n]] \neq 0$

$$\Rightarrow \quad E[u^2[n]] = E[(x[n] + \mathrm{j}y[n])^2] = E[(x[n])^2 + \mathrm{j}2 \cdot x[n] \cdot y[n] - (y[n])^2]$$
$$= \underbrace{E[x^2[n]] - E[y^2[n]]}_{0} + \mathrm{j}2\,\underbrace{E[x[n]y[n]]}_{0}$$

$\Rightarrow$ **In the follwoing:**

- ergodicity
- zero mean, $\mu_x[n] = 0$
- proper

Therefore the Correlations show this properties:

- $r_u[-k] = (r_u[k])^*$
- $r_{uw}[-k] = (r_{wu}[k])^*$

Include the Properties of the real case, because its proper. In general this is not satisfied.

2.5.3 Correlation Matrix

Signal Vector

$$\underline{\boldsymbol{u}}[n] := \begin{bmatrix} u[n] \\ u[n-1] \\ \vdots \\ u[n-M+1] \end{bmatrix} \in \mathbb{C}^{M\times 1}$$

Correlation Matrix

$$\begin{aligned}
\underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}} &:= E[\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n]] \in \mathbb{C}^{M\times M} \\
&= E\left[\begin{bmatrix} u[n] \\ u[n-1] \\ \vdots \\ u[n-M+1] \end{bmatrix} \begin{bmatrix} u^*[n] & u^*[n-1] & \cdots & u^*[n-M+1] \end{bmatrix}\right] \\
&= \begin{bmatrix} r_u[0] & r_u[1] & r_u[2] & \cdots & r_u[M-1] \\ r_u^*[1] & r_u[0] & r_u[1] & \cdots & r_u[M-2] \\ r_u^*[2] & r_u^*[1] & r_u[0] & \cdots & r_u[M-3] \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ r_u^*[M-1] & r_u^*[M-2] & r_u[M-3] & \cdots & r_u[0] \end{bmatrix}
\end{aligned}$$

Note:

- $\underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}}$ is a Toeplitz Matrix
- $\underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}} = \underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}}^H$, always true
- $\underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}}$ is Gram-ian

  $\rightarrow$ EVD Exist: $\underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}} = \underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{V}}^{-1} = \underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{V}}^H, \quad \underset{\sim}{\boldsymbol{\Lambda}} \geq \underset{\sim}{\boldsymbol{0}} \quad \Rightarrow \quad \underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}} \geq \underset{\sim}{\boldsymbol{0}}$

positive semidefinite: $\underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}} \geq \underset{\sim}{\boldsymbol{0}} \iff \forall \underline{\boldsymbol{w}} \in \mathbb{C}^{M\times 1}: \quad \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}} \underline{\boldsymbol{w}} \geq 0$

$$\Rightarrow \quad \underbrace{\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{V}}}_{\underline{\boldsymbol{z}}^H} \underset{\sim}{\boldsymbol{\Lambda}} \underbrace{\underset{\sim}{\boldsymbol{V}}^H \underline{\boldsymbol{w}}}_{\underline{\boldsymbol{z}}} = \underline{\boldsymbol{z}}^H \underset{\sim}{\boldsymbol{\Lambda}} \underline{\boldsymbol{z}} = \sum_{i=1}^{M} |z_i|^2 \lambda_i \geq 0$$

Example:

Complex Envelope:

$$u[n] = \alpha e^{\mathrm{j}\omega T n} + \underbrace{\eta[n]}_{\text{noise}}: \quad \alpha, \omega, T : const; \quad \omega, T \in \mathbb{R}; \quad \alpha \in \mathbb{C}$$

Real Signal

$$\mathrm{Re}\left\{u[n] \cdot e^{\mathrm{j}\underbrace{2\pi f_0}_{\omega_0} T n}\right\} = \mathrm{Re}\left\{\alpha e^{\mathrm{j}(\omega+\omega_0)Tn}\right\} = |\alpha| \cos((\omega+\omega_0)Tn + \arg\alpha)$$

with $\quad \omega_0$: carrier frequency

Noise:

$$\mu_\eta[n] = 0, \quad r_\eta[k] = \begin{cases} \sigma_\eta^2 & k = 0 \\ 0 & \text{else} \end{cases}$$

$r_\eta[0] = E[\eta[n] \cdot \eta^*[n]] = E[|\eta[n]|^2] = \sigma_\eta^2$

Correlation of the Signal

$$r_u[k] = E[u[k] \cdot u^*[n-k]] = E[(\alpha e^{\mathrm{j}\omega T n} + \eta[n])(\alpha^* e^{-\mathrm{j}\omega T(n-k)} + \eta^*[n-k])]$$
$$= E[|\alpha|^2 e^{\mathrm{j}\omega T(n-n+k)}] + E[\alpha e^{\mathrm{j}\omega T n}\eta^*[n-k]] + E[\eta[n]\alpha^* e^{-\mathrm{j}\omega T(n-k)}] + E[\eta[n]\eta^*[n-k]]$$
$$= |\alpha|^2 e^{\mathrm{j}\omega T k} + \begin{cases} \sigma_\eta^2 & k=0 \\ 0 & \text{else} \end{cases} = |\alpha|^2 e^{\mathrm{j}\omega T k} + \sigma_\eta^2 \cdot \delta_F; \qquad \delta_F = \begin{cases} 1 & k=0 \\ 0 & \text{else} \end{cases}$$

Signal-to-Noise Ratio (SNR)

$$r_u[k] = |\alpha|^2 e^{\mathrm{j}\omega T k} + \sigma_\eta^2 \cdot \delta_F = |\alpha|^2 (e^{\mathrm{j}\omega T k} + \tfrac{1}{SNR} \cdot \delta_F) \quad \text{with} \quad SNR := \tfrac{|\alpha|^2}{\sigma_\eta^2}$$

Correlation Matrix

$$\underset{\sim}{\boldsymbol{R}}_{\underline{\boldsymbol{u}}} = |\alpha|^2 \begin{bmatrix} 1+\frac{1}{SNR} & e^{j\omega T} & e^{j2\omega T} & \cdots & e^{j(M-1)\omega T} \\ e^{-j\omega T} & 1+\frac{1}{SNR} & e^{j\omega T} & \cdots & e^{j(M-2)\omega T} \\ e^{-j2\omega T} & e^{-j\omega T} & 1+\frac{1}{SNR} & \cdots & e^{j(M-3)\omega T} \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ e^{-j(M-1)\omega T} & e^{-j(M-2)\omega T} & e^{-j(M-3)\omega T} & \cdots & 1+\frac{1}{SNR} \end{bmatrix}$$

3 Optimum Linear Filtering

- Discrete Time
- Proper random processes
- Mimimum mean square error (MMSE), Linear processing MMSE (LMMSE)
- Theory by N. Wiener (1948) and A. Kolmogorov (1941)

The design of a Wiener Filter requires a priori information about the statistics of the data to be processed. The filter is optimum only when the statistical characteristics of the input data match the priori information on which the design of the filter is based on. When this information is not known completely, however, it may not be possible to design the Wiener Filter or else the design may no loner be optimum.A straightforward approach that we may use in such situation is the "estimate and plug" procedure. This is a two-stage process, whereby the filter first "estimates" the statistical parameters of the relevant signals and then "plugs" the results so obtained into a non recursive formula for computing the filter parameters. For real-time operation, this procedure has the disadvantage of requiring excessively elaborate and costy hardware.
A widre variety of recursive algorithms have been developed in the literature for the operation of linear adaptive filters. In the final analysis, the choise of the algorithm is determinded by one or more of the following factors:

- rate of convergence
- misadjustment
- tracking
- robustness
- computaional requirements

$x[n] = \sum_{k=0}^{K} h[k] \cdot s[n-k];\qquad K\text{: Channel memory } (0 \leq K)$

$y[n] = \sum_{k=0}^{M-1} w_k^* \cdot u[n-k];\qquad M\text{: Number of filter coefficients (filter order) } (1 \leq M)$

$$\underline{\boldsymbol{x}}[n] = \begin{bmatrix} x[n] \\ x[n-1] \\ \vdots \\ x[n-M+1] \end{bmatrix}; \qquad \underline{\boldsymbol{s}}[n] = \begin{bmatrix} s[n] \\ s[n-1] \\ \vdots \\ s[n-N+1] \end{bmatrix}$$

$$\underline{\boldsymbol{x}}[n] = \underbrace{\begin{bmatrix} h[0] & h[1] & h[2] & \cdots & h[K] & & 0 & \cdots & 0 & \\ 0 & h[0] & h[1] & h[2] & \cdots & h[K-1] & h[K] & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ 0 & \cdots & 0 & h[0] & h[1] & h[2] & & \cdots & h[K-1] & h[K] \end{bmatrix}}_{\underset{\sim}{\boldsymbol{H}}} \cdot \underline{\boldsymbol{s}}[n]$$

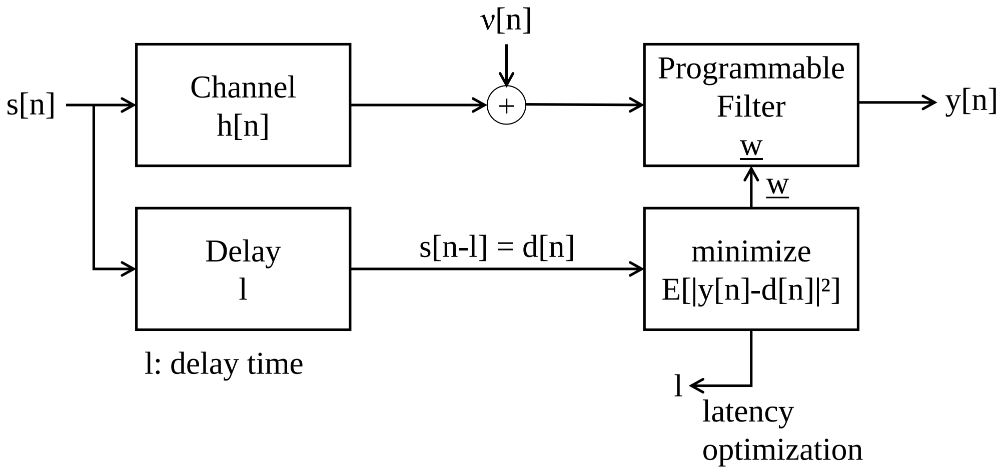

Figure 25: Basic block diagram of an Optimum Linear Filter

$\underset{\sim}{\boldsymbol{H}} \in \mathbb{C}^{M \times N}$
$N = M + K \geq M$
$\Rightarrow$ wide Matrix

$$\underline{\boldsymbol{w}} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_{M-1} \end{bmatrix}; \qquad \underline{\boldsymbol{u}}[n] = \underline{\boldsymbol{x}}[n] + \underline{\boldsymbol{\nu}}[n]; \qquad \underline{\boldsymbol{\nu}}[n]\text{: Noise}$$

$$y[n] = \underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n] = \underline{\boldsymbol{w}}^H (\underline{\boldsymbol{x}}[n] + \underline{\boldsymbol{\nu}}[n])$$

$y[n] = \underline{\boldsymbol{w}}^H (\underset{\sim}{\boldsymbol{H}} \underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n])$
with:
$y[n]$: Filter output
$\underline{\boldsymbol{w}}^H$: Filter vector
$\underset{\sim}{\boldsymbol{H}}$: Channel matrix
$\underline{\boldsymbol{s}}[n]$: Tx-signal vector
$\underline{\boldsymbol{\nu}}[n]$: Noise vector

3.1 Correlation between filteroutput and error in the optimum

- Error signal: $e[n] = d[n] - y[n]$
- Cost function: $J = E[|e[n]|^2]$
- Optimization: $\underline{\boldsymbol{w}}_{opt} = \arg \min\limits_{\underline{\boldsymbol{w}}} J$

$$J = E[e^*[n] \cdot e[n]] = E[(d^*[n] - y^*[n])(d[n] - \underbrace{y[n]}_{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n]})]$$

$$\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*} = E[e^*[n](-\underline{\boldsymbol{u}}[n])] \overset{!}{=} 0 \text{ (in the optimum)}$$

$$E[e^*[n] y[n]] = E[e^*[n] \underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n]] \underbrace{=}_{\underline{\boldsymbol{w}}^H \text{ deterministic}} \underline{\boldsymbol{w}}^H \, E[e^*[n] \underline{\boldsymbol{u}}[n]] = 0$$

in the optimum $\Leftrightarrow E[y[n] e^*[n]] = 0$ (Filteroutput is uncorrelated with error)

If you want to know, whether a different filter is optimal or not, test if the error and the output are uncorrelated.

3.2 Determine Optimal Filter Weight w_{opt}

$J = E[(d^*[n] - y^*[n])(d[n] - y[n])] = \sigma_d^2 - \underline{\boldsymbol{w}}^H \underline{\boldsymbol{p}} - \underline{\boldsymbol{p}}^H \underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}}$

$\sigma_d^2 = E[|d[n]|^2]$

$\underline{\boldsymbol{p}} = E[\underline{\boldsymbol{u}}[n] d^*[n]];$ correlation vector

$\underset{\sim}{\boldsymbol{R}} = E[\underline{\boldsymbol{u}}[n] \underline{\boldsymbol{u}}^H[n]];$ corrrelation matrix

Note: $\underset{\sim}{\boldsymbol{R}} \geq 0$ usually $\underset{\sim}{\boldsymbol{R}} > 0 \Rightarrow \underset{\sim}{\boldsymbol{R}}^{-1}$ exists

$\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H$

$\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*} = -\underline{\boldsymbol{p}} + \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}} \overset{!}{=} 0$

$$\boxed{\underline{\boldsymbol{w}}_{opt} = \underset{\sim}{\boldsymbol{R}}^{-1} \underline{\boldsymbol{p}}}$$

$\underline{\boldsymbol{w}}_{opt}^H = \underline{\boldsymbol{p}}^H (\underset{\sim}{\boldsymbol{R}}^{-1})^H = \underline{\boldsymbol{p}}^H (\underset{\sim}{\boldsymbol{R}}^H)^{-1} = \underline{\boldsymbol{p}}^H \underset{\sim}{\boldsymbol{R}}^{-1}$

$$J \big|_{\underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_{opt}} = \sigma_d^2 - \underbrace{\underline{\boldsymbol{p}}^H \underset{\sim}{\boldsymbol{R}}^{-1} \underline{\boldsymbol{p}}}_{= \underline{\boldsymbol{w}}_{opt} \underline{\boldsymbol{p}}} \underbrace{- \underline{\boldsymbol{p}}^H \underset{\sim}{\boldsymbol{R}}^{-1} \underline{\boldsymbol{p}} + \underbrace{\underline{\boldsymbol{p}}^H \underset{\sim}{\boldsymbol{R}}^{-1} \underset{\sim}{\boldsymbol{R}} \underset{\sim}{\boldsymbol{R}}^{-1} \underline{\boldsymbol{p}}}_{= \underline{\boldsymbol{p}}^H \underset{\sim}{\boldsymbol{R}}^{-1} \underline{\boldsymbol{p}}}}_{=0} = \sigma_d^2 (1 - \underbrace{\frac{\underline{\boldsymbol{p}}^H \underset{\sim}{\boldsymbol{R}}^{-1} \underline{\boldsymbol{p}}}{\sigma_d^2}}_{|\rho|^2}) \qquad 0 \leq |\rho|^2 \leq 1$$

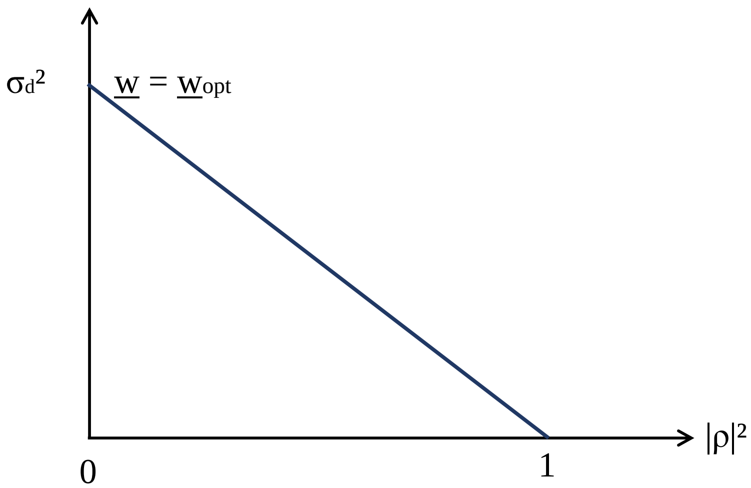

Figure 26: Correlation betweem σ_d^2 and $|\rho|^2$

$$\begin{aligned}
J &= \sigma_d^2 - \underline{\boldsymbol{w}}^H\underline{\boldsymbol{p}} - \underline{\boldsymbol{p}}^H\underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}\\
&= \sigma_d^2 - \underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{R}}\underbrace{\underset{\sim}{\boldsymbol{R}}^{-1}\underline{\boldsymbol{p}}}_{\underline{\boldsymbol{w}}_{opt}} - \underbrace{\underline{\boldsymbol{p}}^H\underset{\sim}{\boldsymbol{R}}^{-1}}_{\underline{\boldsymbol{w}}_{opt}^H}\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}} + \underline{\boldsymbol{p}}^H\underbrace{\underset{\sim}{\boldsymbol{R}}^{-1}\underset{\sim}{\boldsymbol{R}}\underset{\sim}{\boldsymbol{R}}^{-1}}_{\underset{\sim}{\boldsymbol{R}}^{-1}}\underline{\boldsymbol{p}} - \underline{\boldsymbol{p}}^H\underset{\sim}{\boldsymbol{R}}^{-1}\underline{\boldsymbol{p}}\\
&= \sigma_d^2 - \underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}_{opt} - \underline{\boldsymbol{w}}_{opt}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}_{opt}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}_{opt} - \underline{\boldsymbol{p}}^H\underset{\sim}{\boldsymbol{R}}^{-1}\underline{\boldsymbol{p}}\\
&= \underbrace{\sigma_d^2 - \underline{\boldsymbol{p}}^H\underset{\sim}{\boldsymbol{R}}^{-1}\underline{\boldsymbol{p}}}_{J_{min}} + \underbrace{(\underline{\boldsymbol{w}}^H - \underline{\boldsymbol{w}}_{opt}^H)}_{\Delta\underline{\boldsymbol{w}}^H}\underset{\sim}{\boldsymbol{R}}\underbrace{(\underline{\boldsymbol{w}} - \underline{\boldsymbol{w}}_{opt})}_{\Delta\underline{\boldsymbol{w}}}\\
&= J_{\min} + \underbrace{\Delta\underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{R}}\Delta\underline{\boldsymbol{w}}}_{>0 \quad \forall\Delta\underline{\boldsymbol{w}}\neq 0 \quad \text{because } \underset{\sim}{\boldsymbol{R}}>0} \qquad ; \qquad \underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_{opt} + \Delta\underline{\boldsymbol{w}}
\end{aligned}$$

EVD:

$$\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{Q}}^H; \qquad \underset{\sim}{\boldsymbol{\Lambda}} = \begin{bmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_M\end{bmatrix} > \underline{\boldsymbol{0}}$$

$$J = J_{\min} + \underbrace{\Delta\underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{Q}}}_{\underline{\boldsymbol{v}}^H}\underset{\sim}{\boldsymbol{\Lambda}}\underbrace{\underset{\sim}{\boldsymbol{Q}}^H\Delta\underline{\boldsymbol{w}}}_{\underline{\boldsymbol{v}}} = J_{\min}\underline{\boldsymbol{v}}^H\underset{\sim}{\boldsymbol{\Lambda}}\underline{\boldsymbol{v}} = J_{\min} + \sum_{m=1}^{M}\lambda_m|v_m|^2$$

$$\underline{\boldsymbol{u}}[n] = \underset{\sim}{\boldsymbol{H}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n]$$

$$\begin{aligned}
\underset{\sim u}{\boldsymbol{R}} &:= E[\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n]] = E[(\underset{\sim}{\boldsymbol{H}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n])(\underline{\boldsymbol{s}}^H[n]\underset{\sim}{\boldsymbol{H}}^H + \underline{\boldsymbol{\nu}}^H[n])]\\
&= \underset{\sim}{\boldsymbol{H}}\underbrace{E[\underline{\boldsymbol{s}}[n]\underline{\boldsymbol{s}}^H[n]]}_{\underset{\sim s}{\boldsymbol{R}}}\underset{\sim}{\boldsymbol{H}}^H + \underset{\sim}{\boldsymbol{H}}\underbrace{E[\underline{\boldsymbol{s}}[n]\underline{\boldsymbol{\nu}}^H[n]]}_{\underset{\sim s\nu}{\boldsymbol{R}}} + \underbrace{E[\underline{\boldsymbol{\nu}}[n]\underline{\boldsymbol{s}}^H[n]]}_{\underset{\sim \nu s}{\boldsymbol{R}} = \underset{\sim s\nu}{\boldsymbol{R}}^H}\underset{\sim}{\boldsymbol{H}}^H + \underbrace{E[\underline{\boldsymbol{\nu}}[n]\underline{\boldsymbol{\nu}}^H[n]]}_{\underset{\sim \nu}{\boldsymbol{R}}}
\end{aligned}$$

In the following: $\underset{\sim \nu s}{\boldsymbol{R}} = 0$

$$\boxed{\underset{\sim u}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{H}}\underset{\sim s}{\boldsymbol{R}}\underset{\sim}{\boldsymbol{H}}^H + \underset{\sim \nu}{\boldsymbol{R}} := \underset{\sim}{\boldsymbol{R}}}$$

$$\underline{\boldsymbol{p}} = E[\underline{\boldsymbol{u}}[n]d^*[n]] = E[(\underset{\sim}{\boldsymbol{H}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n])d^*[n]]; \qquad d[n] = s[n-l]$$

$$\begin{aligned}
&= \underset{\sim}{\boldsymbol{H}}\,E[\underline{\boldsymbol{s}}[n]d^*[n]] + \underbrace{E[\underline{\boldsymbol{\nu}}[n]\cdot s^*[n-l]]}_{=\underline{\boldsymbol{0}}}\\
&= \underset{\sim}{\boldsymbol{H}}\,E[\underline{\boldsymbol{s}}[n]\underbrace{s^*[n-l]}_{\underline{\boldsymbol{s}}^H[n]\cdot\underline{\boldsymbol{e}}_{l+1}}]
\end{aligned}$$

$$\underline{\boldsymbol{s}}^H = \begin{bmatrix} s^*[n] & s^*[n-1] & s^*[n-2] & \cdots & s^*[n-l] & s^*[n-l-1] & \cdots & s^*[n-N+1]\end{bmatrix}$$

$$\underline{\boldsymbol{e}}_{l+1}^T = \begin{bmatrix} 0 & 0 & 0 & \cdots & \underbrace{1}_{l+1} & 0 & \cdots & 0\end{bmatrix}$$

$$\underline{\boldsymbol{p}} = \underset{\sim}{\boldsymbol{H}}\,E[\underline{\boldsymbol{s}}[n]\underline{\boldsymbol{s}}^H[n]]\underline{\boldsymbol{e}}_{l+1}$$

$\underline{\boldsymbol{p}} = \underset{\sim}{\boldsymbol{H}}\underset{\sim}{\boldsymbol{R}}_s \underline{\boldsymbol{e}}_{l+1}$

$\underline{\boldsymbol{w}}_{opt} = \underset{\sim}{\boldsymbol{R}}_u^{-1}\underline{\boldsymbol{p}} = (\underset{\sim}{\boldsymbol{H}}\underset{\sim}{\boldsymbol{R}}_s \underset{\sim}{\boldsymbol{H}}^H + \underset{\sim}{\boldsymbol{R}}_\nu)^{-1}\underset{\sim}{\boldsymbol{H}}\underset{\sim}{\boldsymbol{R}}_s \underline{\boldsymbol{e}}_{l+1}$

$\underset{\sim}{\boldsymbol{H}} \in \mathbb{C}^{M\times N}$

3.2.1 Filter Weight w_{opt} for tall Channel matrices

Example: Receiver with two antennas

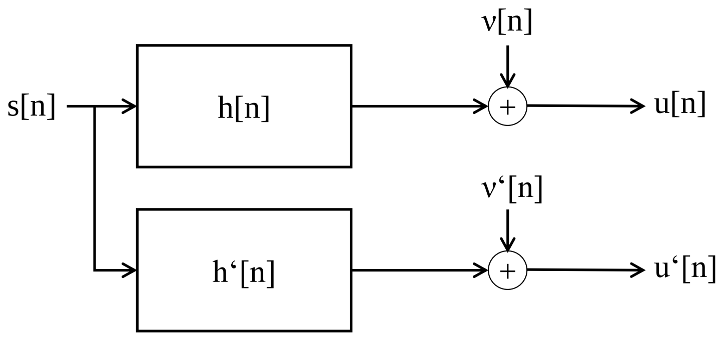

Figure 27: Block diagram of a receiver with two inputs (antennas)

$\underline{\boldsymbol{u}}[n] = \underset{\sim}{\boldsymbol{H}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n]$

$\underline{\boldsymbol{u}}'[n] = \underset{\sim}{\boldsymbol{H}}'\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}'[n]$

$$\underbrace{\begin{bmatrix} \underline{\boldsymbol{u}}[n] \\ \underline{\boldsymbol{u}}'[n] \end{bmatrix}}_{\tilde{\underline{\boldsymbol{u}}}[n]} = \underbrace{\begin{bmatrix} \underset{\sim}{\boldsymbol{H}} \\ \underset{\sim}{\boldsymbol{H}}' \end{bmatrix}}_{\underset{\sim}{\tilde{\boldsymbol{H}}}} \underline{\boldsymbol{s}}[n] + \underbrace{\begin{bmatrix} \underline{\boldsymbol{\nu}}[n] \\ \underline{\boldsymbol{\nu}}'[n] \end{bmatrix}}_{\tilde{\underline{\boldsymbol{\nu}}}[n]}$$

$\tilde{\underline{\boldsymbol{u}}}[n] = \underset{\sim}{\tilde{\boldsymbol{H}}}\underline{\boldsymbol{s}}[n] + \tilde{\underline{\boldsymbol{\nu}}}[n]; \qquad \underset{\sim}{\tilde{\boldsymbol{H}}} \in \mathbb{C}^{2M\times N}$

$\underset{\sim}{\tilde{\boldsymbol{H}}}$ could have more rows than columns:

$(\underset{\sim}{\tilde{\boldsymbol{H}}}\underset{\sim}{\tilde{\boldsymbol{R}}}_s \underset{\sim}{\tilde{\boldsymbol{H}}}^H) \in \mathbb{C}^{2M\times 2M}$

$\Rightarrow$ If $2M > N$ we have a problem with the rank.

Sherman-Morrison-Woodburry Identity

Given $\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{M\times M}$, rank$\underset{\sim}{\boldsymbol{A}} = M$ and $\underset{\sim}{\boldsymbol{C}} \in \mathbb{C}^{N\times N}$, rank$\underset{\sim}{\boldsymbol{C}} = N$ and $\underset{\sim}{\boldsymbol{B}} \in \mathbb{C}^{M\times N}$ and $\underset{\sim}{\boldsymbol{D}} \in \mathbb{C}^{N\times M}$, then:

$(\underset{\sim}{\boldsymbol{A}} + \underset{\sim}{\boldsymbol{B}}\underset{\sim}{\boldsymbol{C}}\underset{\sim}{\boldsymbol{D}})^{-1} = \underset{\sim}{\boldsymbol{A}}^{-1} - \underset{\sim}{\boldsymbol{A}}^{-1}\underset{\sim}{\boldsymbol{B}}(\underset{\sim}{\boldsymbol{C}}^{-1} + \underset{\sim}{\boldsymbol{D}}\underset{\sim}{\boldsymbol{A}}^{-1}\underset{\sim}{\boldsymbol{B}})^{-1}\underset{\sim}{\boldsymbol{D}}\underset{\sim}{\boldsymbol{A}}^{-1}$

Proof:

$$(\utilde{\boldsymbol{A}}^{-1} - \utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}}(\utilde{\boldsymbol{C}}^{-1} + \utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}})^{-1}\utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1})(\utilde{\boldsymbol{A}} + \utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}}\utilde{\boldsymbol{D}})$$

$$= \underbrace{\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{A}}}_{\utilde{\boldsymbol{I}}} + \utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}}\utilde{\boldsymbol{D}} - \utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}}(\utilde{\boldsymbol{C}}^{-1} + \utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}})^{-1}\utilde{\boldsymbol{C}}^{-1}\utilde{\boldsymbol{C}}\utilde{\boldsymbol{D}}\underbrace{\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{A}}}_{\utilde{\boldsymbol{I}}} - \utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}}(\utilde{\boldsymbol{C}}^{-1} + \utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}})^{-1}\utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}}\utilde{\boldsymbol{C}}\utilde{\boldsymbol{D}}$$

$$= \utilde{\boldsymbol{I}} + \utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}}(\utilde{\boldsymbol{I}}\underbrace{-(\utilde{\boldsymbol{C}}^{-1} + \utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}})^{-1}\utilde{\boldsymbol{C}}^{-1} - (\utilde{\boldsymbol{C}}^{-1} + \utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}})^{-1}\utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}}}_{-\underbrace{(\utilde{\boldsymbol{C}}^{-1} + \utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}})^{-1}(\utilde{\boldsymbol{C}}^{-1} + \utilde{\boldsymbol{D}}\utilde{\boldsymbol{A}}^{-1}\utilde{\boldsymbol{B}})}_{\utilde{\boldsymbol{I}}}})\utilde{\boldsymbol{C}}\utilde{\boldsymbol{D}}$$

$$= \utilde{\boldsymbol{I}} \qquad \square$$

Applying the Sherman-Morrison-Woodbury Identity on $\underline{\boldsymbol{w}}_{opt}$ with $\utilde{\boldsymbol{B}} = \utilde{\boldsymbol{H}}$, $\utilde{\boldsymbol{C}} = \utilde{\boldsymbol{R}}_s$, $\utilde{\boldsymbol{D}} = \utilde{\boldsymbol{H}}^H$ and $\utilde{\boldsymbol{A}} = \utilde{\boldsymbol{R}}_\nu$ we get:

$$\begin{aligned}
\underline{\boldsymbol{w}}_{opt} &= (\utilde{\boldsymbol{H}}\utilde{\boldsymbol{R}}_s\utilde{\boldsymbol{H}}^H + \utilde{\boldsymbol{R}}_\nu)^{-1}\utilde{\boldsymbol{H}}\utilde{\boldsymbol{R}}_s\underline{\boldsymbol{e}}_{l+1} \\
&= (\utilde{\boldsymbol{R}}_\nu^{-1} - \utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}}(\utilde{\boldsymbol{R}}_s^{-1} + \utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}})^{-1}\utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1})\utilde{\boldsymbol{H}}\utilde{\boldsymbol{R}}_s\underline{\boldsymbol{e}}_{l+1} \\
&= \utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}}(\underbrace{\utilde{\boldsymbol{I}}}_{(\utilde{\boldsymbol{R}}_s^{-1}+\utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}})^{-1}(\utilde{\boldsymbol{R}}_s^{-1}+\utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}})}(-(\utilde{\boldsymbol{R}}_s^{-1} + \utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}})^{-1}\utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}})\utilde{\boldsymbol{R}}_s\underline{\boldsymbol{e}}_{l+1} \\
&= \utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}}(\utilde{\boldsymbol{R}}_s^{-1} + \utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}})^{-1}\underbrace{(\utilde{\boldsymbol{R}}_s^{-1} + \underbrace{\utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}} - \utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}}}_{=0})\utilde{\boldsymbol{R}}_s}_{=\utilde{\boldsymbol{I}}}\underline{\boldsymbol{e}}_{l+1}
\end{aligned}$$

$$\boxed{\begin{aligned}
\underline{\boldsymbol{w}}_{opt} &= \utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}}\underbrace{(\utilde{\boldsymbol{R}}_s^{-1} + \utilde{\boldsymbol{H}}^H\utilde{\boldsymbol{R}}_\nu^{-1}\utilde{\boldsymbol{H}})^{-1}}_{N\times N}\underline{\boldsymbol{e}}_{l+1} && (1) \\
&= \underbrace{(\utilde{\boldsymbol{R}}_\nu + \utilde{\boldsymbol{H}}\utilde{\boldsymbol{R}}_s\utilde{\boldsymbol{H}}^H)^{-1}}_{M\times M}\utilde{\boldsymbol{H}}\utilde{\boldsymbol{R}}_s\underline{\boldsymbol{e}}_{l+1} && (2)
\end{aligned}}$$

$\utilde{\boldsymbol{H}} \in \mathbb{C}^{M\times N}$

If $M < N$ use equation (2): M equations in M unknowns

If $N < M$ use equation (1): N equations in N unknowns

Note: processor & memory load

$\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}: \quad \underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{n\times n}$

Solve for $\underline{\boldsymbol{x}}$: computational load $\sim n^3$

memory $\sim n^2$

Note: Inter-Symbol-Interference

$$y[n] = \underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n] = \underline{\boldsymbol{w}}(\underset{\sim}{\boldsymbol{H}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n]) = \underline{\boldsymbol{w}}^H(\underbrace{\begin{bmatrix}\underline{\boldsymbol{h}}_1 & \underline{\boldsymbol{h}}_2 & \cdots & \underline{\boldsymbol{h}}_N\end{bmatrix}}_{\underset{\sim}{\boldsymbol{H}}} \underbrace{\underline{\boldsymbol{s}}[n]}_{\begin{bmatrix} s[n] \\ s[n-1] \\ \vdots \\ s[n-M+1]\end{bmatrix}} + \underline{\boldsymbol{\nu}}[n])$$

$d[n] = s[n-1]$

$$y[n] = \underline{\boldsymbol{w}}^H(\underline{\boldsymbol{h}}_{l+1} \underbrace{\underline{\boldsymbol{s}}[n-l]}_{\text{signal of interest}} + \underbrace{\sum_{k=1;k\neq l+1}^{N} \underline{\boldsymbol{h}}_k \underline{\boldsymbol{s}}[n-k]}_{\text{Inter-Symbol-Interference}} + \underbrace{\underline{\boldsymbol{\nu}}[n])}_{\text{Noise}}$$

3.3 Filter with different kinds of Noise

3.3.1 Special case: White Noise & White Signal

$\underset{\sim}{\boldsymbol{R}}_s = \sigma_d^2 \underset{\sim}{\boldsymbol{I}}; \qquad \underset{\sim}{\boldsymbol{R}}_\nu = \sigma_\nu^2 \underset{\sim}{\boldsymbol{I}}$

$\underline{\boldsymbol{w}}_{opt} = (\frac{\sigma_u^2}{\sigma_d^2}\underset{\sim}{\boldsymbol{I}} + \underset{\sim}{\boldsymbol{H}}\underset{\sim}{\boldsymbol{H}}^H)^{-1}\underset{\sim}{\boldsymbol{H}}\underline{\boldsymbol{e}}_{l+1}$

3.3.2 Case 1: High SNR (signal to noise ratio)

$\underset{\sim}{\boldsymbol{H}} \in \mathbb{C}^{M\times N}; \qquad M < N$

rank$\underset{\sim}{\boldsymbol{H}} = M$ (full rank $\underset{\sim}{\boldsymbol{H}}$)

$$\underset{\sim}{\boldsymbol{H}} = \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H \underbrace{=}_{\mathrm{rank}\underset{\sim}{\boldsymbol{H}}=M} \underset{\sim}{\boldsymbol{U}} \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H$$

$$\begin{aligned}
(\underset{\sim}{\boldsymbol{H}}\underset{\sim}{\boldsymbol{H}}^H)^{-1}\underset{\sim}{\boldsymbol{H}} &= (\underset{\sim}{\boldsymbol{U}} \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H \underset{\sim}{\boldsymbol{V}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{U}}^H)^{-1} \underset{\sim}{\boldsymbol{U}} \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H \\
&= (\underset{\sim}{\boldsymbol{U}} \underset{\sim}{\boldsymbol{\Sigma}}_1^2 \underset{\sim}{\boldsymbol{U}}^H)^{-1} \underset{\sim}{\boldsymbol{U}} \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H \\
&= \underset{\sim}{\boldsymbol{U}} \underset{\sim}{\boldsymbol{\Sigma}}_1^{-2} \underbrace{\underset{\sim}{\boldsymbol{U}}^H \underset{\sim}{\boldsymbol{U}}}_{\underset{\sim}{\boldsymbol{I}}} \underset{\sim}{\boldsymbol{\Sigma}}_1 \underset{\sim}{\boldsymbol{V}}_1^H \\
&= \underset{\sim}{\boldsymbol{U}} \underset{\sim}{\boldsymbol{\Sigma}}_1^{-1} \underset{\sim}{\boldsymbol{V}}_1^H = \underset{\sim}{\boldsymbol{U}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1^{-1} \underset{\sim}{\boldsymbol{V}}_1^H = (\underset{\sim}{\boldsymbol{V}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_1^{-1} \underset{\sim}{\boldsymbol{U}}^H)^H \\
&= (\underset{\sim}{\boldsymbol{H}}^+)^H = (\underset{\sim}{\boldsymbol{H}}^H)^+
\end{aligned}$$

$$\lim_{\frac{\sigma_d^2}{\sigma_\nu^2}\to\infty} \underline{\boldsymbol{w}}_{opt} = (\underset{\sim}{\boldsymbol{H}}\underset{\sim}{\boldsymbol{H}}^H)^{-1}\underset{\sim}{\boldsymbol{H}}\underline{\boldsymbol{e}}_{l+1} = \underbrace{(\underset{\sim}{\boldsymbol{H}}^H)^+}_{\text{Moore-Penrose-Pseudo-Inverse}} \underline{\boldsymbol{e}}_{l+1}$$

No optimal solution because:

Try: $\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} \overset{!}{=} \underline{\boldsymbol{e}}_{l+1}^T$; $\quad y[n] = \underbrace{\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}}}_{\underline{\boldsymbol{e}}_{l+1}^T} \underline{\boldsymbol{s}}[n] = \underline{\boldsymbol{s}}[n-1]$

However it does not work!

$\underset{\sim}{\boldsymbol{H}}^H \underline{\boldsymbol{w}} = \underline{\boldsymbol{e}}_{l+1}$

$\underset{\sim}{\boldsymbol{H}} \in \mathbb{C}^{N\times M} \qquad$ with $M < N$

Number of degrees of freedom $= M$

Number of constraints $= N > M$

$\Rightarrow$ No exact solution exists!

$\Rightarrow$ LS-Solution (least square)

$$\boxed{\underline{\boldsymbol{w}}_{LS} = (\underset{\sim}{\boldsymbol{H}}^H)^+ \underline{\boldsymbol{e}}_{l+1}}$$

For high SNR the optimum filter tries to suppress the interference but lack DoF (Degrees of Freedom) to do it perfectly. $\underline{\boldsymbol{w}}_{LS} = \arg\min_{\underline{\boldsymbol{w}}} \lvert\lvert \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} - \underline{\boldsymbol{e}}_{l+1} \rvert\rvert_2^2$

The higher the dimensions, the better the outcome of the filter.

Work-around: Block processing

at Tx: pre-pend K guard symbols

at Rx: discard first K received symbols

Example:

$K = 2$ (number of guard symbols), $\quad M = 4$ (number of payload symbols)

$$\begin{bmatrix} u[n] \\ u[n-1] \\ u[n-2] \\ u[n-3] \end{bmatrix} = \underline{\boldsymbol{\nu}}[n] + \begin{bmatrix} h_0 & h_1 & h_2 & 0 & 0 & 0 \\ 0 & h_0 & h_1 & h_2 & 0 & 0 \\ 0 & 0 & h_0 & h_1 & h_2 & 0 \\ 0 & 0 & 0 & h_0 & h_1 & h_2 \end{bmatrix} \begin{bmatrix} s[n] \\ s[n-1] \\ s[n-2] \\ s[n-3] \\ s[n] \\ s[n-1] \end{bmatrix}$$

$$\underset{\sim}{\mathring{\boldsymbol{H}}} = \begin{bmatrix} h_0 & h_1 & h_2 & 0 \\ 0 & h_0 & h_1 & h_2 \\ h_2 & 0 & h_0 & h_1 \\ h_1 & h_2 & 0 & h_0 \end{bmatrix} \text{(cyclic matrix)}$$

$\underset{\sim}{\mathring{\boldsymbol{H}}}$ is invertible $\quad \Rightarrow \quad$ with no noise, we get a perfect solution with losing bandwidth efficiency

$$\underline{\boldsymbol{s}}[n] = \begin{bmatrix} s[n] \\ s[n-1] \\ s[n-2] \\ s[n-3] \end{bmatrix}$$

Example MATLAB-code: `Q=fft(eye(M))/sqrt(M)`

$$\underset{\sim}{\boldsymbol{Q}}^H \begin{bmatrix} H \end{bmatrix} \begin{bmatrix} Q \end{bmatrix} = \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & a_3 & \\ & & & a_4 \end{bmatrix}$$

$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix} = \text{fft}(\begin{bmatrix} h_0 & h_1 & h_2 & 0 \end{bmatrix}) \Rightarrow$ OFDM (Orthogonal Frequency Division Multiplexing)

3.3.3 Case 2: Low SNR (High Noise)

$\frac{\sigma_\nu^2}{\sigma_d^2} \to \infty$

$y[n] \approx \underline{\boldsymbol{w}}^H (\underline{\boldsymbol{h}}_{l+1} s[n-l] + \underline{\boldsymbol{\nu}}[n])$

$SNR = \frac{E[|y[n]|^2 | \underline{\boldsymbol{\nu}}[n] = \underline{\mathbf{0}}]}{E[|y[n]|^2 | s[n-l] = 0]} = \frac{\sigma_d^2 |\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}_{l+1}|^2}{\sigma_\nu^2 \underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}}$

Calculation of the SNR with respect to variances of signal and noise

System: $y[n] = \underline{\boldsymbol{w}}^H \underline{\boldsymbol{u}}[n] = \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} \underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n]$

with $\quad \underline{\boldsymbol{s}}[n] = \underline{\boldsymbol{e}}_{l+1} \cdot d[n]$

$$
\begin{aligned}
&E[|y[n]|^2 | \underline{\boldsymbol{\nu}}[n] = \underline{\mathbf{0}}] \\
=&E[(\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} \underline{\boldsymbol{s}}[n])^* (\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} \underline{\boldsymbol{s}}[n])] \\
=&E[(\underline{\boldsymbol{s}}[n]^H \underset{\sim}{\boldsymbol{H}}^H \underline{\boldsymbol{w}})(\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} \underline{\boldsymbol{s}}[n])] \\
=&E[\underline{\boldsymbol{s}}[n]^H \underset{\sim}{\boldsymbol{H}}^H \underline{\boldsymbol{w}}\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} \underline{\boldsymbol{s}}[n]] \\
=&E[d[n]^* d[n] \cdot \underline{\boldsymbol{e}}_{l+1}^H \underset{\sim}{\boldsymbol{H}}^H \underline{\boldsymbol{w}}\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{H}} \underline{\boldsymbol{e}}_{l+1}] \\
=&E[|d[n]|^2 \cdot \underbrace{\underline{\boldsymbol{h}}_{l+1}^H \underline{\boldsymbol{w}}}_{\in \mathbb{C}^{1\times 1}} \underbrace{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}_{l+1}}_{\in \mathbb{C}^{1\times 1}}] \\
=&E[\underbrace{|d[n]|^2}_{\text{random}} \cdot \underbrace{\left| \underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}_{l+1} \right|^2}_{\text{deterministic}}] \\
=&E[|d[n]|^2] \cdot \left| \underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}_{l+1} \right|^2] \\
=&\sigma_d^2 \cdot \left| \underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}_{l+1} \right|^2
\end{aligned}
$$

$$
\begin{aligned}
&E[|y[n]|^2 | s[n-l] = 0] \\
=&E[(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n])^* (\underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n])] \\
=&E[(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n])^H (\underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n])] \\
=&E[\underline{\boldsymbol{\nu}}[n]^H \underline{\boldsymbol{w}}\underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n]] \\
=&E[\mathrm{tr}(\underline{\boldsymbol{\nu}}[n]^H \underline{\boldsymbol{w}}\underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n])] \\
=&E[\mathrm{tr}(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{\nu}}[n] \underline{\boldsymbol{\nu}}[n]^H \underline{\boldsymbol{w}})] \\
=&\,\mathrm{tr}(\underline{\boldsymbol{w}}^H E[\underline{\boldsymbol{\nu}}[n] \underline{\boldsymbol{\nu}}[n]^H] \underline{\boldsymbol{w}}) \\
=&\,\mathrm{tr}(\underline{\boldsymbol{w}}^H (\sigma_\nu^2 \underset{\sim}{\boldsymbol{I}}) \underline{\boldsymbol{w}}) \\
=&\sigma_\nu^2 \cdot \mathrm{tr}(\underbrace{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}}_{\in \mathbb{C}^{1\times 1}}) = \sigma_\nu^2 \cdot \underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}
\end{aligned}
$$

Mathmatical derivation of Cauchy-Schwarz-Inequality:

$\underline{\boldsymbol{z}} := \underline{\boldsymbol{h}} - \underline{\boldsymbol{w}} \frac{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}}$

$$
\begin{aligned}
\underline{\boldsymbol{z}}^H \underline{\boldsymbol{z}} &= (\underline{\boldsymbol{h}}^H - \frac{\underline{\boldsymbol{h}}^H \underline{\boldsymbol{w}}}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}} \underline{\boldsymbol{w}})(\underline{\boldsymbol{h}} - \underline{\boldsymbol{w}} \frac{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}}) \\
&= \underline{\boldsymbol{h}}^H \underline{\boldsymbol{h}} - \frac{(\underline{\boldsymbol{h}}^H \underline{\boldsymbol{w}})(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}})}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}} - \frac{(\underline{\boldsymbol{h}}^H \underline{\boldsymbol{w}})(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}})}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}} + \frac{(\underline{\boldsymbol{h}}^H \underline{\boldsymbol{w}})(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}) \underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}}{(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}})(\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}})} \\
&= \underline{\boldsymbol{h}}^H \underline{\boldsymbol{h}} - \frac{|\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}|^2}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}} \geq 0
\end{aligned}
$$

Cauchy-Schwarz-Inequality:
$\frac{\lvert\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}\rvert^2}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}} \leq \underline{\boldsymbol{h}}^H \underline{\boldsymbol{h}}$

$SNR = \frac{\sigma_d^2}{\sigma_\nu^2} \frac{|\underline{\boldsymbol{w}}^H \underline{\boldsymbol{h}}_{l+1}|^2}{\underline{\boldsymbol{w}}^H \underline{\boldsymbol{w}}} \leq \frac{\sigma_d^2}{\sigma_\nu^2} \underline{\boldsymbol{h}}_{l+1}^H \underline{\boldsymbol{h}}_{l+1}$

$\underline{\boldsymbol{w}}_{MF} = \arg\max_{\underline{\boldsymbol{w}}} SNR = \mathtt{const} \cdot \underline{\boldsymbol{h}}_{l+1}$ with $\mathtt{const} \neq 0$ where MF stands for "matched filter"

try: $SNR \,|_{\underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_{MF}} = \frac{\sigma_d^2}{\sigma_\nu^2} \frac{\left| \underline{\boldsymbol{h}}_{l+1}^H \underline{\boldsymbol{h}}_{l+1} \right|^2 \cdot |\mathtt{const}|^2}{\underline{\boldsymbol{h}}_{l+1}^H \underline{\boldsymbol{h}}_{l+1} \cdot |\mathtt{const}|^2} = \frac{\sigma_d^2}{\sigma_\nu^2} \underline{\boldsymbol{h}}_{l+1}^H \underline{\boldsymbol{h}}_{l+1}$

$\underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_{MF}$ achieves an upper bound of the SNR.

Optimum Filter Solution:

$$\underline{\boldsymbol{w}}_{opt} = \underbrace{\frac{\sigma_d^2}{\sigma_\nu^2}}_{\texttt{const}} \underline{\boldsymbol{h}}_{l+1} = \underline{\boldsymbol{w}}_{MF}$$

$$\lim_{\frac{\sigma_d^2}{\sigma_\nu^2}\to 0} \underline{\boldsymbol{w}}_{opt} = \underline{\boldsymbol{w}}_{MF} \sim \underline{\boldsymbol{h}}_{l+1}$$

3.4 Steepest Descent Algorithm (SDA)

$\underline{\boldsymbol{w}}_{opt} = \underset{\sim}{\boldsymbol{R}}^{-1}\underline{\boldsymbol{p}}, \qquad \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}_{opt} = \underline{\boldsymbol{p}}$

One method to solve this problem woulbd be the Gauß-Elimination but since $\underset{\sim}{\boldsymbol{R}}$ and $\underline{\boldsymbol{p}}$ usually change it is neccessary to recalulate everything when ever something changes. This would lead to an enormouse effort. So an iterative solution would be more productive.

Iterative solution:

$\underline{\boldsymbol{w}}[n+1] = \underline{\boldsymbol{w}}[n] + \Delta\underline{\boldsymbol{w}}[n]$

$J = \sigma_d^2 - \underline{\boldsymbol{w}}^H\underline{\boldsymbol{p}} - \underline{\boldsymbol{p}}^H\underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}^H\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}; \quad \underset{\sim}{\boldsymbol{R}} > \underline{\boldsymbol{0}}; \qquad \underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H$

$dJ = \left(\frac{\partial J}{\partial \underline{\boldsymbol{w}}}\right)^T d\underline{\boldsymbol{w}} + \left(\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*}\right)^T d\underline{\boldsymbol{w}}^*$

with $\left(\frac{\partial J}{\partial \underline{\boldsymbol{w}}}\right)^T d\underline{\boldsymbol{w}} = \left(\left(\frac{\partial J^*}{\partial \underline{\boldsymbol{w}}^*}\right)^T d\underline{\boldsymbol{w}}^*\right)^* \underbrace{=}_{J \underbrace{=}_{J\in\mathbb{R}} J^*} \left(\left(\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*}\right)^T d\underline{\boldsymbol{w}}^*\right)^*$ we get:

$dJ = 2\,\mathrm{Re}\left\{\left(\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*}\right)^T d\underline{\boldsymbol{w}}^*\right\}$

$dJ = 2\,\mathrm{Re}\left\{\left(\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*}\right)^H d\underline{\boldsymbol{w}}\right\} \le 2\left|\left(\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*}\right)^H d\underline{\boldsymbol{w}}\right|$

$d\underline{\boldsymbol{w}} = \frac{\partial J}{\partial \underline{\boldsymbol{w}}^*}\cdot \texttt{const}^*; \qquad \texttt{const} < 0$

$\Delta\underline{\boldsymbol{w}}[n] = -\mu\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*[n]}; \qquad \mu > 0$

$\frac{\partial J}{\partial \underline{\boldsymbol{w}}^*[n]} = -\underline{\boldsymbol{p}} + \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}[n]$

$$\begin{aligned}\underline{\boldsymbol{w}}[n+1] &= \underline{\boldsymbol{w}}[n] + \Delta\underline{\boldsymbol{w}}[n] = \underline{\boldsymbol{w}}[n] - \mu\left(\underline{\boldsymbol{p}} + \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}[n]\right); \qquad \mu > 0\\ &= (\underset{\sim}{\boldsymbol{I}} - \mu\underset{\sim}{\boldsymbol{R}})\underline{\boldsymbol{w}}[n] + \mu\underline{\boldsymbol{p}}\end{aligned}$$

Does $\underline{\boldsymbol{w}}[n]$, $\underline{\boldsymbol{w}}[n+1], \ldots$ converge to $\underline{\boldsymbol{w}}_{opt}$?

$\underline{\boldsymbol{c}}[n] = \underline{\boldsymbol{w}}[n] - \underline{\boldsymbol{w}}_{opt}$

$\underbrace{c[n+1] + \underline{\boldsymbol{w}}_{opt}}_{=\underline{\boldsymbol{w}}[n+1]} = (\underset{\sim}{\boldsymbol{I}} - \mu\underset{\sim}{\boldsymbol{R}})\,\underbrace{\underline{\boldsymbol{c}}[n] + \underline{\boldsymbol{w}}_{opt}}_{\underline{\boldsymbol{w}}[n]} + \mu\underline{\boldsymbol{p}}$

$$c[n+1] = (\underset{\sim}{\boldsymbol{I}} - \mu\underset{\sim}{\boldsymbol{R}})\underline{\boldsymbol{c}}[n] - \underbrace{\mu\underset{\sim}{\boldsymbol{R}}\overbrace{\underline{\boldsymbol{w}}_{opt}}^{\underset{\sim}{\boldsymbol{R}}^{-1}\underline{\boldsymbol{p}}} + \mu\underline{\boldsymbol{p}}}_{=\underline{\boldsymbol{0}}}$$

$$\boxed{c[n+1] = (\underset{\sim}{\boldsymbol{I}} - \mu\underset{\sim}{\boldsymbol{R}})\underline{\boldsymbol{c}}[n]}$$

Does $c[n]$, $c[n+1]$, $\ldots$ converge to $\underline{\boldsymbol{0}}$?

EVD: $\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{Q}}^{-1} \underbrace{=}_{\underset{\sim}{\boldsymbol{R}}=\underset{\sim}{\boldsymbol{R}}^H} \underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{Q}}^H$

$$\underset{\sim}{\boldsymbol{\Lambda}} = \begin{bmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_M \end{bmatrix}; \quad \lambda_i > 0; \quad i \in \{1, 2, \ldots, M\}; \qquad (\underset{\sim}{\boldsymbol{R}} > 0)$$

$$\underline{\boldsymbol{c}}[n+1] = (\underset{\sim}{\boldsymbol{I}} - \mu\underbrace{\underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{Q}}^H}_{=\underset{\sim}{\boldsymbol{R}}})\underline{\boldsymbol{c}}[n] = (\underbrace{\underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{Q}}^H}_{=\underset{\sim}{\boldsymbol{I}}} - \mu\underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{Q}}^H)\underline{\boldsymbol{c}}[n]$$

$$= \underset{\sim}{\boldsymbol{Q}}(\underset{\sim}{\boldsymbol{I}} - \mu\underset{\sim}{\boldsymbol{\Lambda}})\underset{\sim}{\boldsymbol{Q}}^H\underline{\boldsymbol{c}}[n]$$

$$\underbrace{\underset{\sim}{\boldsymbol{Q}}^H\underline{\boldsymbol{c}}[n+1]}_{\underline{\boldsymbol{z}}[n+1]} = (\underset{\sim}{\boldsymbol{I}} - \mu\underset{\sim}{\boldsymbol{\Lambda}})\underbrace{\underset{\sim}{\boldsymbol{Q}}^H\underline{\boldsymbol{c}}[n]}_{\underline{\boldsymbol{z}}[n]}$$

$$\boxed{\underline{\boldsymbol{z}}[n+1] = (\underset{\sim}{\boldsymbol{I}} - \mu\underset{\sim}{\boldsymbol{\Lambda}})\underline{\boldsymbol{z}}[n]}$$

$$\underline{\boldsymbol{z}}[n] = \begin{bmatrix} z_1[n] & z_2[n] & \cdots & z_M[n] \end{bmatrix}^T$$

$$\boxed{z_k[n+1] = (1 - \mu\lambda_k)z_k[n]; \qquad 1 \le k \le M}$$

$z_k[n]$, $z_k[n+1]\ldots$ converge to 0 if and only if $|1 - \mu\lambda_k| < 1; \qquad \forall k$

case 1: $1 - \mu\lambda_k \ge 0 \qquad \left[\mu \le \frac{1}{\lambda_k}\right]$

then $1 - \mu\lambda_k < 1$

$$-\mu\lambda_k < 0; \qquad \mu > 0, \quad \lambda_l > 0$$

$$-\mu < 0$$

$$\mu > 0$$

$$\boxed{0 < \mu \le \frac{1}{\lambda_k}}$$

case 2: $1 - \mu\lambda_k \leq 0 \qquad \left[\mu \geq \frac{1}{\lambda_k}\right]$

then $\;-1 + \mu\lambda_k < 1$

$$\mu\lambda_k < 2$$

$$\mu < \frac{2}{\lambda_k}$$

$\frac{1}{\lambda_k} \leq \mu \leq \frac{2}{\lambda_k}$

Case 1 & Case 2:
$\Rightarrow 0 < \mu < \frac{2}{\lambda_k} \Leftrightarrow z_k[n],\, z_k[n+1],\, \ldots$ converges to 0.

Key results:
$\underline{\boldsymbol{w}}[n],\, \underline{\boldsymbol{w}}[n+1], \ldots$ converge to $\underline{\boldsymbol{w}}_{opt}$ iff
$0 \leq \mu \leq \frac{2}{\lambda_{max}} \qquad \lambda_{max} =$ maximum Eigenvalue of $\underset{\sim}{\boldsymbol{R}}$

Computation of λ_{max} is costly!

$$\underbrace{\sum_{k=1}^{M} \lambda_k}_{\mathrm{tr}\underset{\sim}{\boldsymbol{R}}} > \lambda_{max} < \mathrm{tr}\,\underset{\sim}{\boldsymbol{R}}$$

$$0 < \mu < \frac{2}{\mathrm{tr}\underset{\sim}{\boldsymbol{R}}} < \frac{2}{\lambda_{max}}$$

$\mathrm{tr}\underset{\sim}{\boldsymbol{R}}$ is cheap to compute. Therfor in practice $0 < \mu < \frac{2}{\mathrm{tr}\underset{\sim}{\boldsymbol{R}}}$ can be used. This is sufficeint for convergence. But usually one is interested in using an optimum step-size μ to achive optimum performance of the algorithm.

3.4.1 Optimum step-size μ

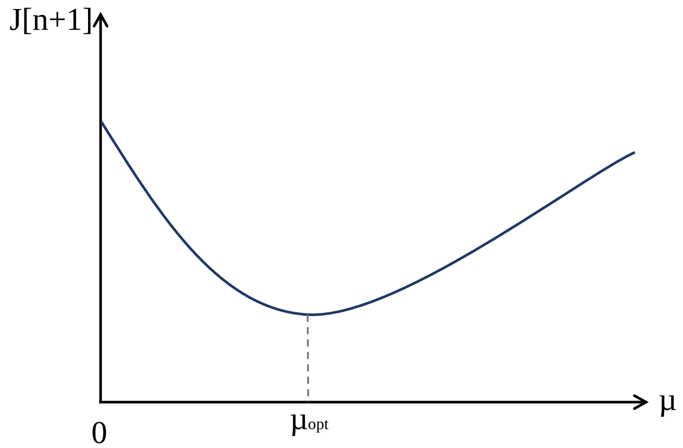

Figure 28: Example characteristics of the cost function $J[n+1]$ plotted over the step-size μ

Solve $\frac{\partial J[n+1]}{\partial \mu^*} = 0$ for $\mu = \mu_{opt}$:

$J[n+1] = \sigma_d^2 - \underline{\boldsymbol{P}}^H \underline{\boldsymbol{w}}[n+1] - \underline{\boldsymbol{w}}^H[n+1]\underline{\boldsymbol{P}} + \underline{\boldsymbol{w}}^H[n+1]\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}[n+1]$

$\underline{\boldsymbol{w}}[n+1] = \underline{\boldsymbol{w}}[n] + \mu \underline{\boldsymbol{r}}[n] \quad$ with $\underline{\boldsymbol{r}}[n] = \underline{\boldsymbol{p}} - \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}[n]$

$$\boxed{\mu_{opt} = \frac{\underline{\boldsymbol{r}}^H[n]\underline{\boldsymbol{r}}[n]}{\underline{\boldsymbol{r}}^H[n]\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{r}}[n]}}$$

Note: The optimum step-size μ_{opt} is (usually) different for each step.

3.4.2 Summary: Steepest Descent Algorithm

Find the optimal $\underline{\boldsymbol{w}}_{opt}$ for $\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}} = \underline{\boldsymbol{p}}$

Input: $\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{R}}^H > \underline{\boldsymbol{0}}, \qquad \underline{\boldsymbol{p}}$

Output: $\underline{\boldsymbol{w}}$ that solves $\underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}} = \underline{\boldsymbol{p}}$ approximately.
This $\underline{\boldsymbol{w}}$ is the minimum of $J = \sigma_d^2 - \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}} - \underline{\boldsymbol{p}}^H \underline{\boldsymbol{w}} + \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}$

Algorithm:

1. Init: $\underline{\boldsymbol{w}} \leftarrow 0$
2. Iteration: $\underline{\boldsymbol{r}} \leftarrow \underline{\boldsymbol{p}} - \underset{\sim}{\boldsymbol{R}}\underline{\boldsymbol{w}}$
   $\mu \leftarrow \frac{\underline{\boldsymbol{r}}^H \underline{\boldsymbol{r}}}{\underline{\boldsymbol{r}}^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{r}}}$
   $\underline{\boldsymbol{w}} \leftarrow \underline{\boldsymbol{w}} + \mu \underline{\boldsymbol{r}}$
3. Go to step 2 until $\underline{\boldsymbol{r}}^H \underline{\boldsymbol{r}} < \varepsilon$

with optimum μ	fixed μ
$4M^2 + 5M$	$2M^2 + 4M$

Table 2: Number of operations $(\cdot, +)$ per iteration

MATLAB experiment to show the number of steps for optimum μ:

Random $\underset{\sim}{\boldsymbol{R}} \in \mathbb{C}^{M \times M}, \quad \underset{\sim}{\boldsymbol{R}} > 0, \quad \frac{\lambda_{max}}{\lambda_{min}} \leq 10$

Random $\underline{\boldsymbol{p}} \in \mathbb{C}^{M \times 1}$

Run SDA with optimum μ, count interations

Repeat many times → <u>median</u> iteartion number

M	10	100	1000
number of iterations (median)	32	36	40

Table 3: Results of MATLAB experiment to show the number of steps for optimum μ

$\Rightarrow$ number of iterations $= 28 + 4\log_{10} M$

This algorithm is perfect for parallel computation

Example: Linear Constrained Optimization

$\min \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}, \quad \text{s. t. } \underset{\sim}{\boldsymbol{B}} \underline{\boldsymbol{w}} = \underline{\boldsymbol{c}}; \qquad \underset{\sim}{\boldsymbol{A}} > \underline{\boldsymbol{0}}$

$\underset{\sim}{\boldsymbol{B}} = \begin{bmatrix} \underset{\sim}{\boldsymbol{U}}_1 & \underset{\sim}{\boldsymbol{U}}_2 \end{bmatrix} \cdot \begin{bmatrix} \underset{\sim}{\boldsymbol{\Sigma}}_1 & \underset{\sim}{\boldsymbol{0}} \\ \underset{\sim}{\boldsymbol{0}} & \underset{\sim}{\boldsymbol{0}} \end{bmatrix} \cdot \begin{bmatrix} \underset{\sim}{\boldsymbol{V}}_1^H \\ \underset{\sim}{\boldsymbol{V}}_2^H \end{bmatrix}$

with $\underset{\sim}{\boldsymbol{B}} \in \mathbb{C}^{q\times M}$, $\underset{\sim}{\boldsymbol{\Sigma}}_1 \in \mathbb{C}^{q\times q}$, $\underset{\sim}{\boldsymbol{V}}_1^H \in \mathbb{C}^{q\times M}$ and $\underset{\sim}{\boldsymbol{V}}_2^H \in \mathbb{C}^{(M-q)\times M}$

$q \leq M; \qquad \operatorname{null} \underset{\sim}{\boldsymbol{B}} = \operatorname{im} \underset{\sim}{\boldsymbol{V}}_2$

$\operatorname{rank} \underset{\sim}{\boldsymbol{B}} = q$

$$\underset{\sim}{\boldsymbol{B}} \underline{\boldsymbol{w}} = \underline{\boldsymbol{c}}: \quad \underline{\boldsymbol{w}} \in \left\{ \underbrace{\underset{\sim}{\boldsymbol{B}}^+ \underline{\boldsymbol{c}}}_{\underline{\boldsymbol{w}}_q} - \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{w}}_a \mid \underline{\boldsymbol{w}}_a \in \mathbb{C}^{(M-q)\times 1} \right\}$$

$\underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_q - \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{w}}_a$

$\underline{\boldsymbol{w}}_q$ is the optimum solution. It is a particular solution of the linear equation system. Under constraint equation system

$\underline{\boldsymbol{w}}_a$ is used to project the null space

$$\begin{aligned}
\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}} &= (\underline{\boldsymbol{w}}_q^H - \underline{\boldsymbol{w}}_a^H \underset{\sim}{\boldsymbol{V}}_2^H) \underset{\sim}{\boldsymbol{A}} (\underline{\boldsymbol{w}}_q - \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{w}}_a) \\
&= \underbrace{\underline{\boldsymbol{w}}_q^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}_q}_{\sigma_d^2} - \underbrace{\underline{\boldsymbol{w}}_q^H \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{V}}_2}_{\underline{\boldsymbol{p}}^H} \underline{\boldsymbol{w}}_a - \underline{\boldsymbol{w}}_a^H \underbrace{\underset{\sim}{\boldsymbol{V}}_2^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}_q}_{\underline{\boldsymbol{p}}} + \underline{\boldsymbol{w}}_a^H \underbrace{\underset{\sim}{\boldsymbol{V}}_2^H \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{V}}_2}_{\underset{\sim}{\boldsymbol{R}}} \underline{\boldsymbol{w}}_a \\
&= \sigma_d^2 - \underline{\boldsymbol{p}}^H \underline{\boldsymbol{w}}_a - \underline{\boldsymbol{w}}_a^H \underline{\boldsymbol{p}} + \underline{\boldsymbol{w}}_a^H \underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}}_a
\end{aligned}$$

$\Rightarrow$ substitution for iterative algorithm

Optimization of the cost function with respect to $\underline{\boldsymbol{w}}_a$ ($\underset{\sim}{\boldsymbol{V}}_2$), which represents the nullspace of the constraints.

$$\begin{aligned}
\frac{\partial}{\partial \underline{\boldsymbol{w}}_a^*} \left(\underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}} \right) &= \frac{\partial}{\partial \underline{\boldsymbol{w}}_a^*} \left(\underline{\boldsymbol{w}}_q^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}_q - \underline{\boldsymbol{w}}_q^H \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{w}}_a - \underline{\boldsymbol{w}}_a^H \underset{\sim}{\boldsymbol{V}}_2^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}_q + \underline{\boldsymbol{w}}_a^H \underset{\sim}{\boldsymbol{V}}_2^H \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{w}}_a \right) \\
&= - \underbrace{\underset{\sim}{\boldsymbol{V}}_2^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}_q}_{\underline{\boldsymbol{p}}} + \underbrace{\underset{\sim}{\boldsymbol{V}}_2^H \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{V}}_2}_{\underset{\sim}{\boldsymbol{R}}} \underline{\boldsymbol{w}}_a \overset{!}{=} 0
\end{aligned}$$

Is the same Problem as we already solved with the SDA ($\underset{\sim}{\boldsymbol{R}} \underline{\boldsymbol{w}}_a = \underline{\boldsymbol{p}}$).

1. Compute SVD of $\underset{\sim}{\boldsymbol{R}}$: $\underline{\boldsymbol{w}}_q = \underset{\sim}{\boldsymbol{B}}^+ \underline{\boldsymbol{c}}$; $\quad \underset{\sim}{\boldsymbol{V}}_2$; $\qquad \underset{\sim}{\boldsymbol{B}}^+ = \underset{\sim}{\boldsymbol{V}}_1 \underset{\sim}{\boldsymbol{\Sigma}}_{\boldsymbol{1}}^{\boldsymbol{-1}} \underset{\sim}{\boldsymbol{U}}_1^H$
2. $\underset{\sim}{\boldsymbol{R}} \leftarrow \underset{\sim}{\boldsymbol{V}}_2 \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{V}}_2$; $\quad \underline{\boldsymbol{p}} \leftarrow \underset{\sim}{\boldsymbol{V}}_2^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}_q$

3. Run SDA for $\underline{\boldsymbol{w}}_a$

4. $\underline{\boldsymbol{w}} \leftarrow \underline{\boldsymbol{w}}_q - \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{w}}_a$

3.4.3 Steepest Descent Procedure

Procedure runs permanently with a exponential weighted estimation of correlation matrix and vector.

$\underset{\sim}{\boldsymbol{R}} = E[\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n]], \qquad \underline{\boldsymbol{p}} = E[\underline{\boldsymbol{u}}[n]d^*[n]]$

Problem: $\underset{\sim}{\boldsymbol{R}}$ and $\underline{\boldsymbol{p}}$ are unknown and changing. Estimation and tracking is necessary.

$\rightarrow$ Averaging in <u>time</u>.

Idea: $\underset{\sim}{\hat{\boldsymbol{R}}}[n] = \frac{1}{S}\sum\limits_{k=0}^{S-1} \underline{\boldsymbol{u}}[n-k]\underline{\boldsymbol{u}}^H[n-k]$

Generalize: $\underset{\sim}{\hat{\boldsymbol{R}}}[n] = \frac{\sum\limits_{k=0}^{\infty} a_k \underline{\boldsymbol{u}}[n-k]\underline{\boldsymbol{u}}^H[n-k]}{\sum\limits_{k=0}^{\infty} a_k}$

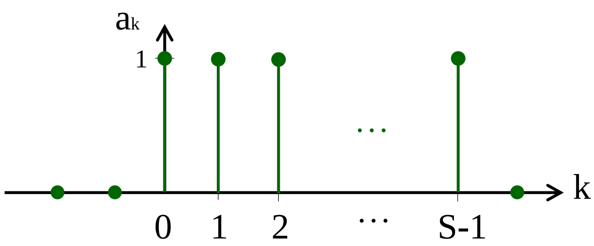

Figure 29: Square window weighting

The weighting coefficients a_k can be chosen in different ways. One may be to do a square window weighting (see figure 29), where in total S previous values of $\underline{\boldsymbol{u}}$ are taken into account and weighted equally (moving average). But usually it is better to weight the older values less than newer ones. Therefor an exponential weighting is used commonly.

Exponential weighting (Exponential Window): $a_k = \eta^k; \quad |\eta| \leq 1$

$$\underset{\sim}{\hat{\boldsymbol{R}}}[n] = \frac{\sum\limits_{k=0}^{\infty} \eta^k \underline{\boldsymbol{u}}[n-k]\underline{\boldsymbol{u}}^H[n-k]}{\underbrace{\sum\limits_{k=0}^{\infty} \eta^k}_{\frac{1}{1-\eta}}}$$

To simplify the implementation of the calculation a recursive equation should be derived:

$\underset{\sim}{\hat{\boldsymbol{R}}}[n] = (1-\eta) \sum\limits_{k=0}^{\infty} \eta^k \underline{\boldsymbol{u}}[n-k]\underline{\boldsymbol{u}}^H[n-k]$

$$
\begin{aligned}
\underset{\sim}{\hat{\boldsymbol{R}}}[n+1] &= (1-\eta)\sum_{k=0}^{\infty}\eta^k \underline{\boldsymbol{u}}[n+1-k]\underline{\boldsymbol{u}}^H[n+1-k] \\
&= (1-\eta)\sum_{k+1=0}^{\infty}\eta^{k+1}\underline{\boldsymbol{u}}[n-k]\underline{\boldsymbol{u}}^H[n-k] \\
&= (1-\eta)\eta\sum_{k=-1}^{\infty}\eta^{k}\underline{\boldsymbol{u}}[n-k]\underline{\boldsymbol{u}}^H[n-k] \\
&= \underbrace{\eta\,(1-\eta)\sum_{k=0}^{\infty}\eta^{k}\underline{\boldsymbol{u}}[n-k]\underline{\boldsymbol{u}}^H[n-k]}_{\underset{\sim}{\hat{\boldsymbol{R}}}[n]} + (1-\eta)\underbrace{\eta\eta^{-1}}_{1}\underline{\boldsymbol{u}}[n+1]\underline{\boldsymbol{u}}^H[n+1]
\end{aligned}
$$

$$
\boxed{\underset{\sim}{\hat{\boldsymbol{R}}}[n+1] = \eta\underset{\sim}{\hat{\boldsymbol{R}}}[n] + (1-\eta)\underline{\boldsymbol{u}}[n+1]\underline{\boldsymbol{u}}^H[n+1]}
$$

$$
\underline{\hat{\boldsymbol{p}}}[n+1] = \eta\underline{\hat{\boldsymbol{p}}}[n](1-\eta)\underline{\boldsymbol{u}}[n+1]d^*[n+1]
$$

SDP (Steepest Descent Procedure):

$\Rightarrow$ with optimal step size and estimation of correlation matrix and correlation vector

$\Rightarrow$ procedure never stops.

Input: $\underline{\boldsymbol{u}}[n]$, $d[n]$

Output: sequence of $\underline{\boldsymbol{w}}[n]$

1. Init: $\underset{\sim}{\hat{\boldsymbol{R}}} \leftarrow \underset{\sim}{\mathbf{0}}$, $\underline{\hat{\boldsymbol{p}}} \leftarrow \underline{\mathbf{0}}$, $\underline{\boldsymbol{w}} \leftarrow \underline{\mathbf{0}}$, $n \leftarrow 0$
2. Estimation: $\underset{\sim}{\hat{\boldsymbol{R}}} \leftarrow \eta\underset{\sim}{\hat{\boldsymbol{R}}} + (1-\eta)\underline{\boldsymbol{u}}[n+1]\underline{\boldsymbol{u}}^H[n+1]$

   $\underline{\hat{\boldsymbol{p}}} \leftarrow \eta\underline{\hat{\boldsymbol{p}}}[n](1-\eta)\underline{\boldsymbol{u}}[n+1]d^*[n+1]$
3. Weight update: $\underline{\hat{\boldsymbol{r}}} \leftarrow \underline{\hat{\boldsymbol{p}}} - \underset{\sim}{\hat{\boldsymbol{R}}}\underline{\boldsymbol{w}}$

   $\hat{\mu} \leftarrow \frac{\underline{\hat{\boldsymbol{r}}}^H\underline{\hat{\boldsymbol{r}}}}{\underline{\hat{\boldsymbol{r}}}H\underset{\sim}{\hat{\boldsymbol{R}}}\underline{\hat{\boldsymbol{r}}}}$

   $\underline{\boldsymbol{w}} \leftarrow \underline{\boldsymbol{w}}\hat{\mu}\underline{\hat{\boldsymbol{r}}}$
4. Output $\underline{\boldsymbol{w}}$ for $\underline{\boldsymbol{w}}[n+1]$
5. $n \leftarrow n+1$, go to step 2

Note: This is called procedure and not algorithm since an algorithm needs to stop (finish) after a finite time (by definition). That's not the case here.

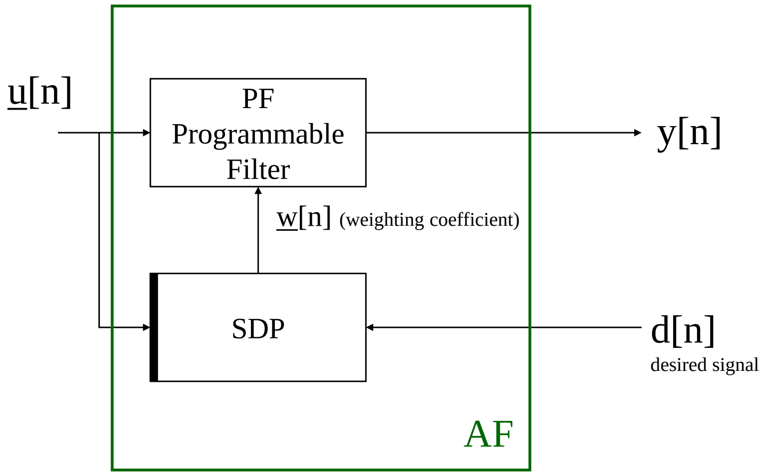

Figure 30: Block diagram of an adaptive filter (AF) using the "Steepest Descent Procedure"

Note: η close to 0:

Pro: Track fast changing channels better

Con: Poor estimation for slow changing channels

η close to 1:

Pro: Good estimation for slow changing channels

Con: Bad tracking of fast changing channels

Note: LMS-Procdure:

"LMS" $= \lim_{\eta \to 0}$ "SDP"

$\underset{\sim}{\hat{\boldsymbol{R}}}[n+1] = \underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n], \quad \hat{\underline{\mathbf{p}}}[n] = \underline{\boldsymbol{u}}[n]d^*[n]$

$\underline{\boldsymbol{w}}[n+1] = \underline{\boldsymbol{w}}[n] + \mu\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{e}}^*[n]$

$\underline{\boldsymbol{e}}[n] = d[n] - \underline{\boldsymbol{w}}^H[n]\underline{\boldsymbol{u}}[n]$

$\mu = \frac{1}{||\underline{\boldsymbol{u}}[n]||_2^2 + a}; \qquad a > 0$

or $\mu = \texttt{const}$

Application: linearly constraint minimum variance problem

$y[n] = \underline{\boldsymbol{w}}^H\underline{\boldsymbol{u}}[n]$

$\min \quad \underbrace{E[|y[n]|^2]}_{\underline{\boldsymbol{w}}^H \underbrace{E[\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n]]}_{\underset{\sim}{\boldsymbol{R}} = \underset{\sim}{\boldsymbol{A}}} \underline{\boldsymbol{w}}} \quad , \text{ s. t. } \underset{\sim}{\boldsymbol{B}}\underline{\boldsymbol{w}} = \underline{\boldsymbol{c}}$

$\underline{\boldsymbol{w}} = \underline{\boldsymbol{w}}_q - \underset{\sim}{\boldsymbol{V}}_2\underline{\boldsymbol{w}}_a$

$y[n] = \underline{\boldsymbol{w}}_q^H \underline{\boldsymbol{u}}[n] - \underline{\boldsymbol{w}}_a^H \underset{\sim}{\boldsymbol{V}}_2 \underline{\boldsymbol{u}}[n]$

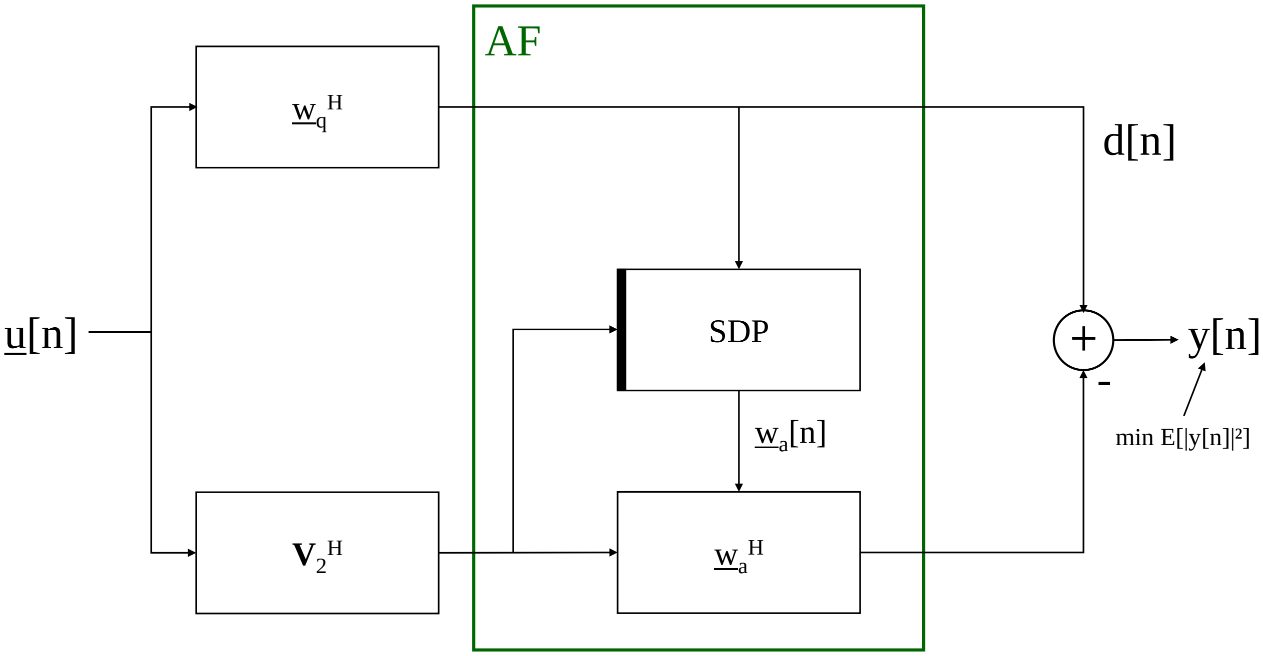

Figure 31: Linearly constraint minimum variance problem

3.4.4 Example: Digital Spectrum Analyzer (linear)

$u[n] = \sum_{k=1}^{d} s_k \cdot e^{j2\pi f_k T n} + \nu[n]$

With following variables:

d:	number of complex sinusoids
$f_1,\ f_2\ \ldots\ f_d$:	frequencies
$\lvert s_1\rvert,\ \lvert s_2\rvert\ \ldots\ \lvert s_d\rvert$:	amplitudes
$\arg S_1,\ \arg S_2,\ \ldots\ \arg S_d$:	phase
T:	sampling time
$\nu[n]$:	noise

We want to know: $d,\ f_1,\ f_2,\ \ldots\ f_d,\ \lvert s_1\rvert,\ \lvert s_2\rvert\ \ldots\ \lvert s_d\rvert$

$$\underline{\boldsymbol{u}}[n] = \begin{bmatrix} u[n] \\ u[n-1] \\ \vdots \\ u[n-(M-1)] \end{bmatrix} = \begin{bmatrix} \sum\limits_{k=1}^{d} s_k \cdot e^{j2\pi f_k T\cdot n} + \nu[n] \\ \sum\limits_{k=1}^{d} s_k \cdot e^{j2\pi f_k T\cdot (n-1)} + \nu[n-1] \\ \vdots \\ \sum\limits_{k=1}^{d} s_k \cdot e^{j2\pi f_k T\cdot (n-(M-1))} + \nu[n-(M-1)] \end{bmatrix}$$

$$= \sum_{k=1}^{d} s_k \underbrace{\begin{bmatrix} 1 \\ e^{-j2\pi f_k T} \\ \vdots \\ e^{-(M-1)j2\pi f_k T} \end{bmatrix}}_{\underline{\boldsymbol{a}}(f_k)} e^{j2\pi f_k T\cdot n} + \begin{bmatrix} \nu[n] \\ \nu[n-1] \\ \vdots \\ \nu[n-(M-1)] \end{bmatrix}$$

$$\boxed{\underline{\boldsymbol{u}}[n] = \sum_{k=1}^{d} s_k \underline{\boldsymbol{a}}(f_k) e^{j2\pi f_k T n} + \underline{\boldsymbol{\nu}}[n]}$$

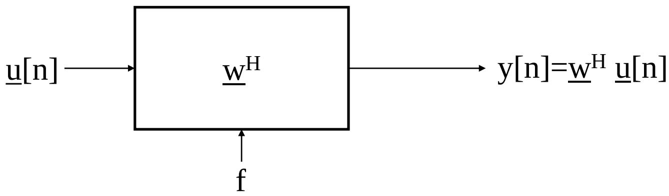

Figure 32: Block diagram of a Digital Spectrum Analyzer

1. $\underline{\boldsymbol{w}}^H \underline{\boldsymbol{a}}(f) = 1 \qquad$ or $\underbrace{\underline{\boldsymbol{a}}^H(f)}_{\underset{\sim}{\boldsymbol{B}}} \underline{\boldsymbol{w}} = 1$

2. min $\quad \underbrace{E[|y[n]|^2]}_{\underline{\boldsymbol{w}}^H \underbrace{E[\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n]]}_{\underset{\sim}{\boldsymbol{A}}} \underline{\boldsymbol{w}}}$

$$\mathcal{L} = \underline{\boldsymbol{w}}^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}} + \lambda(\underline{\boldsymbol{a}}^H(f)\underline{\boldsymbol{w}} - 1) + (\underline{\boldsymbol{w}}^H \underline{\boldsymbol{a}}(f) - 1)\lambda^*$$

$$\boxed{\underline{\boldsymbol{w}}_{opt} = \frac{\underset{\sim}{\boldsymbol{A}}^{-1}\underline{\boldsymbol{a}}(f)}{\underline{\boldsymbol{a}}^H(f)\underset{\sim}{\boldsymbol{A}}^{-1}\underline{\boldsymbol{a}}(f)}}$$

$$E[|y[n]|^2] = \underline{\boldsymbol{w}}_{opt}^H \underset{\sim}{\boldsymbol{A}} \underline{\boldsymbol{w}}_{opt} = \frac{1}{\underline{\boldsymbol{a}}^H(f)\underset{\sim}{\boldsymbol{A}}^{-1}\underline{\boldsymbol{a}}(f)}$$

$d = 2, \quad M = 3, \quad \sigma_\nu^2 = 0.001$

$S_1 = 1 + j, \quad S_2 = -1 + 2j, \quad T = 0.25\,\mu\mathrm{S}$

$|S_1|^2 = 2, \quad |S_2|^2 = 5$

$f_1 = 1\,\mathrm{MHz}, \quad f_2 = 1.7\,\mathrm{MHz}$

Parameter: $\underline{\boldsymbol{a}}(f_1) = \begin{bmatrix} 1 \\ e^{-j2\pi f_1 T} \\ e^{-j4\pi f_1 T} \end{bmatrix} \qquad \underline{\boldsymbol{a}}(f_2) = \begin{bmatrix} 1 \\ e^{-j2\pi f_2 T} \\ e^{-j4\pi f_2 T} \end{bmatrix}$

Variable: $\underline{\boldsymbol{a}}(f) = \begin{bmatrix} 1 \\ e^{-j2\pi f T} \\ e^{-j4\pi f T} \end{bmatrix}$

$$\underset{\sim}{\boldsymbol{A}} = E[\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n]]$$
$$= |S_1|^2 \underline{\boldsymbol{a}}_1 \underline{\boldsymbol{a}}_1^H + |S_2|^2 \underline{\boldsymbol{a}}_2 \underline{\boldsymbol{a}}_2^H + 0.001\underset{\sim}{\boldsymbol{I}} + S_1 S_2^* e^{j2\pi(f_1 - f_2)Tn} + S_1^* S_2 e^{-j2\pi(f_1 - f_2)Tn}$$

$$\underset{\sim}{\boldsymbol{A}} \leftarrow \underbrace{\underset{\sim}{\bar{\boldsymbol{A}}}}_{\text{time average}} = |S_1|^2 \underline{\boldsymbol{a}}_1 \underline{\boldsymbol{a}}_1^H + |S_2|^2 \underline{\boldsymbol{a}}_2 \underline{\boldsymbol{a}}_2^H + 0.001\underset{\sim}{\boldsymbol{I}}$$

$A_{11} = A_{22} = A_{33} = 7.001$

$A_{21} = A_{32} = A_{12}^* = A_{23}^* = -4.455 - j\,4.27$

$A_{31} = A_{13}^* = 0.9389 + j\,4.0451$

$$\underset{\sim}{\boldsymbol{A}} = E[\underline{\boldsymbol{u}}[n]\underline{\boldsymbol{u}}^H[n]] = E\left[\left(\sum_{k=1}^{2} s_k \underline{\boldsymbol{a}}(f_k) e^{j2\pi f_k T n} + \underline{\boldsymbol{\nu}}[n]\right)\left(\sum_{k=1}^{2} s_k \underline{\boldsymbol{a}}(f_k) e^{j2\pi f_k T n} + \underline{\boldsymbol{\nu}}[n]\right)^H\right]$$
$$= E\left[\left(s_1 \underline{\boldsymbol{a}}(f_1) e^{j2\pi f_1 T n} + s_2 \underline{\boldsymbol{a}}(f_2) e^{j2\pi f_2 T n} + \underline{\boldsymbol{\nu}}[n]\right)\left(s_1^* \underline{\boldsymbol{a}}(f_1)^H e^{-j2\pi f_1 T n} + s_2^* \underline{\boldsymbol{a}}(f_2)^H e^{-j2\pi f_2 T n} + \underline{\boldsymbol{\nu}}[n]^H\right)\right]$$
$$= |s_1|\, \underline{\boldsymbol{a}}_1 \underline{\boldsymbol{a}}_1^H + |s_2|\, \underline{\boldsymbol{a}}_2 \underline{\boldsymbol{a}}_2^H + \sigma_\nu^2 \underset{\sim}{\boldsymbol{I}} + s_1 s_2^* \underline{\boldsymbol{a}}(f_1) \underline{\boldsymbol{a}}(f_2)^H e^{j2\pi(f_1 - f_2)Tn} + s_2 s_1^* \underline{\boldsymbol{a}}(f_2) \underline{\boldsymbol{a}}(f_1)^H e^{j2\pi(f_2 - f_1)Tn}$$

$$\underset{\sim}{\bar{\boldsymbol{A}}} = \begin{bmatrix} 7.001 & -4.455 + \mathrm{j}4.27 & 0.9389 - \mathrm{j}4.0451 \\ -4.455 - \mathrm{j}4.27 & 7.001 & -4.455 + \mathrm{j}4.27 \\ 0.9389 + \mathrm{j}4.0451 & -4.455 - \mathrm{j}4.27 & 7.001 \end{bmatrix}$$

$$\underset{\sim}{\bar{\boldsymbol{A}}}^{-1} = 100 \cdot \begin{bmatrix} 2.0858 & 1.8569 - \mathrm{j}3.0297 & -0.9460 - \mathrm{j}1.8553 \\ 1.8569 + \mathrm{j}3.0297 & 6.0603 & 1.8569 - \mathrm{j}3.0297 \\ -0.9460 + \mathrm{j}1.8553 & 1.8569 + \mathrm{j}3.0297 & 2.0858 \end{bmatrix}$$

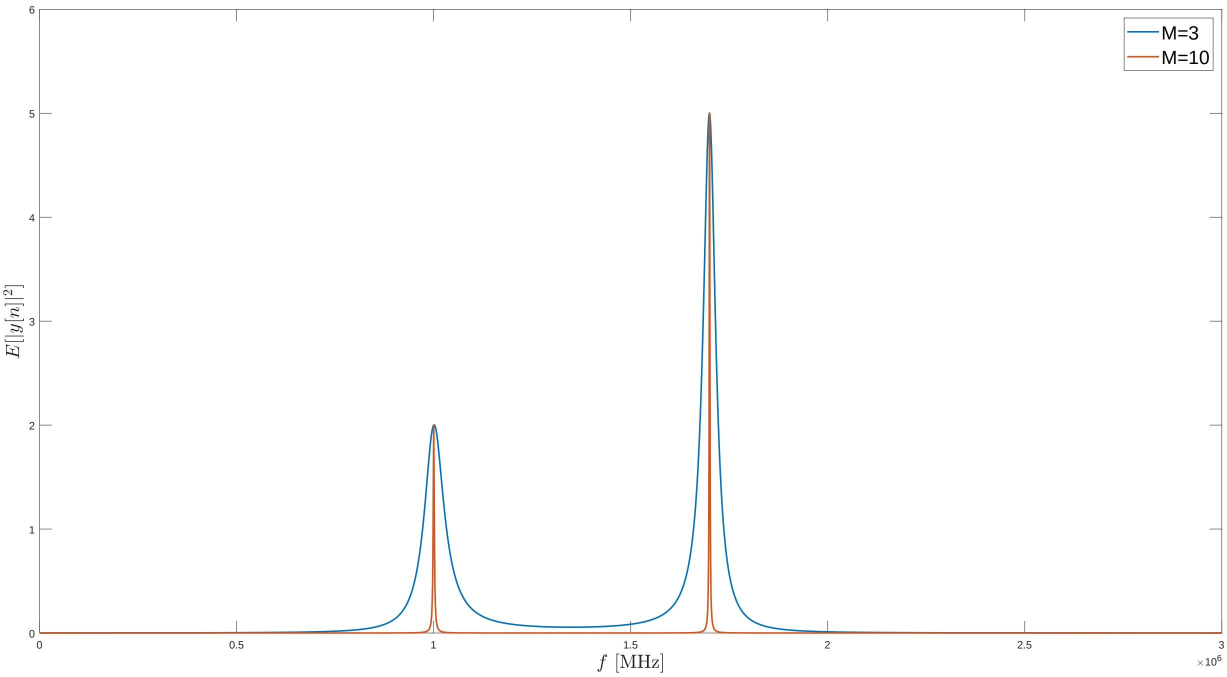

Figure 33: Plot of the result for two difference memory depth ($M = 3$ and $M = 10$)

4 Spatial Filtering

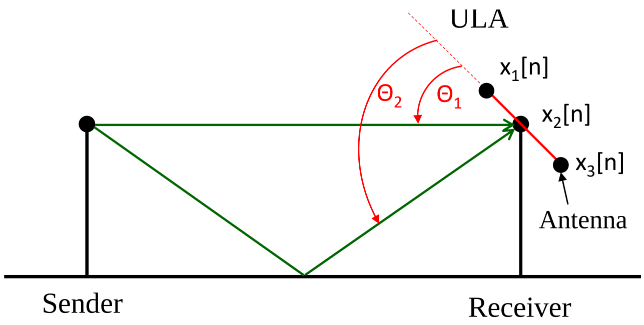

Figure 34: Eaxample of a spatial filtering scenario: Sender with 1 antenna and Receiver with 3 antennas (as ULA)

$$\underline{\boldsymbol{x}}[n] = \begin{bmatrix} x_1[n] \\ x_2[n] \\ \vdots \\ x_M[n] \end{bmatrix} = \sum_{k=1}^{d} \underline{\boldsymbol{a}}(\Theta_k) s_k[n] + \underline{\boldsymbol{\nu}}[n]$$

$$\underline{\boldsymbol{a}}(\Theta) = \begin{bmatrix} 1 \\ e^{-j2\pi\frac{\Delta}{\lambda}\cos(\Theta)} \\ e^{-j2\cdot 2\pi\frac{\Delta}{\lambda}\cos(\Theta)} \\ \vdots \\ e^{-j(M-1)2\pi\frac{\Delta}{\lambda}\cos(\Theta)} \end{bmatrix}$$

$$\underline{\boldsymbol{x}}[n] = \underbrace{\begin{bmatrix} \underline{\boldsymbol{a}}(\Theta_1) & \underline{\boldsymbol{a}}(\Theta_2) & \cdots & \underline{\boldsymbol{a}}(\Theta_d) \end{bmatrix}}_{\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{M\times d}} \cdot \underbrace{\begin{bmatrix} s_1[n] \\ s_2[n] \\ \vdots \\ s_d[n] \end{bmatrix}}_{\underline{\boldsymbol{s}}[n] \in \mathbb{C}^{M\times 1}} + \underbrace{\underline{\boldsymbol{\nu}}}_{\in \mathbb{C}^{M\times 1}}$$

$$\underline{\boldsymbol{x}}[n] = \underset{\sim}{\boldsymbol{A}} \cdot \underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n]$$

With:

$\underline{\boldsymbol{x}}[n]$: Observation Snapshot

$\underset{\sim}{\boldsymbol{A}}$: Steering Matrix

$\underline{\boldsymbol{s}}[n]$: Signal Vector

$\underline{\boldsymbol{\nu}}[n]$: Noise

Observation Matrix (Space/time):

$$\underset{\sim}{\boldsymbol{X}} := \begin{bmatrix} \underline{\boldsymbol{x}}[n] & \underline{\boldsymbol{x}}[n+1] & \cdots & \underline{\boldsymbol{x}}[n+N-1] \end{bmatrix} \in \mathbb{C}^{M\times N}$$

$$\underset{\sim}{\boldsymbol{S}} := \begin{bmatrix} \underline{\boldsymbol{s}}[n] & \underline{\boldsymbol{s}}[n+1] & \cdots & \underline{\boldsymbol{s}}[n+N-1] \end{bmatrix} \in \mathbb{C}^{d\times N}$$
$$\underset{\sim}{\boldsymbol{\nu}} := \begin{bmatrix} \underline{\boldsymbol{\nu}}[n] & \underline{\boldsymbol{\nu}}[n+1] & \cdots & \underline{\boldsymbol{\nu}}[n+N-1] \end{bmatrix} \in \mathbb{C}^{M\times N}$$

N snapshots

$$\boxed{\underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{\nu}}}$$

Given: $\underset{\sim}{\boldsymbol{X}}$ and $\underset{\sim}{\boldsymbol{A}}$

and possibly 2^{nd}-order statistics of $\underset{\sim}{\boldsymbol{S}}$ and/or $\underset{\sim}{\boldsymbol{\nu}}$

Find: Estimate $\underset{\sim}{\hat{\boldsymbol{S}}}$ of $\underset{\sim}{\boldsymbol{S}}$

Here: Linear estimations $\underset{\sim}{\hat{\boldsymbol{S}}} = \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}}$

4.1 Case 1: Only X and A are known - Least Square

Idea: $\underset{\sim}{\boldsymbol{X}} \approx \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}}$ (we assume that $\underset{\sim}{\boldsymbol{\nu}}$ is small)

$$\underset{\sim}{\hat{\boldsymbol{S}}} = \arg\min_{\underset{\sim}{\boldsymbol{S}}} ||\underset{\sim}{\boldsymbol{X}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}}||_F^2$$

$\rightarrow$ Least squares solution

$$||\underset{\sim}{\boldsymbol{B}}||_F^2 = \sum_{n,m} |B_{n,m}|^2 = \sum_{k=1}^{M} \underline{\boldsymbol{b}}_k^H \underline{\boldsymbol{b}}_k = \sum_{k=1}^{M} ||\underline{\boldsymbol{b}}_k||_2^2 = \operatorname{tr} \underset{\sim}{\boldsymbol{B}}^H \underset{\sim}{\boldsymbol{B}} = \operatorname{tr} \begin{bmatrix} \underline{\boldsymbol{b}}_1^H \\ \underline{\boldsymbol{b}}_2^H \\ \vdots \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{b}}_1 & \underline{\boldsymbol{b}}_2 & \cdots \end{bmatrix} = \underline{\boldsymbol{b}}_1^H \underline{\boldsymbol{b}}_1 + \underline{\boldsymbol{b}}_2^H \underline{\boldsymbol{b}}_2 + \cdots$$

Frobenius norm of the error:

$$\varepsilon = ||\underset{\sim}{\boldsymbol{X}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}}||_F^2 = \operatorname{tr}\left((\underset{\sim}{\boldsymbol{X}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}})^H(\underset{\sim}{\boldsymbol{X}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}})\right) = \operatorname{tr}\left((\underset{\sim}{\boldsymbol{X}}^H - \underset{\sim}{\boldsymbol{S}}^H\underset{\sim}{\boldsymbol{A}}^H)(\underset{\sim}{\boldsymbol{X}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}})\right)$$
$$= \operatorname{tr}(\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H) - \operatorname{tr}(\underset{\sim}{\boldsymbol{X}}^H\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}}) - \operatorname{tr}(\underset{\sim}{\boldsymbol{S}}^H \underbrace{\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{X}}}_{\underset{\sim}{\boldsymbol{B}}}) + \operatorname{tr}(\underset{\sim}{\boldsymbol{S}}^H\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{X}})$$

Definition: Derivation of a scalar fuction with respect to a matrix:

$$\frac{\partial\varepsilon}{\partial\underset{\sim}{\boldsymbol{S}}^*} := \begin{bmatrix} \frac{\partial\varepsilon}{\partial\underline{\boldsymbol{S}}_1^*} & \frac{\partial\varepsilon}{\partial\underline{\boldsymbol{S}}_2^*} & \cdots & \frac{\partial\varepsilon}{\partial\underline{\boldsymbol{S}}_N^*} \end{bmatrix}$$

with substitution: $\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{B}}$

$$\frac{\partial \operatorname{tr}(\underset{\sim}{\boldsymbol{S}}^H \underset{\sim}{\boldsymbol{B}})}{\partial \underline{\boldsymbol{s}}_i^*} = \frac{\partial \operatorname{tr} \begin{bmatrix} \underline{\boldsymbol{s}}_1^H \\ \vdots \\ \underline{\boldsymbol{s}}_N^H \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{b}}_1 & \cdots & \underline{\boldsymbol{b}}_N \end{bmatrix}}{\partial \underline{\boldsymbol{s}}_i^*} = \frac{\underline{\boldsymbol{s}}_1^H \underline{\boldsymbol{b}}_1 + \ldots + \underline{\boldsymbol{s}}_N^H \underline{\boldsymbol{b}}_N}{\partial \underline{\boldsymbol{s}}_i^*} = \underline{\boldsymbol{b}}_i$$

$$\frac{\partial \operatorname{tr}(\underset{\sim}{\boldsymbol{S}}^H \underset{\sim}{\boldsymbol{B}})}{\partial \underset{\sim}{\boldsymbol{S}}^*} = \begin{bmatrix} \underline{\boldsymbol{b}}_1 & \underline{\boldsymbol{b}}_2 & \cdots & \underline{\boldsymbol{b}}_N \end{bmatrix} = \underset{\sim}{\boldsymbol{B}}$$

$$\frac{\partial \varepsilon}{\partial \underset{\sim}{\boldsymbol{S}}^H} = -\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{X}} + \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{S}} \overset{!}{=} \underset{\sim}{\boldsymbol{0}}$$

$$\boxed{\underset{\sim}{\hat{\boldsymbol{S}}}_{LS} = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{A}}^+ \underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}}}$$

$\underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{M \times d}$

$\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} \in \mathbb{C}^{d \times d}$, Assume rank $\underset{\sim}{\boldsymbol{A}} = d$ $\Rightarrow$ Number of incoming wavefronts: d

rank $\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} =$ rank $\underset{\sim}{\boldsymbol{A}} = d$

$\Rightarrow$ all d steering vectors $\underline{\boldsymbol{a}}(\Theta_1) \ldots \underline{\boldsymbol{a}}(\Theta_d)$ are L.I.D.

Calculation with realistic signal (Add observation noise)

$$\underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{\nu}}$$

$$\underset{\sim}{\hat{\boldsymbol{S}}}_{\boldsymbol{LS}} = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H (\underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{\nu}})$$

$$\underset{\sim}{\hat{\boldsymbol{S}}}_{\boldsymbol{LS}} = \underbrace{(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}}}_{\underset{\sim}{\boldsymbol{I}}} \underset{\sim}{\boldsymbol{S}} + (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{\nu}}$$

$$\boxed{\begin{array}{l} \underset{\sim}{\hat{\boldsymbol{S}}}_{\boldsymbol{LS}} = \underset{\sim}{\boldsymbol{S}} + \underbrace{(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{\nu}}}_{\text{Estimation Noise}} \\ \underset{\sim}{\boldsymbol{\nu}}\text{: Observation Noise} \end{array}}$$

Note: Mean value:

$$E[\underset{\sim}{\hat{\boldsymbol{S}}}_{\boldsymbol{LS}}] = E[\underset{\sim}{\boldsymbol{S}}] + (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underbrace{E[\underset{\sim}{\boldsymbol{\nu}}]}_{=0} = E[\underset{\sim}{\boldsymbol{S}}]$$

$\Rightarrow$ <u>unbiased</u> estimate, because the arithmetic mean value of the noise is zero

4.1.1 Signal-to-Noise Ratio (SNR)

$$SNR = \frac{E[||\underset{\sim}{\hat{\boldsymbol{S}}}_{LS}||_F^2 \mid \underset{\sim}{\boldsymbol{\nu}} = 0]}{E[||\underset{\sim}{\hat{\boldsymbol{S}}}_{LS}||_F^2 \mid \underset{\sim}{\boldsymbol{S}} = 0]} = \frac{E[||\underset{\sim}{\boldsymbol{S}}||_F^2]}{E[||(\underset{\sim}{\boldsymbol{A}} \underset{\sim}{\boldsymbol{A}}^H)^{-1} \underset{\sim}{\boldsymbol{\nu}}||_F^2]} = \frac{E[\operatorname{tr} \underset{\sim}{\boldsymbol{S}}^H \underset{\sim}{\boldsymbol{S}}]}{E[\operatorname{tr} \underset{\sim}{\boldsymbol{A}}^+ \underset{\sim}{\boldsymbol{\nu}} \underset{\sim}{\boldsymbol{\nu}}^H (\underset{\sim}{\boldsymbol{A}}^+)^H]} = \frac{\operatorname{tr} E[\underset{\sim}{\boldsymbol{S}}^H \underset{\sim}{\boldsymbol{S}}]}{\operatorname{tr} E[\underset{\sim}{\boldsymbol{A}}^+ \underset{\sim}{\boldsymbol{\nu}} \underset{\sim}{\boldsymbol{\nu}}^H (\underset{\sim}{\boldsymbol{A}}^+)^H]}$$
$$= \frac{\operatorname{tr} E[\underset{\sim}{\boldsymbol{S}}^H \underset{\sim}{\boldsymbol{S}}]}{\operatorname{tr}(\underset{\sim}{\boldsymbol{A}}^+ E[\underset{\sim}{\boldsymbol{\nu}} \underset{\sim}{\boldsymbol{\nu}}^H] (\underset{\sim}{\boldsymbol{A}}^+)^H)}$$

Numinator of SNR:

$$\operatorname{tr} E[\underset{\sim}{\boldsymbol{S}}^H \underset{\sim}{\boldsymbol{S}}] = \operatorname{tr} E\left[\begin{bmatrix} \underline{\boldsymbol{s}}^H[n] \\ \underline{\boldsymbol{s}}^H[n+1] \\ \vdots \\ \underline{\boldsymbol{s}}^H[n+N-1] \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{s}}[n] & \underline{\boldsymbol{s}}[n+1] & \cdots & \underline{\boldsymbol{s}}[n+N-1] \end{bmatrix}\right]$$

$$= E\left[\operatorname{tr}(\underline{\boldsymbol{s}}^H[n]\underline{\boldsymbol{s}}[n]) + \operatorname{tr}(\underline{\boldsymbol{s}}^H[n+1]\underline{\boldsymbol{s}}[n++1]) + \ldots + \operatorname{tr}(\underline{\boldsymbol{s}}^H[n+N-1]\underline{\boldsymbol{s}}[n+N-1])\right]$$

$$= E\left[\sum_{k=0}^{N-1} \underline{\boldsymbol{s}}^H[n+k]\underline{\boldsymbol{s}}[n+1]\right]$$

$$= N \cdot E[\underline{\boldsymbol{s}}^H[n]\underline{\boldsymbol{s}}[n]]$$

$$= N \cdot E[\operatorname{tr}(\underline{\boldsymbol{s}}^H[n]\underline{\boldsymbol{s}}[n])]$$

$$= N \cdot E[\operatorname{tr}(\underline{\boldsymbol{s}}[n]\underline{\boldsymbol{s}}^H[n])]$$

$$= N \cdot \operatorname{tr}(\underbrace{E[\underline{\boldsymbol{s}}[n]\underline{\boldsymbol{s}}^H[n]]}_{\underset{\sim}{\boldsymbol{R}}_s})$$

$\Rightarrow$ Signal correlation matrix: $\underset{\sim}{\boldsymbol{R}}_s$

Interim results:

$SNR = \frac{N \operatorname{tr} \underset{\sim}{\boldsymbol{R}}_s}{\operatorname{tr}(\underset{\sim}{\boldsymbol{A}}^+ E[\underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{\nu}}^H](\underset{\sim}{\boldsymbol{A}}^+)^H)}$

Denominator of SNR:

$$E[\underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{\nu}}^H] = E\left[\begin{bmatrix} \underline{\boldsymbol{\nu}}[n] & \underline{\boldsymbol{\nu}}[n+1] & \cdots & \underline{\boldsymbol{\nu}}[n+N-1] \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{\nu}}^H[n] \\ \underline{\boldsymbol{\nu}}^H[n+1] \\ \vdots \\ \underline{\boldsymbol{\nu}}^H[n+N-1] \end{bmatrix}\right]$$

$$= E[\underline{\boldsymbol{\nu}}[n]\underline{\boldsymbol{\nu}}^H[n] + \underline{\boldsymbol{\nu}}[n+1]\underline{\boldsymbol{\nu}}^H[n+1] + \ldots + \underline{\boldsymbol{\nu}}[n+N-1]\underline{\boldsymbol{\nu}}^H[n+N-1]]$$

$$= N \cdot \underbrace{E[\underline{\boldsymbol{\nu}}[n]\underline{\boldsymbol{\nu}}^H[n]]}_{\underset{\sim}{\boldsymbol{R}}_\nu} \qquad \text{Assume: Noise is stationary}$$

$\Rightarrow$ Noise correlation matrix: $\underset{\sim}{\boldsymbol{R}}_\nu$

$$\boxed{\begin{array}{l} SNR = \frac{\operatorname{tr} \underset{\sim}{\boldsymbol{R}}_s}{\operatorname{tr}(\underset{\sim}{\boldsymbol{A}}^+ \underset{\sim}{\boldsymbol{R}}_\nu (\underset{\sim}{\boldsymbol{A}}^+)^H)} = \frac{\operatorname{tr} \underset{\sim}{\boldsymbol{R}}_s}{\operatorname{tr}((\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{A}} (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1})} \\ \text{with: } \underset{\sim}{\boldsymbol{A}}^+ = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \end{array}}$$

4.1.1.1 Special case: White signal and white noise

$$\underset{\sim}{\boldsymbol{R}}_s = \sigma_s^2 \underset{\sim}{\boldsymbol{I}}_d, \qquad \text{white signal} \quad \Rightarrow \quad \mathrm{tr}(\underset{\sim}{\boldsymbol{R}})_s = d \cdot \sigma_s^2$$

$$\underset{\sim}{\boldsymbol{R}}_\nu = \sigma_\nu^2 \underset{\sim}{\boldsymbol{I}}_M, \qquad \text{white noise}$$

$$SNR = \frac{d \cdot \sigma_s^2}{\mathrm{tr}((\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \sigma_\nu^2 \underset{\sim}{\boldsymbol{A}} (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1})} = \frac{d \cdot \frac{\sigma_s^2}{\sigma_\nu^2}}{\mathrm{tr}(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1}}$$

$$\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} \underbrace{=}_{\text{EVD}} \underset{\sim}{\boldsymbol{Q}} \underset{\sim}{\boldsymbol{\Lambda}} \underset{\sim}{\boldsymbol{Q}}^H; \qquad \underset{\sim}{\boldsymbol{Q}}^{-1} = \underset{\sim}{\boldsymbol{Q}}^H$$

$$(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} = \underset{\sim}{\boldsymbol{Q}} \underset{\sim}{\boldsymbol{\Lambda}}^{-1} \underset{\sim}{\boldsymbol{Q}}^H$$

$$\mathrm{tr}(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} = \mathrm{tr}(\underset{\sim}{\boldsymbol{Q}} \underbrace{\underset{\sim}{\boldsymbol{\Lambda}}^{-1} \underset{\sim}{\boldsymbol{Q}}^H}_{\underset{\sim}{\boldsymbol{D}}}) = \mathrm{tr}(\underset{\sim}{\boldsymbol{Q}}^H \underset{\sim}{\boldsymbol{Q}} \underset{\sim}{\boldsymbol{\Lambda}}^{-1}) = \mathrm{tr}\,\Lambda^{-1}$$

$$\boxed{SNR = \frac{d \cdot \frac{\sigma_s^2}{\sigma_\nu^2}}{\sum\limits_{k=1}^{d} \frac{1}{\lambda_k}} \qquad \text{with} \quad \lambda_k \text{ Eigenvalues of } \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}}}$$

4.1.1.2 Optimum case for LS with white noise and signal

SNR is maximum for $\sum \frac{1}{\lambda_k}$ is minimum:

$$\Rightarrow \quad \sum_{k=1}^{d} \frac{1}{\lambda_k} \text{ is minimum, s. t. } \sum_{k=1}^{d} \lambda_k = c = \texttt{const}$$

$$\sum_{k=1}^{d} \lambda_k = \mathrm{tr}\,\Lambda = \mathrm{tr}\,\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} = \sum_{k=1}^{d} ||\underline{\boldsymbol{a}}(\Theta_1)||_2^2 = M \cdot d$$

$$\mathcal{L} = \sum_{k=1}^{d} \frac{1}{\lambda_k} + \underbrace{\mu}_{\text{Lagrange-multiplier}} \left(\sum_{k=1}^{d} \lambda_k - c\right)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = -\frac{1}{\lambda_i^2} + \mu \stackrel{!}{=} 0; \qquad \lambda_i = \frac{1}{\sqrt{\mu}}$$

$$\lambda_1 = \lambda_2 = \ldots = \lambda_d = \frac{c}{d} = M$$

<u>Check if we found a minimum:</u>

$$\sum_{k=1}^{d} \frac{1}{\lambda_k} = \frac{1}{\frac{c}{d}+x} + \frac{1}{\frac{c}{d}-x} + \frac{d}{c}(d-2)$$

Taylor-Series with respect to x around 0:

$$\sum_{k=1}^{d} \frac{1}{\lambda_k} = \frac{d^2}{c} + \underbrace{\frac{2d}{c} \sum_{n=1}^{\infty} \underbrace{(\frac{d}{c}x)^{2n}}_{\geq 0}}_{\geq 0}$$

$\Rightarrow$ Minimum for $x = 0$ $\quad \Rightarrow$ maximum SNR

<u>Alternative to show that it is a minimum:</u>

$$\frac{d \frac{\sigma_s^2}{\sigma_\nu^2}}{\sum\limits_{k=1}^{d} \frac{1}{\lambda_k}} \leq d \frac{\sigma_s^2}{\sigma_\nu^2} \lambda_{min}$$

For white noise and white signal the Least-Squares-Estimation works best if and only if

$\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} = \texttt{const} \cdot \underset{\sim}{\boldsymbol{I}}$

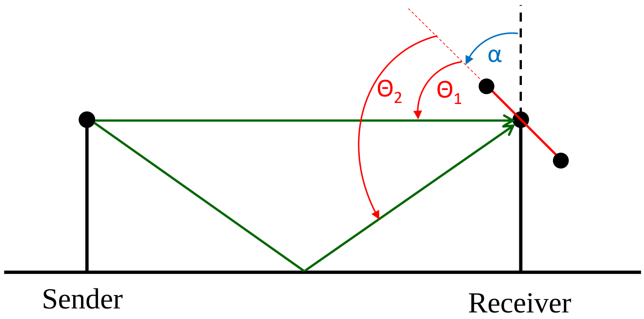

Figure 35: Example of a spatial filtering with ULA and tilt angle α

Choose α_{opt} such that $\underline{\boldsymbol{a}}^H(\Theta_1)\underline{\boldsymbol{a}}(\Theta_2) = 0$ (see figure 35).

$$\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} \underline{\boldsymbol{a}}^H(\Theta_1) \\ \underline{\boldsymbol{a}}^H(\Theta_2) \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{a}}(\Theta_1) & \underline{\boldsymbol{a}}(\Theta_2) \end{bmatrix} = \begin{bmatrix} M & 0 \\ 0 & M \end{bmatrix} = M\underset{\sim}{\boldsymbol{I}}$$

Example:

M=2 $\quad\Rightarrow\quad$ 2 antennas

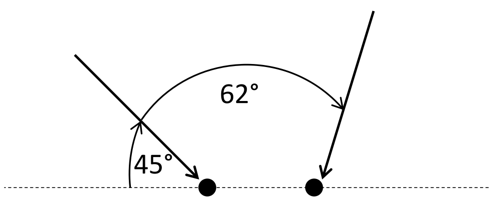

Figure 36: Example for $M = 2$ and different angles of the arriving signals

$\underline{\boldsymbol{a}}^H(\Theta_1)\underline{\boldsymbol{a}}(\Theta_2) = 0$

4.2 Case 2: Improved Least-Squares - BLUE

→ **B**est **L**inear **U**nbiased **E**stimator (BLUE)

$$\underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{\nu}}$$

$$\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} = \underbrace{\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}}}_{\underset{\sim}{\boldsymbol{I}}\text{ (unbiased)}} \underset{\sim}{\boldsymbol{S}} + \underbrace{\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}}_{\text{Estimation Noise}}$$

$$\underset{\sim}{\boldsymbol{W}}_{BLUE} = \arg\min_{\underset{\sim}{\boldsymbol{W}}} E[||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}||_F^2], \quad \text{s. t.} \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}_d$$

$$E[||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}^H||_F^2] = E[\mathrm{tr}(\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{\nu}}^H \underset{\sim}{\boldsymbol{W}})] = \mathrm{tr}\, E[\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{\nu}}^H \underset{\sim}{\boldsymbol{W}}] = \mathrm{tr}(\underset{\sim}{\boldsymbol{W}}^H \overbrace{E[\underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{\nu}}^H]}^{N \cdot \underset{\sim}{\boldsymbol{R}}_\nu} \underset{\sim}{\boldsymbol{W}})$$

$$= N \cdot \mathrm{tr}(\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{W}})$$

$$\underset{\sim}{\boldsymbol{W}} = \begin{bmatrix} \underline{\boldsymbol{w}}_1 & \underline{\boldsymbol{w}}_2 & \cdots & \underline{\boldsymbol{w}}_d \end{bmatrix}$$

$$\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}_d \Leftrightarrow \quad \underline{\boldsymbol{w}}_1^H \underset{\sim}{\boldsymbol{A}} = \underline{\boldsymbol{e}}_1^T, \quad \underline{\boldsymbol{w}}_2^H \underset{\sim}{\boldsymbol{A}} = \underline{\boldsymbol{e}}_2^T, \quad \ldots, \quad \underline{\boldsymbol{w}}_d^H \underset{\sim}{\boldsymbol{A}} = \underline{\boldsymbol{e}}_d^T$$

Lagrange function with equality constraints:

$$\mathcal{L} = \mathrm{tr}(\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{W}}) + 2\,\mathrm{Re}\left\{(\underline{\boldsymbol{e}}_1^T - \underline{\boldsymbol{w}}_1^H \underset{\sim}{\boldsymbol{A}})\underline{\boldsymbol{\lambda}}_1\right\} + \ldots + 2\,\mathrm{Re}\left\{(\underline{\boldsymbol{e}}_d^T - \underline{\boldsymbol{w}}_d^H \underset{\sim}{\boldsymbol{A}})\underline{\boldsymbol{\lambda}}_d\right\}$$

with $\begin{bmatrix} \underline{\boldsymbol{\lambda}}_1 & \underline{\boldsymbol{\lambda}}_2 & \cdots & \underline{\boldsymbol{\lambda}}_d \end{bmatrix} =: \underset{\sim}{\boldsymbol{\lambda}}$

$$\text{with} \quad \mathrm{tr}\left((\underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}})\underset{\sim}{\boldsymbol{\lambda}}\right) = \mathrm{tr} \begin{bmatrix} \underline{\boldsymbol{e}}_1^T - \underline{\boldsymbol{w}}_1^H \underset{\sim}{\boldsymbol{A}} \\ : \\ \underline{\boldsymbol{e}}_d^T - \underline{\boldsymbol{w}}_d^H \underset{\sim}{\boldsymbol{A}} \end{bmatrix} \begin{bmatrix} \underline{\boldsymbol{\lambda}}_1 & \cdots & \underline{\boldsymbol{\lambda}}_d \end{bmatrix}$$

$$= (\underline{\boldsymbol{e}}_1^T - \underline{\boldsymbol{w}}_1^H \underset{\sim}{\boldsymbol{A}})\lambda_1 + (\underline{\boldsymbol{e}}_2^T - \underline{\boldsymbol{w}}_2^H \underset{\sim}{\boldsymbol{A}})\lambda_2 + \ldots + (\underline{\boldsymbol{e}}_d^T - \underline{\boldsymbol{w}}_d^H \underset{\sim}{\boldsymbol{A}})\lambda_d$$

$$\mathcal{L} = \mathrm{tr}(\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{W}}) + 2\,\mathrm{Re}\left\{(\mathrm{tr}(\underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}})\underset{\sim}{\boldsymbol{\lambda}}\right\}$$

$$= \mathrm{tr}(\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{W}}) + \mathrm{tr}((\underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}})\underset{\sim}{\boldsymbol{\lambda}}) + \mathrm{tr}(\underset{\sim}{\boldsymbol{\lambda}}^H(\underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{W}})$$

$$\frac{\partial \mathcal{L}}{\partial \underset{\sim}{\boldsymbol{W}}^*} = \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{W}} - \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{\lambda}} \overset{!}{=} 0 \Rightarrow \underset{\sim}{\boldsymbol{W}} = \underset{\sim}{\boldsymbol{R}}_\nu^{-1} \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{\lambda}}, \quad \underset{\sim}{\boldsymbol{W}}^H = \underset{\sim}{\boldsymbol{\lambda}}^H \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_\nu^{-1}$$

$$\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}} = \underset{\sim}{\boldsymbol{\lambda}}^H \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_\nu^{-1} \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}$$

$$\underset{\sim}{\boldsymbol{\lambda}}^H = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_\nu^{-1} \underset{\sim}{\boldsymbol{A}})^{-1}$$

$$\boxed{\underset{\sim}{\boldsymbol{W}}^H_{BLUE} = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_\nu^{-1} \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_\nu^{-1}}$$

$\underset{\sim}{\boldsymbol{R}}_{\nu} = \texttt{const} \cdot \underset{\sim}{\boldsymbol{I}} \Rightarrow \underset{\sim}{\boldsymbol{W}}_{BLUE} = \underset{\sim}{\boldsymbol{W}}_{LS} \quad \Rightarrow \quad \text{white noise}$

Reconstruction of the signal:

$$\underset{\sim}{\hat{\boldsymbol{S}}} = \underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{W}}^{H}_{BLUE} \underset{\sim}{\boldsymbol{\nu}}$$

Calculate SNR:

$$E[||\underset{\sim}{\boldsymbol{W}}^{H}_{BLUE} \underset{\sim}{\boldsymbol{\nu}}||_F^2] = \operatorname{tr} E[\underset{\sim}{\boldsymbol{W}}_{BLUE} \underset{\sim}{\boldsymbol{\nu}} \underset{\sim}{\boldsymbol{\nu}}^H \underset{\sim}{\boldsymbol{W}}_{BLUE}] = N \operatorname{tr}(\underset{\sim}{\boldsymbol{W}}^{H}_{BLUE} \underset{\sim}{\boldsymbol{R}}_{\nu} \underset{\sim}{\boldsymbol{W}}_{BLUE})$$

$$SNR_{BLUE} = \frac{\operatorname{tr} \underset{\sim}{\boldsymbol{R}}_s}{\operatorname{tr}(\underset{\sim}{\boldsymbol{W}}^{H}_{BLUE} \underset{\sim}{\boldsymbol{R}}_{\nu} \underset{\sim}{\boldsymbol{W}}_{BLUE})}$$

With: $\underset{\sim}{\boldsymbol{W}}^{H}_{BLUE} \underset{\sim}{\boldsymbol{R}}_{\nu} \underset{\sim}{\boldsymbol{W}}_{BLUE} = \underbrace{\underbrace{(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{\nu}^{-1} \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{\nu}^{-1}}_{\underset{\sim}{\boldsymbol{W}}_{BLUE}} \underbrace{\underset{\sim}{\boldsymbol{R}}_{\nu} \underset{\sim}{\boldsymbol{R}}_{\nu}^{-1}}_{\underset{\sim}{\boldsymbol{I}}} \underset{\sim}{\boldsymbol{A}}}_{\underset{\sim}{\boldsymbol{I}}} (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{\nu}^{-1} \underset{\sim}{\boldsymbol{A}})$

$$\boxed{SNR_{BLUE} = \frac{\operatorname{tr} \underset{\sim}{\boldsymbol{R}}_s}{\operatorname{tr}((\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{\nu}^{-1} \underset{\sim}{\boldsymbol{A}})^{-1})}}$$

$$\boxed{SNR_{LS} = \frac{\operatorname{tr} \underset{\sim}{\boldsymbol{R}}_s}{\operatorname{tr}(\underset{\sim}{\boldsymbol{A}}^+ \underset{\sim}{\boldsymbol{R}}_{\nu} (\underset{\sim}{\boldsymbol{A}}^+)^H)}; \qquad \underset{\sim}{\boldsymbol{A}}^+ = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H}$$

BLUE and LS method are identical if:

$$\underset{\sim}{\boldsymbol{R}}_{\nu} = \texttt{const} \cdot \underset{\sim}{\boldsymbol{I}} \quad \text{or} \quad M = d \Leftrightarrow SNR_{LS} = SNR_{BLUE}$$

$\Rightarrow$ Only improvement for:

- More sensors ($d < M$)
- $\underset{\sim}{\boldsymbol{R}}_{\nu} \neq \texttt{const} \cdot \underset{\sim}{\boldsymbol{I}}$ (no white noise, but colored noise)

4.2.1 Introduction signal correlation matrix as a substitution of the noise correlation matrix

$$\underset{\sim}{\boldsymbol{W}}^{H}_{BLUE} = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{\nu}^{-1} \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{\nu}^{-1}$$

$\Rightarrow$ We need to know $\underset{\sim}{\boldsymbol{A}}$ and $\underset{\sim}{\boldsymbol{R}}_{\nu}$ or just $\underset{\sim}{\boldsymbol{A}}$ and allow *latency time*.

Why can we trade in $\underset{\sim}{\boldsymbol{R}}_{\nu}$ for latency time?

$\underset{\sim}{\boldsymbol{W}}^{H}_{BLUE} = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{x}^{-1} \underset{\sim}{\boldsymbol{A}})^{-1} \underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{R}}_{x}^{-1}$ where $\underset{\sim}{\boldsymbol{R}}_{x} = E[\underline{\boldsymbol{x}}[n] \underline{\boldsymbol{x}}^H[n]]$ is the signal correlation matrix

The orginal optimization is:

$$\underset{\sim}{\boldsymbol{W}}_{BLUE} = \arg\min E[||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}||_F^2], \quad \text{s. t. } \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}$$

Now substitude Noise $\underset{\sim}{\boldsymbol{\nu}}$ with $\underset{\sim}{\boldsymbol{X}}$:

$\underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{\nu}} \quad \Rightarrow \quad$ with optimal $\underset{\sim}{\boldsymbol{W}}^H \quad \Rightarrow \qquad \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} \mid_{\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}} = \underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}$

substitutional optimization:

$$E[\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} \underset{\sim}{\boldsymbol{X}}^H \underset{\sim}{\boldsymbol{W}}] \mid_{\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}} = E\left[(\underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}})(\underset{\sim}{\boldsymbol{S}}^H + \underset{\sim}{\boldsymbol{\nu}}^H \underset{\sim}{\boldsymbol{W}})\right]$$

$$\underbrace{=}_{E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{\nu}}]=0} E[\underset{\sim}{\boldsymbol{S}}^H \underset{\sim}{\boldsymbol{S}}] + \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{W}}$$

Note: $E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{\nu}}] = 0$ means that noise and signal are uncorrelated, which can be assumed usually.

$$\underbrace{\operatorname{tr} E[\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} \underset{\sim}{\boldsymbol{X}}^H \underset{\sim}{\boldsymbol{W}}]}_{E[||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}}||_F^2} \mid_{\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}} = \operatorname{tr}(E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^H]) + \underbrace{\operatorname{tr}(\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{R}}_\nu \underset{\sim}{\boldsymbol{W}})}_{E[||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}||_F^2}$$

If $\underset{\sim}{\boldsymbol{W}}\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}$ is met, then $\min ||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}}||_F^2$ will minimize $E[||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}|_F^2]$ because $\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{\nu}}$.

$\operatorname{tr}(E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^H]) = \texttt{const}$

So we can exchange $\underset{\sim}{\boldsymbol{R}}_\nu$ with $\underset{\sim}{\boldsymbol{R}}_x$.

Estimate $\underset{\sim}{\boldsymbol{R}}_x$ as:

$$\hat{\underset{\sim}{\boldsymbol{R}}}_x = \frac{1}{N} \sum_{k=0}^{N-1} \underline{\boldsymbol{x}}[n+k]\underline{\boldsymbol{x}}^H[n+k] = \frac{1}{N} \underset{\sim}{\boldsymbol{X}} \underset{\sim}{\boldsymbol{X}}^H$$

There is no need to know the noise, but we need to know N snapshots to fill $\underset{\sim}{\boldsymbol{X}}$.

$\Rightarrow$ therefore latency time!

$$\hat{\underset{\sim}{\boldsymbol{W}}}^H_{BLUE} = \left(\underset{\sim}{\boldsymbol{A}}^H \left(\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H\right)^{-1} \underset{\sim}{\boldsymbol{A}}\right)^{-1} \underset{\sim}{\boldsymbol{A}}^H \left(\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H\right)^{-1}$$

Note: $\hat{\underset{\sim}{\boldsymbol{W}}}^H_{BLUE}$ needs to be computed before we can use the filter. This causes *latency.*

$\hat{\underset{\sim}{\boldsymbol{S}}}_{BLUE} = \hat{\underset{\sim}{\boldsymbol{W}}}^H_{BLUE} \underset{\sim}{\boldsymbol{X}} \rightarrow$ Latency!

$\underset{\sim}{\boldsymbol{W}}^H_{LS} = (\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}})$

$\hat{\underset{\sim}{\boldsymbol{S}}}_{LS} = \underset{\sim}{\boldsymbol{W}}^H_{LS} \underset{\sim}{\boldsymbol{X}} \rightarrow$ No latency!

4.3 Case 3: LMMSE-Estimator (Linear Minimum Mean Square Error)

$\underset{\sim}{\boldsymbol{W}}_{LMMSE} = \arg\min_{\underset{\sim}{\boldsymbol{W}}} \underbrace{E[||\hat{\underset{\sim}{\boldsymbol{S}}} - \underset{\sim}{\boldsymbol{S}}||_F^2]}_{J}$

$\hat{\underset{\sim}{\boldsymbol{S}}} = \underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} \qquad$ Estimate Signal Matrix

Cost Function:

$$
\begin{aligned}
J &= E[||\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} - \underset{\sim}{\boldsymbol{S}}||_F^2] \\
&= E\left[\operatorname{tr}\left((\underset{\sim}{\boldsymbol{W}}^H \underset{\sim}{\boldsymbol{X}} - \underset{\sim}{\boldsymbol{S}})(\underset{\sim}{\boldsymbol{X}}^H \underset{\sim}{\boldsymbol{W}} - \underset{\sim}{\boldsymbol{S}})\right)\right] \\
&= \operatorname{tr}\left(\underset{\sim}{\boldsymbol{W}}^H E[\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H]\underset{\sim}{\boldsymbol{W}}\right) - \operatorname{tr}\left(\underset{\sim}{\boldsymbol{W}} E[\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{S}}^H]\right) - \operatorname{tr}\left(E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{X}}^H]\underset{\sim}{\boldsymbol{W}}\right) + \operatorname{tr}\left(E\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^H]\right)
\end{aligned}
$$

Singal Correlation Matrix:

$E[\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H] = N\underset{\sim}{\boldsymbol{R}}_x$

Channel output Correlation Matrix:

$E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^H] = N\underset{\sim}{\boldsymbol{R}}_S$

Correlation between Signal and Channel Output:

$E[\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{S}}^H] = E\left[(\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}} - \underset{\sim}{\boldsymbol{\nu}})\underset{\sim}{\boldsymbol{S}}^H\right] = \underset{\sim}{\boldsymbol{A}}E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^H] + E[\underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{S}}^H]$

With assumimg, that noise and signal are uncorrelated ($E[\underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{S}}^H] = \underset{\sim}{\boldsymbol{0}}$) we get:

$E[\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{S}}^H] = \underset{\sim}{\boldsymbol{A}}E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^H] = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s N$

With the property $\underset{\sim}{\boldsymbol{R}}_s = \underset{\sim}{\boldsymbol{R}}_s^H$ (Recall: All correlation matrices are hermitian matrices) we get for our cost function J:

$J = N\operatorname{tr}\left(\underset{\sim}{\boldsymbol{W}}^H\underset{\sim}{\boldsymbol{R}}_x\underset{\sim}{\boldsymbol{W}}\right) - N\operatorname{tr}\left(\underset{\sim}{\boldsymbol{W}}^H\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\right) - N\operatorname{tr}\left(\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{W}}\right) + N\operatorname{tr}\left(\underset{\sim}{\boldsymbol{R}}_s\right)$

Now what is our filter matrice $\underset{\sim}{\boldsymbol{W}}_{LMMSE}$?

$\frac{\partial J}{\partial \underset{\sim}{\boldsymbol{W}}^*} = N\underset{\sim}{\boldsymbol{R}}_x\underset{\sim}{\boldsymbol{W}} - N\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s \overset{!}{=} 0$

$$\boxed{\underset{\sim}{\boldsymbol{W}}_{LMMSE} = \underset{\sim}{\boldsymbol{R}}_x^{-1}\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s \qquad \text{or} \qquad \underset{\sim}{\boldsymbol{W}}_{LMMSE}^H = \underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{R}}_x^{-1}}$$

Derive Correlation Matrices

$$\underset{\sim}{\boldsymbol{R}}_x = E[\underline{\boldsymbol{x}}[n]\underline{\boldsymbol{x}}^H[n]] \underbrace{=}_{\text{stationarity}} \frac{1}{N}\sum_{k=0}^{N-1} E[x[n+k]x^h[n+k]]$$
$$= \frac{1}{N}E[\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H]$$
$$= \frac{1}{N}E\left[(\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}}+\underset{\sim}{\boldsymbol{\nu}})(\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}}^H+\underset{\sim}{\boldsymbol{\nu}}^H)\right]; \qquad \underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}}+\underset{\sim}{\boldsymbol{\nu}}$$
$$= \frac{1}{N}\left(\underset{\sim}{\boldsymbol{A}}\underbrace{E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{S}}^H]}_{N\underset{\sim}{\boldsymbol{R}}_s}+\underset{\sim}{\boldsymbol{A}}\underbrace{E[\underset{\sim}{\boldsymbol{S}}\underset{\sim}{\boldsymbol{\nu}}^H]}_{\underset{\sim}{\boldsymbol{0}}}+\underbrace{E[\underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{S}}^H]}_{\underset{\sim}{\boldsymbol{0}}}\underset{\sim}{\boldsymbol{A}}^H+\underbrace{E[\underset{\sim}{\boldsymbol{\nu}}\underset{\sim}{\boldsymbol{\nu}}^H]}_{N\underset{\sim}{\boldsymbol{R}}_\nu}\right)$$

$$\boxed{\underset{\sim}{\boldsymbol{R}}_x = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H+\underset{\sim}{\boldsymbol{R}}_\nu}$$

Assueme: $\underset{\sim}{\boldsymbol{A}}$ has got full column rank $\Rightarrow \underset{\sim}{\boldsymbol{A}}^+\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}$ therefore we can obtain $\underset{\sim}{\boldsymbol{R}}_S$

$$\boxed{\underset{\sim}{\boldsymbol{R}}_s = \underset{\sim}{\boldsymbol{A}}^+(\underset{\sim}{\boldsymbol{R}}_x-\underset{\sim}{\boldsymbol{R}}_\nu)(\underset{\sim}{\boldsymbol{A}}^+)^H \quad \Rightarrow \quad \text{it's not neccessary to know } \underset{\sim}{\boldsymbol{R}}_S}$$

$$\boxed{\begin{aligned}\underset{\sim}{\boldsymbol{W}}^H_{LMMSE} &= \underset{\sim}{\boldsymbol{A}}^+(\underset{\sim}{\boldsymbol{R}}_x-\underset{\sim}{\boldsymbol{R}}_\nu)(\underset{\sim}{\boldsymbol{A}}^+)^H\underset{\sim}{\boldsymbol{R}}_x^{-1}\\ &= (\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}})^{-1}\underset{\sim}{\boldsymbol{A}}^H(\underset{\sim}{\boldsymbol{R}}_x-\underset{\sim}{\boldsymbol{R}}_\nu)\underset{\sim}{\boldsymbol{A}}(\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}})^{-1}\underset{\sim}{\boldsymbol{R}}_x^{-1}\end{aligned}}$$

$\underset{\sim}{\hat{\boldsymbol{R}}}_x = \frac{1}{N}\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H \qquad \underset{\sim}{\boldsymbol{R}}_\nu = \underset{\sim}{\boldsymbol{R}}_x|_{\underset{\sim}{\boldsymbol{S}}=\underset{\sim}{\boldsymbol{0}}}$;

$\underset{\sim}{\boldsymbol{R}}_x^{-1}$ has to be regular (squares matrix with det $\neq 0$) $\quad\Rightarrow\quad N \geq M$

Where M is the number of sensors and N the number of snapshots we take.

$\underset{\sim}{\boldsymbol{R}}_\nu = \frac{1}{N}\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H|_{\underset{\sim}{\boldsymbol{S}}=\underset{\sim}{\boldsymbol{0}}}$

$\Rightarrow$Possibility: Run BLUE, LS and LMMSE in parallel and choose the one that performes best.

Decide on the parameter checksum / error routine which has the least error.

5 Estimate the Steering Matrix

Estimation of $\underset{\sim}{\boldsymbol{A}}$ without knowledge of $\underset{\sim}{\boldsymbol{S}}$:

$\underset{\sim}{\boldsymbol{X}} = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{S}} + \underset{\sim}{\boldsymbol{\nu}}$

To determine $\underset{\sim}{\boldsymbol{A}}$ we could use pilots like we did already earlier in this lecture. But here we want to derive a method that works without sending known symbols over the channel. This is called "Blind Channel Estimation".

Non-linear estimation for $\underset{\sim}{\boldsymbol{A}}$:

$\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} \underline{\boldsymbol{a}}[\Theta_1] & \underline{\boldsymbol{a}}[\Theta_2] & \cdots & \underline{\boldsymbol{a}}[\Theta_d] \end{bmatrix} \in \mathbb{C}^{M \times d}$

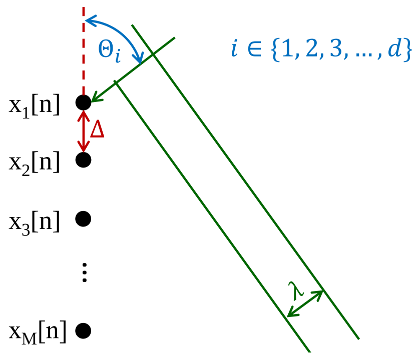

Figure 37: Unform Linear Array (ULA) with M sensors

$$\underline{\boldsymbol{a}}(\Theta_i) = \begin{bmatrix} 1 \\ e^{-\mathrm{j}2\pi\frac{\Delta}{\lambda}\cos(\Theta_i)} \\ e^{-2\mathrm{j}2\pi\frac{\Delta}{\lambda}\cos(\Theta_i)} \\ e^{-3\mathrm{j}2\pi\frac{\Delta}{\lambda}\cos(\Theta_i)} \\ \cdots \\ e^{-(M-1)\mathrm{j}2\pi\frac{\Delta}{\lambda}\cos(\Theta_i)} \end{bmatrix}$$

$\boxed{\mu := 2\pi\frac{\Delta}{\lambda}\cos(\Theta) \qquad \mu_i := 2\pi\frac{\Delta}{\lambda}\cos(\Theta_i) \qquad \mu \text{ is the phase of the signal.}}$

$$\underline{\boldsymbol{a}}(\mu_i) = \begin{bmatrix} 1 \\ e^{-\mathrm{j}\mu_i} \\ e^{-2\mathrm{j}\mu_i} \\ e^{-3\mathrm{j}\mu_i} \\ \vdots \\ e^{-(M-1)\mathrm{j}\mu_i} \end{bmatrix} = \begin{bmatrix} 1 \\ \xi^1 \\ \xi^2 \\ \xi^3 \\ \vdots \\ \xi^{M-1} \end{bmatrix} \text{"Vandermonde vector"}$$

$\Theta = \arccos\left(\frac{\mu}{2\pi\frac{\Delta}{\lambda}}\right)$

$\underline{\boldsymbol{a}}(\mu) = \underline{\boldsymbol{a}}(\mu + 2\pi n); \quad n \in \{0, \pm1, \pm2, \ldots\}$

What is now the corresponding angle? Θ or Θ'?

$\mu + 2\pi n = 2\pi\frac{\Delta}{\lambda}\cos(\Theta')$

$\mu = 2\pi\frac{\Delta}{\lambda}\cos(\Theta)$

$2\pi\frac{\Delta}{\lambda}\cos(\Theta) + 2\pi n = 2\pi\frac{\Delta}{\lambda}\cos(\Theta')$

$\cos(\Theta) + \frac{2\pi n}{2\pi\frac{\Delta}{\lambda}} = \cos(\Theta')$

$\Theta' = \arccos\underbrace{\left(\frac{n}{\frac{\Delta}{\lambda}} + \cos(\Theta)\right)}_{\Phi}; \quad n \in \{0, \pm1, \pm2, \ldots\}$

if $\frac{\Delta}{\lambda} < \frac{1}{2} \Rightarrow n = 0$ only possible for existing values $\Theta = \Theta'$

if $\frac{\Delta}{\lambda} = \frac{1}{2} \Rightarrow (\Theta, \Theta') \in \{(0,\pi), (\pi, 0)\}$

Determine the error of Θ with respect to μ:

$\Theta = \arccos\left(\frac{\mu}{2\pi\frac{\Delta}{\lambda}}\right)$

$\frac{\partial\Theta}{\partial\mu} = \frac{1}{\Delta/\lambda} \cdot \frac{1}{\sqrt{4\pi^2 - \left(\frac{\mu}{\Delta/\lambda}\right)^2}}$

$\min\limits_{\mu} \frac{\partial\Theta}{\partial\mu} = \frac{\partial\Theta}{\partial\mu}\,|_{\mu=0} = \frac{1}{2\pi\frac{\Delta}{\lambda}}$

Estimate μ with error $e_\mu = \pm\Delta\mu$

Then Θ with error $e_\Theta \approx \frac{\partial\Theta}{\partial\mu}\Delta\mu = \frac{1}{\Delta/\lambda} \cdot \frac{1}{\sqrt{4\pi^2 - \left(\frac{\mu}{\Delta/\lambda}\right)^2}}\Delta\mu$

Note that the function for $\frac{\partial\Theta}{\partial\mu}$ has a pole:

$4\pi^2 = \frac{\mu^2}{(\frac{\Delta}{\lambda})^2}; \qquad 4\pi^2(\frac{\Delta}{\lambda})^2 = \mu^2 \Rightarrow \mu = 2\pi\frac{\Delta}{\lambda}$

$\frac{\partial\Theta}{\partial\mu}\,|_{\mu=2\pi\frac{\Delta}{\lambda}} = \infty$

if $\frac{\Delta}{\lambda} = \frac{1}{2} \Rightarrow \mu = \pi$ is worst for estimating $\Theta \Rightarrow$in practice we can therefore allow $\frac{\Delta}{\lambda} = \frac{1}{2}$ because the situition is anyway very bad for estimating Θ.

$\frac{\partial\Theta}{\partial\mu} \leq 2\min\frac{\partial\Theta}{\partial\mu}$

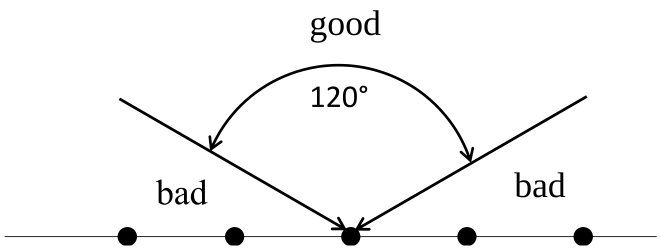

Figure 38: Behaviour of the estimated μ for different angles

d:	Number of arriving wavefronts
	unknown to the receiver
	needs to be estimated
$\Theta_1,\ \Theta_2, \dots \Theta_d$:	Directions of arrival (DoA)

$\mu_i = 2\pi\frac{\Delta}{\lambda}\cos(\Theta_i)$

From now on we choose $\frac{\Delta}{\lambda} = \frac{1}{2}$

Then:

$$\boxed{\mu_i = \pi\cos(\Theta_i)}$$

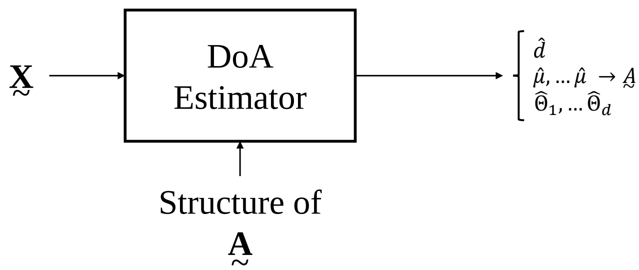

Figure 39: Simple block diagram of a DoA (Dimension of Arrival) Estimator

5.1 Space-Discrete Fourier Transform

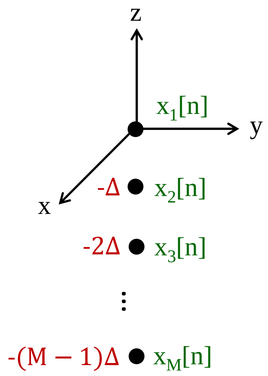

Figure 40: Sensor arrangement for the Space-Discrete Fourier Transform

- sampling in space: $x_k[n] = g(z)\mid_{z=-(k-1)\Delta} \quad \Rightarrow \quad g(z) = field$
- Space Fourier Transform: $G(\tilde{f}) = \int\limits_{-\infty}^{\infty} g(z)e^{-\mathrm{j}2\pi\tilde{f}z}dz$

The "Space Fourier Transform" is related to the Time Fourier Transform, just replace time by space (t by z) and the frequency f by the space frequency $\tilde{f}$.
$\tilde{f}$: space-frequency $\qquad [\tilde{f}] = \frac{1}{m} = m^{-1}$
⇒Problem: g(z) is not known to us everywhere. We just know it at some discrete points. Therefor we sample it:
$h(z) = g(z)\sum\limits_{k=0}^{M-1}\delta(z+\Delta k)$

Recall:
$\delta(z) = 0, \forall z \neq 0$
$\int\limits_{-\infty}^{\infty}\delta(z)dz = 1$

$$
\begin{aligned}
H(\tilde{f}) &= \int_{-\infty}^{\infty} h(z) e^{-\mathrm{j}2\pi\tilde{f}z} dz \\
&= \int_{-\infty}^{\infty} g(z) \left(\sum_{k=0}^{M-1} \delta(z + \Delta k) \right) e^{-\mathrm{j}2\pi\tilde{f}z} dz \\
&= \sum_{k=0}^{M-1} \int_{-\infty}^{\infty} g(z) \delta(z + \Delta k) e^{-\mathrm{j}2\pi\tilde{f}z} dz \\
&= \sum_{k=0}^{M-1} \int_{-\infty}^{\infty} g(-\Delta k) \delta(z + \Delta k) e^{+\mathrm{j}2\pi\tilde{f}\Delta k} dz \\
&= \sum_{k=0}^{M-1} g(-\Delta k) e^{+\mathrm{j}2\pi\tilde{f}\Delta k} \underbrace{\int_{-\infty}^{\infty} \delta(z + \Delta k) dz}_{=1} \\
&= \sum_{k=0}^{M-1} \underbrace{g(-\Delta k)}_{x_{k+1}[n]} e^{+\mathrm{j}2\pi\tilde{f}\Delta k} dz
\end{aligned}
$$

For convenience we define $\mu := 2\pi\tilde{f}\Delta$.

$$\boxed{H(\tilde{f}) = \sum_{k=0}^{M-1} x_{k+1}[n] e^{\mathrm{j}\mu k}}$$

μ: Normalized space angular frequency (dimensionless) Compare:

$\underbrace{2\pi f}_{\omega} T = \omega T = \Omega \qquad 2\pi\tilde{f}\Delta = \mu$

Recall the steering vector $\underline{\boldsymbol{a}}(\mu)$ and the received signal vector $\underline{\boldsymbol{x}}[n]$:

$$
\underline{\boldsymbol{a}}(\mu) = \begin{bmatrix} 1 \\ e^{-\mathrm{j}\mu} \\ e^{-2\mathrm{j}\mu} \\ e^{-3\mathrm{j}\mu} \\ \vdots \\ e^{-(M-1)\mathrm{j}\mu} \end{bmatrix}; \qquad \underline{\boldsymbol{x}}[n] = \begin{bmatrix} x_2[n] \\ x_3[n] \\ x_4[n] \\ \vdots \\ x_M[n] \end{bmatrix}
$$

$H(\mu) = \underline{\boldsymbol{a}}^H(\mu) \underline{\boldsymbol{x}}[n]$

Note: $H(\mu)$ is just an other notation of $H(\tilde{f})$

5.1.0.1 Fourier Periodogram

$$
\begin{aligned}
S(\mu) :&= \frac{1}{N}\sum_{n=1}^{N}\left|\underline{\boldsymbol{a}}^H(\mu)\underline{\boldsymbol{x}}[n]\right|^2 = |\bar{H}(\mu)|^2 \\
&= \frac{1}{N}\sum_{n=1}^{N}\underline{\boldsymbol{a}}^H(\mu)\underline{\boldsymbol{x}}[n]\underline{\boldsymbol{x}}^H[n]\underline{\boldsymbol{a}}(\mu) \\
&= \underline{\boldsymbol{a}}^H(\mu)\underbrace{\left(\sum_{n=1}^{N}\underline{\boldsymbol{x}}[n]\underline{\boldsymbol{x}}^H[n]\right)}_{\underset{\sim}{\hat{\boldsymbol{R}}}_x = \frac{1}{N}\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H}\underline{\boldsymbol{a}}(\mu)
\end{aligned}
$$

$$
\boxed{S(\mu) = \underline{\boldsymbol{a}}^H(\mu)\underset{\sim}{\hat{\boldsymbol{R}}}_x\underline{\boldsymbol{a}}(\mu) \qquad \mu = 2\pi\tfrac{\Delta}{\lambda}\cos(\Theta)}
$$

Search peaks in the periodogram to find μ of incoming wave fronts.

Example:

$d = 1$ (only one arriving wavefront)

no noise, continuous wave $s[n] = \texttt{const} = 1$

$\underline{\boldsymbol{x}}[n] = \underline{\boldsymbol{a}}(\mu_1)s[n] = \underline{\boldsymbol{a}}(\mu_1); \qquad \mu_1 = 2\pi\frac{\Delta}{\lambda}\cos(\Theta_1) \quad$ find: Θ_1

$\underset{\sim}{\hat{\boldsymbol{R}}}_x = \frac{1}{N}\sum\limits_{n=1}^{N}\underline{\boldsymbol{a}}(\mu_1)\underline{\boldsymbol{a}}^H(\mu_1) = \underline{\boldsymbol{a}}(\mu_1)\underline{\boldsymbol{a}}^H(\mu_1)$

$$
\begin{aligned}
S(\mu) &= \underline{\boldsymbol{a}}^H(\mu)\underset{\sim}{\hat{\boldsymbol{R}}}_x\underline{\boldsymbol{a}}(\mu) \\
&= \underline{\boldsymbol{a}}^H(\mu)\underline{\boldsymbol{a}}(\mu)\underline{\boldsymbol{a}}^H(\mu)\underline{\boldsymbol{a}}(\mu) \\
&= \left|\underline{\boldsymbol{a}}^H(\mu)\underline{\boldsymbol{a}}(\mu)\right|^2 \\
&= \left|\sum_{k=0}^{M-1} e^{\mathrm{j}(\mu-\mu_1)k}\right|^2
\end{aligned}
$$

Recall:

$s = 1 + q + q^2 + \ldots + q^{M-1}$

$sq = q + q^2 + \ldots + q^{M-1} + q^M$

$sq - s = q^M - 1 \Rightarrow s = \frac{q^M-1}{q-1}$ (geometric series)

$q = e^{\mathrm{j}(\mu-\mu_1)}$

$$
\begin{aligned}
S(\mu) &= \left| \frac{e^{\mathrm{j}(\mu-\mu_1)M} - 1}{e^{\mathrm{j}(\mu-\mu_1)} - 1} \right|^2 \\
&= \left| \frac{e^{\mathrm{j}(\mu-\mu_1)\frac{M}{2}} \cdot \left(e^{\mathrm{j}(\mu-\mu_1)\frac{M}{2}} - e^{-\mathrm{j}(\mu-\mu_1)\frac{M}{2}} \right)}{e^{\mathrm{j}\frac{(\mu-\mu_1)}{2}} \cdot \left(e^{\mathrm{j}\frac{(\mu-\mu_1)}{2}} - e^{-\mathrm{j}\frac{(\mu-\mu_1)}{2}} \right)} \right|^2 \\
&= \left(\frac{\sin\left(\frac{M}{2}(\mu-\mu_1)\right)}{\sin\left(\frac{1}{2}(\mu-\mu_1)\right)} \right)^2
\end{aligned}
$$

$$
\lim_{\mu \to \mu_1} S(\mu) = M^2
$$

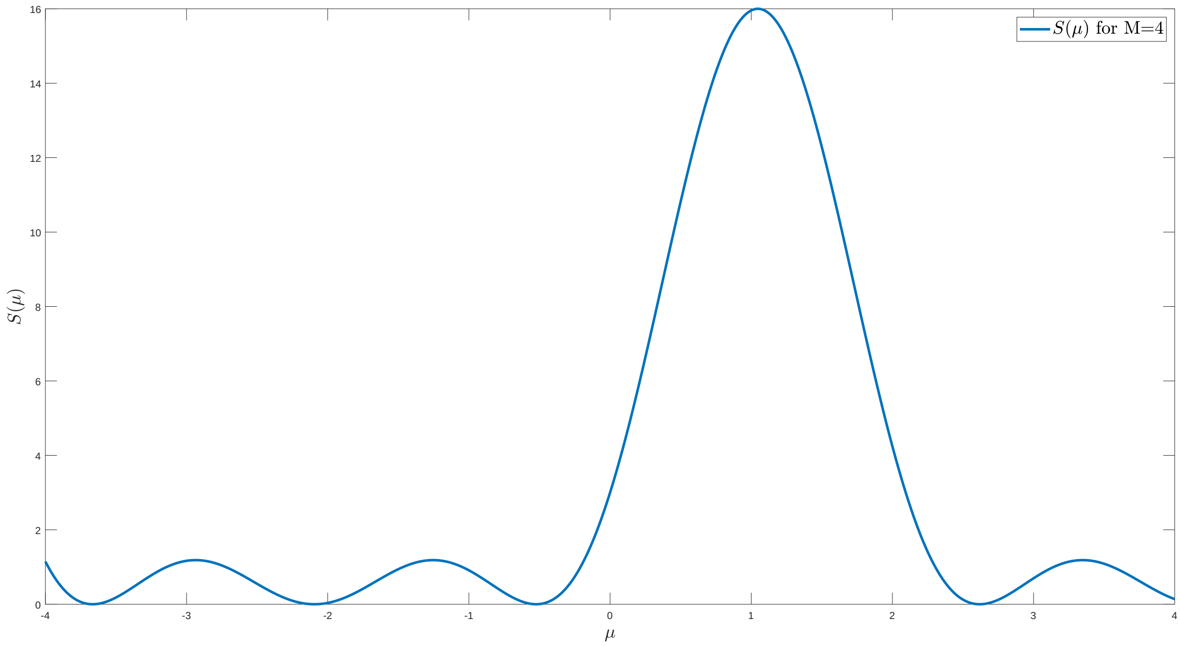

Figure 41: Fourier Periodogram for four sensors ($M = 4$)

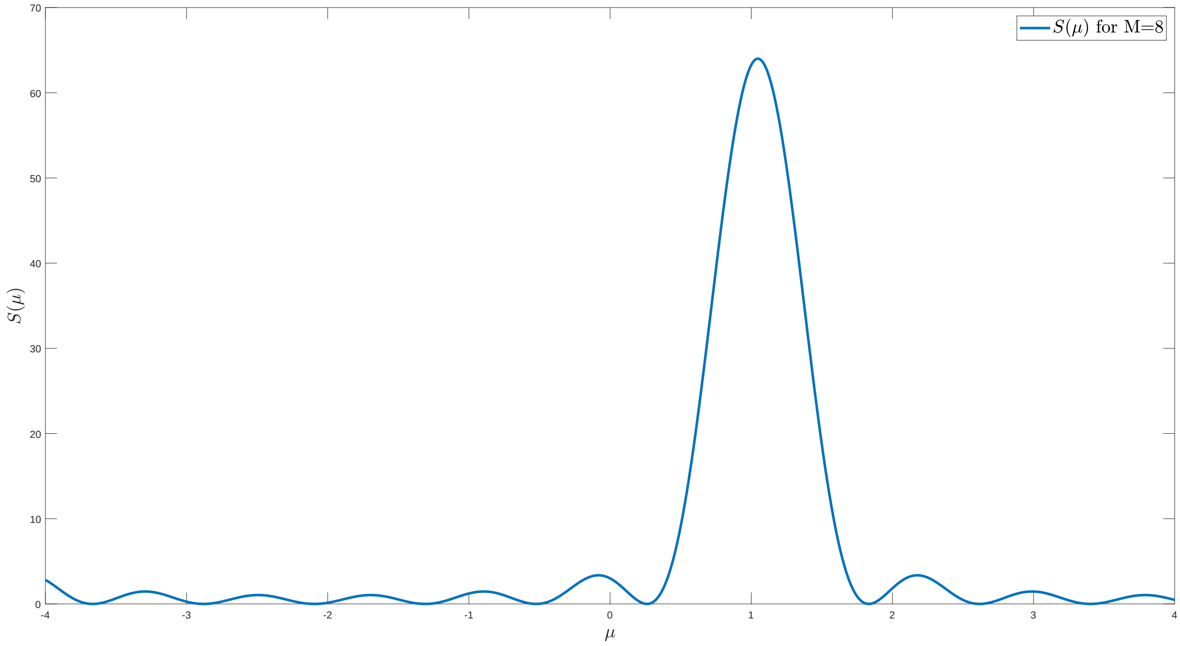

Figure 42: Fourier Periodogram for eight sensors ($M = 8$)

As can be seen from figures 41 and 42 the resolution is limited. This limit (space between maximum and next zero) is called "Rayleigh-limit" ($\Delta\mu$) and depends on the number of sensors M as follows:

$\Delta\mu = \frac{2\pi}{M}$

The only way to increase the resolution (decrease "Rayleigh-limit" $\Delta\mu$) is to increase the number of sensors. However this leads to higher costs. Therefor it is not our preferred solution. Can we do better?

$\Rightarrow$ Yes, e. g. by using the MVDR (minimum variance distortionless response) spectrum estimation:

$S_{MVDR}(\mu) = \frac{1}{\underline{\boldsymbol{a}}^H(\mu)\underset{\sim x}{\hat{\boldsymbol{R}}}^{-1}\underline{\boldsymbol{a}}(\mu)}$

$S_{FPG}(\mu) = \underline{\boldsymbol{a}}^H(\mu)\underset{\sim x}{\hat{\boldsymbol{R}}}\,\underline{\boldsymbol{a}}(\mu)$ Where FPG stands for Fourier Periodogram (for comparison)

For <u>zero noise</u> the MVDR-spectrum has <u>no</u> resolution limitation!

The MVDR is the same algorithm as the Digital spectrum analyzer (subsubsection 3.4.4) but with the spatial sampling.

Can we do better?

$\Rightarrow$ Yes, trade-off of SNR with number N of snapshots taken:

$\rightarrow$ <u>MUSIC</u>-spectrum:

Resolution depends on $N \cdot SNR$ besides on M

5.2 MUSIC-Algorithm (Multiple Signal Classification)

"Music was my first love
And it will be last."

Processing the signals received on an array of sensors for the location of the emitter is of great enough interest to have been treated under many special case assumptions. The general problem considers sensors with arbitrary locations and arbitrary directional characteristics (gain/phase/polarization) in a noise interference environment of arbitrary covariance matrix. The **mu**ltiple **si**gnal **c**lassification (MUSIC) algorithm provides asymptotically unbiased estimates of

1. number of incident wavefronts present
2. directions of arrival (DOA) (or emitter locations)
3. strengths and cross correlations among the incident waveforms
4. noise/interference strength[1]

$\underline{\boldsymbol{x}}[n] = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n]$

$$\begin{aligned}\underset{\sim}{\boldsymbol{A}} &= \begin{bmatrix}\underline{\boldsymbol{a}}(\mu_1) & \underline{\boldsymbol{a}}(\mu_2) & \cdots & \underline{\boldsymbol{a}}(\mu_d)\end{bmatrix}, \qquad \operatorname{rank}\underset{\sim}{\boldsymbol{A}} = d \quad \Rightarrow \quad \text{steering vectors are LID} \\ &= \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{U}}^H \qquad \text{(SVD)}\end{aligned}$$

$\underset{\sim}{\boldsymbol{U}} = \begin{bmatrix}\underline{\boldsymbol{u}}_1 & \underline{\boldsymbol{u}}_2 & \cdots & \underline{\boldsymbol{u}}_d & \underline{\boldsymbol{u}}_{d+1} & \underline{\boldsymbol{u}}_{d+2} & \cdots & \underline{\boldsymbol{u}}_M\end{bmatrix}$

We can partition $\underset{\sim}{\boldsymbol{U}}$ into the two matrices $\underset{\sim}{\boldsymbol{U}}_1$ and $\underset{\sim}{\boldsymbol{U}}_2$ ($\underset{\sim}{\boldsymbol{U}} = \begin{bmatrix}\underset{\sim}{\boldsymbol{U}}_1 & \underset{\sim}{\boldsymbol{U}}_2\end{bmatrix}$) with

$\underset{\sim}{\boldsymbol{U}}_1 = \begin{bmatrix}\underline{\boldsymbol{u}}_1 & \underline{\boldsymbol{u}}_2 & \cdots & \underline{\boldsymbol{u}}_d\end{bmatrix}$

and $\underset{\sim}{\boldsymbol{U}}_2 = \begin{bmatrix}\underline{\boldsymbol{u}}_{d+1} & \underline{\boldsymbol{u}}_{d+2} & \cdots & \underline{\boldsymbol{u}}_M\end{bmatrix}$

Recall the properties of matrices from the SVD:

$\underset{\sim}{\boldsymbol{U}}_1^H\underset{\sim}{\boldsymbol{U}}_1 = \underset{\sim}{\boldsymbol{I}}, \qquad \underset{\sim}{\boldsymbol{U}}_2^H\underset{\sim}{\boldsymbol{U}}_2 = \underset{\sim}{\boldsymbol{I}}, \qquad \underset{\sim}{\boldsymbol{U}}_1^H\underset{\sim}{\boldsymbol{U}}_2 = \underset{\sim}{\boldsymbol{0}}, \qquad \underset{\sim}{\boldsymbol{U}}_2^H\underset{\sim}{\boldsymbol{U}}_1 = \underset{\sim}{\boldsymbol{0}}$

$\underset{\sim}{\boldsymbol{U}}_1\underset{\sim}{\boldsymbol{U}}_1^H + \underset{\sim}{\boldsymbol{U}}_2\underset{\sim}{\boldsymbol{U}}_2^H = \underset{\sim}{\boldsymbol{I}}$

$$\operatorname{im}\underset{\sim}{\boldsymbol{A}} = \operatorname{im}\underset{\sim}{\boldsymbol{U}}_1 \quad \Rightarrow \quad \begin{cases} \forall\mu \in \{\mu_1, \mu_2, \ldots, \mu_d\} : \underline{\boldsymbol{a}}(\mu) \in \operatorname{im}\underset{\sim}{\boldsymbol{A}} \\ \forall\mu \in \{\mu_1, \mu_2, \ldots, \mu_d\} : \exists\underline{\boldsymbol{x}} : \underline{\boldsymbol{a}}(\mu) = \underset{\sim}{\boldsymbol{U}}_1\underline{\boldsymbol{x}} \end{cases}$$

Therefore

$\underset{\sim}{\boldsymbol{U}}_2^H\underline{\boldsymbol{a}}(\mu_i) = \underbrace{\underset{\sim}{\boldsymbol{U}}_2^H\underset{\sim}{\boldsymbol{U}}_1} = \underset{\sim}{\boldsymbol{0}}\underline{\boldsymbol{x}} = \underline{\boldsymbol{0}}$

5.2.1 Ideal MUSIC spectrum

$S_{\text{MUSIC, IDEAL}}(\mu) = \frac{||\underline{\boldsymbol{a}}(\mu)||_2^2}{||\underset{\sim}{\boldsymbol{U}}_2^H \underline{\boldsymbol{a}}(\mu)||_2^2}$
We assueme that $\underline{\boldsymbol{a}}(\mu) \neq 0$ holds for any μ.

To compute the actual spectrum we divide the μs into two parts.

Case 1:
Lets start with $\mu \in \{\mu_1, \ldots, \mu_d\}$: As we derived above $\underset{\sim}{\boldsymbol{U}}_2^H \underline{\boldsymbol{a}}(\mu_i) = 0$ and therefor it's Euclidean norm is also zero. So we get:
$\Rightarrow S_{\text{MUSIC, IDEAL}}(\mu_i) = \infty$

Case 2:
Assume: $\left[\underline{\boldsymbol{a}}(\mu_1) \quad \underline{\boldsymbol{a}}(\mu_2) \quad \cdots \quad \underline{\boldsymbol{a}}(\mu_d) \quad \underline{\boldsymbol{a}}(\mu)\right]$ are L. I. D. with $\mu \notin \{\mu_1, \ldots, \mu_d\}$.
$\underline{\boldsymbol{a}}(\mu) \notin \text{im}\underset{\sim}{\boldsymbol{A}} = \text{im}\underset{\sim}{\boldsymbol{U}}_1$
$\underline{\boldsymbol{a}}(\mu) \in \text{im}\underset{\sim}{\boldsymbol{U}}_2$
$\underline{\boldsymbol{a}}(\mu) = \underset{\sim}{\boldsymbol{U}}_2 \cdot \underline{\boldsymbol{y}}; \qquad \underline{\boldsymbol{y}} \in \mathbb{C}^{(M-d)\times 1}; \qquad M - d \geq 1$

$$\boxed{\begin{array}{l}\text{Out of this we get the restriction that the number of sensors must be at least one more than}\\ \text{the number of arriving wavefronts:}\\ d \leq M-1\end{array}}$$

$\underset{\sim}{\boldsymbol{U}}_2^H \underline{\boldsymbol{a}}(\mu) = \underbrace{\underset{\sim}{\boldsymbol{U}}_2^H \underset{\sim}{\boldsymbol{U}}_2}_{\underset{\sim}{\boldsymbol{I}}} \underline{\boldsymbol{y}} = \underline{\boldsymbol{y}} \neq \underline{\mathbf{0}}$
$\Rightarrow S_{\text{MUSIC, IDEAL}}(\mu) < \infty$ for $\mu \notin \{\mu_1, \ldots, \mu_d\}$

Summarizing both cases we get:
$$S_{\text{MUSIC, IDEAL}}(\mu) = \begin{cases} \infty, & \text{for} \mu \in \{\mu_1, \ldots, \mu_d\} \\ < \infty, & \text{else} \end{cases}$$

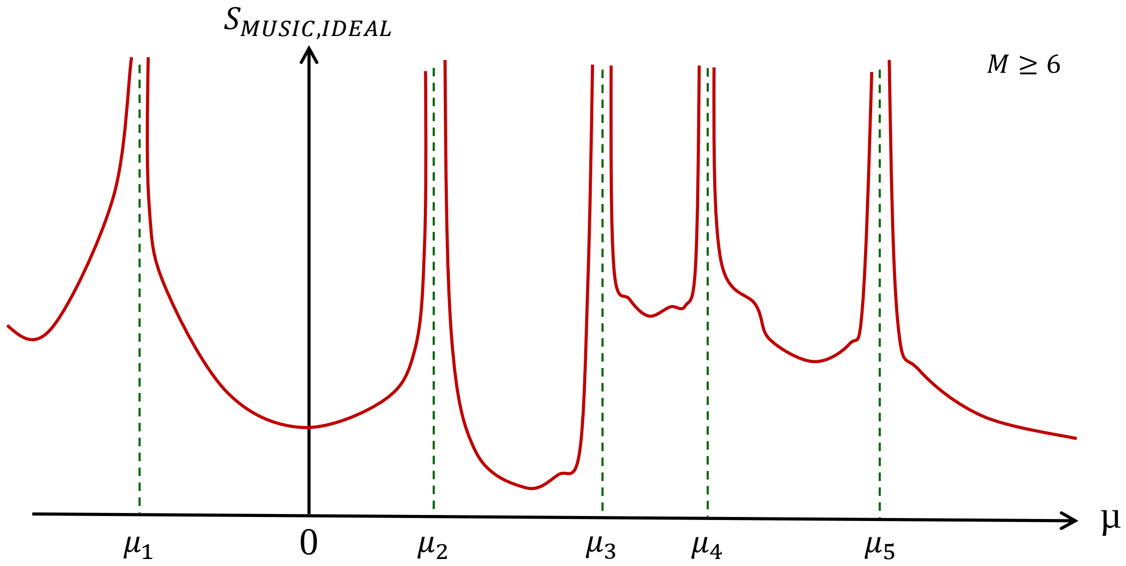

Figure 43: Ideal MUSIC spectrum ⇒no resolution limit!

As we can see in figure 43 there is no resolution limit in the ideal music spectrum. We can be as precise as necessary.

5.2.2 Problems with the MUSIC spectrum

However we got 2 problems:

1. $\begin{bmatrix} \underline{\boldsymbol{a}}(\mu_1) & \underline{\boldsymbol{a}}(\mu_2) & \cdots & \underline{\boldsymbol{a}}(\mu_d) & \underline{\boldsymbol{a}}(\mu) \end{bmatrix}$ may L. D.
2. We don't know $\underset{\sim}{\boldsymbol{U}}_2$ because $\underset{\sim}{\boldsymbol{A}}$ is unknown

5.2.2.1 Problem 1: Linear dependenc of steering vectors

Let's first look a bit deeper into problem 1:
Intuitivly it create linear dependent steering vestors ($\underline{\boldsymbol{a}}$) seems only possible if we place the sensors (antennas) in a well-aranged structure. So we consoder an uniform linear array (ULA) that is such a structure.

$$\underline{\boldsymbol{a}}(\mu_1) = \begin{bmatrix} 1 \\ a \\ a^2 \end{bmatrix}, \qquad \underline{\boldsymbol{a}}(\mu_2) = \begin{bmatrix} 1 \\ b \\ b^2 \end{bmatrix}, \underline{\boldsymbol{a}}(\mu) = \begin{bmatrix} 1 \\ c \\ c^2 \end{bmatrix}$$

where $a = e^{j\mu_1}, \quad b = e^{j\mu_2}, \quad c = e^{j\mu}$ and $a \neq b, \quad a \neq c, b \neq c$

$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \underline{\mathbf{0}}$$

In order to solve this set of linear equations we use the Gaussian elimination:

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & a-b & a-c \\ 0 & a^2-b^2 & a^2-c^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & a-b & a-c \\ 0 & 0 & (c-b)(a-c) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Since $(c-b)(a-c) \neq 0$ we get:

$\gamma = 0$

The same happens for β and α:

$\underbrace{(a-b)}_{\neq 0} \beta = 0 \Rightarrow \beta = 0$

$\alpha = 0$

This means, that even for an ULA we get L. I. D steering vectors if $\Delta < \frac{\lambda}{2}$.

It turns out that it is quiet hard to place the sensors such that the steering vectors are L. D.

So in general problem 1 does not affect us in most of the cases.

5.2.2.2 Problem 2: Determine the steering matrix A

Recall the singular value decomposition (SVD) of matrix $\underset{\sim}{\boldsymbol{A}}$: $\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H$

$\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H = \underset{\sim}{\boldsymbol{U}}\underbrace{\underset{\sim}{\boldsymbol{\Sigma}}\underset{\sim}{\boldsymbol{V}}^H\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{V}}\underset{\sim}{\boldsymbol{\Sigma}}^T}_{\underset{\sim}{\boldsymbol{\Lambda}}}\underset{\sim}{\boldsymbol{U}}^H = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{U}}^H$

Where $\underset{\sim}{\boldsymbol{\Lambda}}$ is the eigenvalue matrix.

Now let's assueme $\underset{\sim}{\boldsymbol{R}}_s$ has full rank, which means rank $\underset{\sim}{\boldsymbol{R}}_s = d$. Therefor the eigenvalue decomposition (EVD) of $\underset{\sim}{\boldsymbol{R}}_s$ exists and is given as follows:

$\underset{\sim}{\boldsymbol{R}}_s = \underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{\Lambda}}'\underset{\sim}{\boldsymbol{Q}}^H; \qquad \underset{\sim}{\boldsymbol{\Lambda}}' > \underline{\boldsymbol{0}}$ (positive definite)

$\underset{\sim}{\boldsymbol{R}}_s^{\frac{1}{2}} = \underset{\sim}{\boldsymbol{Q}}\underset{\sim}{\boldsymbol{\Lambda}}'^{\frac{1}{2}}\underset{\sim}{\boldsymbol{Q}}^H$

Since $\underset{\sim}{\boldsymbol{\Lambda}}'^{\frac{1}{2}}$ is real valued we can conclude: $\left(\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}\right)^H = \underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}$

$\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}\underset{\sim}{\boldsymbol{A}}^H = \underbrace{\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}}_{\underset{\sim}{\boldsymbol{G}}}\underbrace{\left(\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}\right)^H\underset{\sim}{\boldsymbol{A}}^H}_{\underset{\sim}{\boldsymbol{G}}^H} = \underset{\sim}{\boldsymbol{G}}\underset{\sim}{\boldsymbol{G}}^H$

$\Rightarrow \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H$ turns out to be a Gram'ian matrix (see section 2.4.6).

$$\begin{aligned} \text{im}\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H &= \text{im}\underset{\sim}{\boldsymbol{G}}\underset{\sim}{\boldsymbol{G}}^H \\ &= \text{im}\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}} \\ &= \text{im}\underset{\sim}{\boldsymbol{A}} \qquad \text{since } \underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}} \text{ is invertible} \end{aligned}$$

$\text{im}\underset{\sim}{\boldsymbol{A}} = \left\{\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{x}} \mid \underline{\boldsymbol{x}} \in \mathbb{C}^{d\times 1}\right\}$

$\text{im}\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}} = \left\{\underset{\sim}{\boldsymbol{A}}\underbrace{\underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}\underline{\boldsymbol{y}}}_{=\underline{\boldsymbol{x}}} \mid \underline{\boldsymbol{y}} \in \mathbb{C}^{d\times 1}\right\}$

Note: $\underline{\boldsymbol{x}} = \underset{\sim}{\boldsymbol{R}}^{\frac{1}{2}}\underline{\boldsymbol{y}} \Rightarrow \underline{\boldsymbol{y}} = \underset{\sim}{\boldsymbol{R}}^{-\frac{1}{2}}\underline{\boldsymbol{x}}$

rank $\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H$ = rank $\underset{\sim}{\boldsymbol{A}} = d$

$$\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{U}}^H = \begin{bmatrix}\underset{\sim}{\boldsymbol{U}}_1 & \underset{\sim}{\boldsymbol{U}}_2\end{bmatrix} \begin{bmatrix} \lambda_1 & & & & & & \\ & \lambda_2 & & & & & \\ & & \ddots & & & & \\ & & & \lambda_d & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \begin{bmatrix}\underset{\sim}{\boldsymbol{U}}_1^H \\ \underset{\sim}{\boldsymbol{U}}_2^H\end{bmatrix}$$

$\underline{\boldsymbol{x}}[n] = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n]$

If we assueme that noise and signal are uncorrelated (what usually is the case) we get the following:

$$\boxed{E\left[\underline{\boldsymbol{x}}[n]\underline{\boldsymbol{x}}^H[n]\right] = \underset{\sim}{\boldsymbol{R}}_x = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H + \underset{\sim}{\boldsymbol{R}}_\nu}$$

Note that $E\left[\underline{\boldsymbol{x}}[n]\underline{\boldsymbol{x}}^H[n]\right]$ can be estimated from array observations.

$\hat{\underset{\sim}{\boldsymbol{R}}}_x = \frac{1}{N}\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H$

To make things easier we assume the noise to be white which means $\underset{\sim}{\boldsymbol{R}}_\nu = \sigma_\nu^2\underset{\sim}{\boldsymbol{I}}$. So we get:

$$\begin{aligned}\underset{\sim}{\boldsymbol{R}}_x &= \underbrace{\underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H}_{\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Lambda}}\underset{\sim}{\boldsymbol{U}}^H} + \sigma_\nu^2 \underbrace{\underset{\sim}{\boldsymbol{I}}}_{\underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{U}}^H} \\ &= \underset{\sim}{\boldsymbol{U}}\underbrace{\left(\underset{\sim}{\boldsymbol{\Lambda}} + \sigma_\nu^2\underset{\sim}{\boldsymbol{I}}\right)}_{=\underset{\sim}{\boldsymbol{\Lambda}}'}\underset{\sim}{\boldsymbol{U}}^H\end{aligned}$$

$$\boxed{\underset{\sim}{\boldsymbol{R}}_x = \underset{\sim}{\boldsymbol{U}}(\underset{\sim}{\boldsymbol{\Lambda}} + \sigma_\nu^2\underset{\sim}{\boldsymbol{I}})\underset{\sim}{\boldsymbol{U}}^H = \underset{\sim}{\boldsymbol{U}}\underset{\sim}{\boldsymbol{\Lambda}}'\underset{\sim}{\boldsymbol{U}}^H}$$

If we sort the eigenvalues such that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d > 0$ we get the following eigenvalue matrix $\underset{\sim}{\boldsymbol{\Lambda}}'$:

$$\underset{\sim}{\boldsymbol{\Lambda}}' = \begin{bmatrix} \lambda_1 + \sigma_\nu^2 & & & & & & \\ & \lambda_2 + \sigma_\nu^2 & & & & & \\ & & \ddots & & & & \\ & & & \lambda_d + \sigma_\nu^2 & & & \\ & & & & \sigma_\nu^2 & & \\ & & & & & \ddots & \\ & & & & & & \sigma_\nu^2 \end{bmatrix}$$

$\lambda_k' = \lambda_k + \sigma_\nu^2$: Eigenvalue of $\underset{\sim}{\boldsymbol{R}}_x$

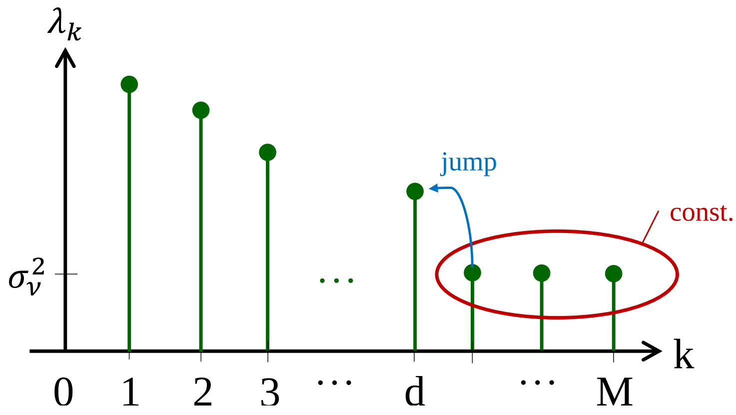

Figure 44: Ideal (all noise EVs have the same height) eigenvalue spectrum for MUSIC

From the picture, we get d and can determine $\underset{\sim}{\boldsymbol{U}}_2$:

$\underset{\sim}{\boldsymbol{U}}_2 = \begin{bmatrix} \underline{\boldsymbol{u}}_{d+1} & \underline{\boldsymbol{u}}_{d+1} & \cdots & \underline{\boldsymbol{u}}_M \end{bmatrix}$

$\underset{\sim}{\boldsymbol{R}}_x$ is still unknown but we can take $\underset{\sim}{\hat{\boldsymbol{R}}}_x$ in its place:

$\underset{\sim}{\hat{\boldsymbol{R}}}_x = \frac{1}{N} \sum_{k=0}^{N-1} \underline{\boldsymbol{x}}[n+k] x^H[n+k] = \frac{1}{N} \underset{\sim}{\boldsymbol{X}} \underset{\sim}{\boldsymbol{X}}^H$

$\underset{\sim}{\boldsymbol{X}} = \begin{bmatrix} \underline{\boldsymbol{x}}[n] & \underline{\boldsymbol{x}}[n+1] & \cdots & \underline{\boldsymbol{x}}[n+N-1] \end{bmatrix}$

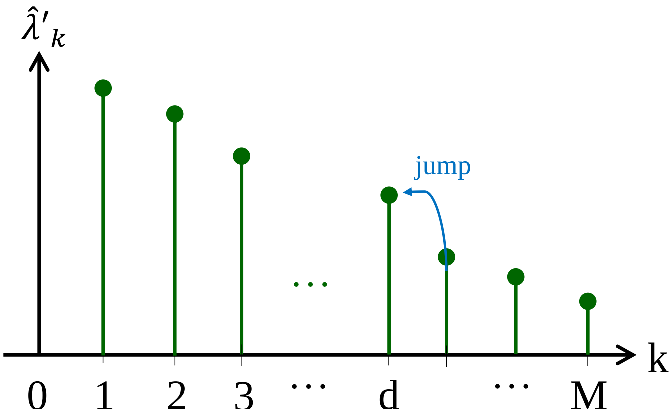

Figure 45: Real (noise EVs have NOT the same height) eigenvalue spectrum for MUSIC

5.2.2.3 Jump detection

As can be seen above a jump detection is needed in order to determine the number of arriving wavefronts d. This can be done by defining a threshold q. If the difference between the m-th eigenvalue and the (m-1)-st is larger than this threshold this is considerd as jump and the value for d is found. But now the qeustion is how to set q such that it is neither to small nor to big.

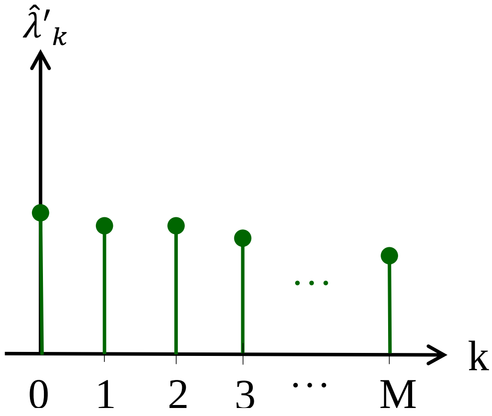

Figure 46: Noise eigenvalues to determine the minimum jump size q

Let's assume that $d = 0$ so we got pure noise. Then we define a probability p of identifying the wrong q:

$p = P_r \left[\frac{\hat{\lambda}'_m}{\hat{\lambda}'_{m-1}} > q \right]$

After that a Monte-Carlo-Experiment can be run. Out of this experiment jump sizes q can be determined (see table 4).

Sensors	M	4	8	8
Snapshots	N	100	10	1000
Probability	p	0.01	0.01	0.01
Threshold	q	1.45	1.34	1.09

Table 4: Results for minimum jump size q (Monte-Carlo-experiment)

5.2.3 Summary: MUSIC-Algorithm

Input: $\underset{\sim}{\boldsymbol{X}} \in \mathbb{C}^{M\times N}$; $\quad M =$ number of sensors, $\quad N =$ number of snapshots
Output: $\mu_1, \mu_2, \ldots, \mu_{\hat{d}}, \hat{d}$
Assume white noise!

1. $\underset{\sim}{\hat{\boldsymbol{R}}}_x = \frac{1}{N}\underset{\sim}{\boldsymbol{X}}\underset{\sim}{\boldsymbol{X}}^H$ (maximum likelihood estimation for Gauss'ian noise)
2. EVD: $\underset{\sim}{\hat{\boldsymbol{R}}}_x = \underset{\sim}{\hat{\boldsymbol{U}}}\underset{\sim}{\hat{\boldsymbol{\Lambda}}}'\underset{\sim}{\hat{\boldsymbol{U}}}^H$
3. Estimate d: "Jump detection"
4. Partition $\underset{\sim}{\hat{\boldsymbol{U}}}$:
   $\underset{\sim}{\hat{\boldsymbol{U}}} = \left[\underline{\hat{\boldsymbol{u}}}_1 \;\cdots\; \underline{\hat{\boldsymbol{u}}}_{\hat{d}} \;\; \underline{\hat{\boldsymbol{u}}}_{\hat{d}+1} \;\cdots\; \underline{\hat{\boldsymbol{u}}}_M\right]$; $\quad \underset{\sim}{\hat{\boldsymbol{U}}}_1 = \left[\underline{\hat{\boldsymbol{u}}}_1 \;\cdots\; \underline{\hat{\boldsymbol{u}}}_{\hat{d}}\right]$; $\quad \underset{\sim}{\hat{\boldsymbol{U}}}_2 = \left[\underline{\hat{\boldsymbol{u}}}_{\hat{d}+1} \;\cdots\; \underline{\hat{\boldsymbol{u}}}_M\right]$
   Remember: $\underset{\sim}{\hat{\boldsymbol{U}}}_2\underset{\sim}{\hat{\boldsymbol{U}}}_2^H = \underset{\sim}{\boldsymbol{I}} - \underset{\sim}{\hat{\boldsymbol{U}}}_1\underset{\sim}{\hat{\boldsymbol{U}}}_1^H$ may sometimes be easier
5. Search for all local maximas of $S_{\text{MUSIC}}(\mu) = \frac{\underline{\boldsymbol{a}}^H(\mu)\underline{\boldsymbol{a}}(\mu)}{\underline{\boldsymbol{a}}^H(\mu)\underset{\sim}{\boldsymbol{U}}_2\underset{\sim}{\boldsymbol{U}}_2^H\underline{\boldsymbol{a}}(\mu)}$
6. if number of peaks is $\begin{cases} < \hat{d} : \hat{d} \leftarrow \hat{d}-1 \text{ go back to 4} \\ > \hat{d} : \hat{d} \leftarrow \hat{d}+1 \text{ go back to 4} \end{cases}$
7. Return the μ_i at the position of the peaks: $\Rightarrow \underset{\sim}{\hat{\boldsymbol{A}}} = \left[\underline{\hat{\boldsymbol{a}}}(\mu_1) \;\; \underline{\hat{\boldsymbol{a}}}(\mu_2) \;\cdots\; \underline{\hat{\boldsymbol{a}}}(\mu_d)\right]$

How to continue?

After computing $\underset{\sim}{\boldsymbol{A}}$ we can proceed in the following way $\underset{\sim}{\boldsymbol{R}}_s$:
$\underset{\sim}{\boldsymbol{R}}_x = \underset{\sim}{\boldsymbol{A}}\underset{\sim}{\boldsymbol{R}}_s\underset{\sim}{\boldsymbol{A}}^H + \underset{\sim}{\boldsymbol{R}}_\nu$
For white noise: $\underset{\sim}{\boldsymbol{R}}_\nu = E[|\underline{\boldsymbol{\nu}}|_2^2]\cdot\underset{\sim}{\boldsymbol{I}}$
LS: $\underset{\sim}{\hat{\boldsymbol{R}}}_s = \underset{\sim}{\hat{\boldsymbol{A}}}^+\left(\underset{\sim}{\hat{\boldsymbol{R}}}_x - \underset{\sim}{\hat{\boldsymbol{R}}}_\nu\right)\left(\underset{\sim}{\hat{\boldsymbol{A}}}^+\right)^H$
with $\underset{\sim}{\boldsymbol{R}}_\nu = \sigma_\nu^2\underset{\sim}{\boldsymbol{I}}$ from $\underset{\sim}{\boldsymbol{\Lambda}}'$
$\Rightarrow$Now we can apply the MMSE-filter to get the signals $\underline{\boldsymbol{s}}[n]$ out of:
$\underline{\boldsymbol{x}}[n] = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{s}}[n] + \underline{\boldsymbol{\nu}}[n]$

5.2.4 Preprocessing if noise is not white (colored noise) - Noise Whitening

So far we assumed that the noise is white. However this may not allwas be the case. To be still able to compute the signal some kind of preprocessing is needed. Therefor the matrix $\underset{\sim}{\boldsymbol{T}}$ is used to "transform" the input signal such that the MUSIC algorithm works anyway, even if the noise is not white.

$\underline{\boldsymbol{x}}[n] = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{s}}[n] + \underset{\sim}{\boldsymbol{\nu}}$
$\underbrace{\underset{\sim}{\boldsymbol{T}}\underline{\boldsymbol{x}}[n]}_{=\underline{\boldsymbol{y}}[n]} = \underset{\sim}{\boldsymbol{T}}\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{s}}[n] + \underset{\sim}{\boldsymbol{T}}\underset{\sim}{\boldsymbol{\nu}}$; $\qquad \det\underset{\sim}{\boldsymbol{T}} \neq 0$ ($\underset{\sim}{\boldsymbol{T}}$ is invertible)

$\underset{\sim}{T} := \underset{\sim}{R}_\nu^{-\frac{1}{2}}$; i. e. $\underset{\sim}{T}\underset{\sim}{T} = \underset{\sim}{R}_\nu^{-1}$ (provided that $\underset{\sim}{R}_\nu^{-1}$ exists)
$E\left[\underline{y}[n]\underline{y}^H[n]\right] = \underset{\sim}{T}\underset{\sim}{A}\underset{\sim}{R}_s\underset{\sim}{A}^H\underset{\sim}{T} + \underset{\sim}{I}$
$\underline{y}[n] = \left[\underset{\sim}{T}\underline{a}(\mu_1) \quad \underset{\sim}{T}\underline{a}(\mu_2) \quad \cdots \quad \underset{\sim}{T}\underline{a}(\mu_d)\right]\underline{s}[n] + \underset{\sim}{T}\underline{\nu}[n]$
Insted of $\underline{x}$ we can feed $\underline{y}$ to the MUSIC algorithm to determine the steering matrix (preprocessing):
$\underset{\sim}{X} \leftarrow \underset{\sim}{T}\underset{\sim}{X}$
$\underline{a}(\mu) \leftarrow \underset{\sim}{T}\underline{a}(\mu)$ (run MUSIC)

⇒Remember that the BLUE filter works better as the LS filter if the noise is not white. The Result of the signal with noise whitening with the LS filter is the BLUE filter.

BLUE: $\underset{\sim}{R}_\nu^{-\frac{1}{2}}\underset{\sim}{R}_x\underset{\sim}{R}_\nu^{-\frac{1}{2}} = \underset{\sim}{R}_\nu^{-\frac{1}{2}}\underset{\sim}{A}\underset{\sim}{R}_s\underset{\sim}{A}^H\underset{\sim}{R}_\nu^{-\frac{1}{2}} + \underbrace{\underset{\sim}{R}_\nu^{-\frac{1}{2}}\underset{\sim}{R}_\nu\underset{\sim}{R}_\nu^{-\frac{1}{2}}}_{=\underset{\sim}{I}}$

$$\hat{\underset{\sim}{R}}_s = \left(\underset{\sim}{R}_\nu^{-\frac{1}{2}}\underset{\sim}{A}\right)^+\left(\hat{\underset{\sim}{R}}_\nu^{-\frac{1}{2}}\hat{\underset{\sim}{R}}_x\hat{\underset{\sim}{R}}_\nu^{-\frac{1}{2}} - \underset{\sim}{I}\right)\left(\underset{\sim}{R}_\nu^{-\frac{1}{2}}\underset{\sim}{A}\right)^{+H}$$

⇒We see that this is an other way to compute the BLUE.

5.2.5 Calibartion of general sensor positions

MUSIC is for a ULA, but in general the positions of the sensors are arbitrary. The steering vector can't be determined analytically in general. Therefore the system is measured.

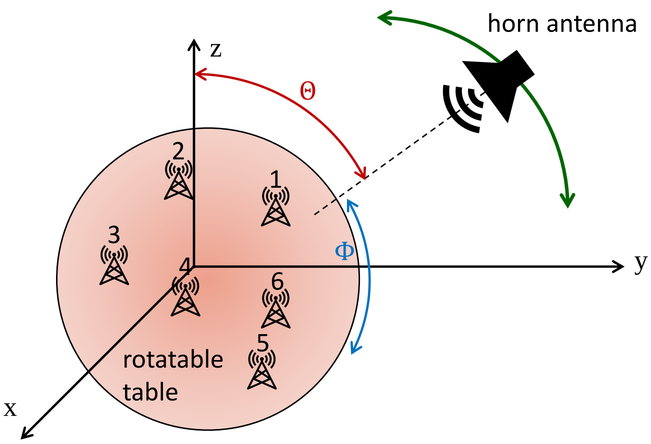

Figure 47: Calibration of an antenna constellation inside an antenna chamber

In order to get information about an antenna constellation one usually calibrates the antennas using a reference antenna inside an antenna champer. Figure 47 depicts such a calibration figurative. After the measurement an antenna table is obtained. For the example from figure 47 it could look like:
Amplitude Ratio:

$$\underline{\boldsymbol{a}}(\Theta,\,\Phi) = \begin{bmatrix} 1 \\ \frac{x_2}{x_1} \\ \frac{x_3}{x_1} \\ \cdots \\ \frac{x_6}{x_1} \end{bmatrix}$$

- get a table of the steering vectors with different $\Theta's$ and $\Phi's$
- between the values is interpolated (linear, quadratic, spline ...)

5.2.6 ULA with rotation

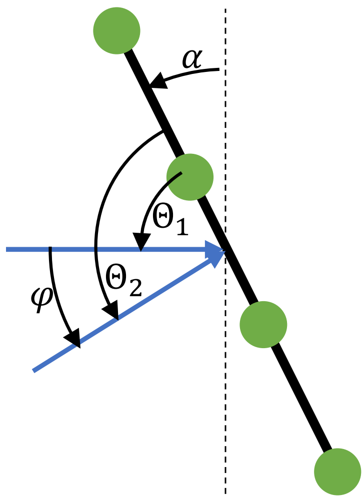

Figure 48: ULA with a rotation angle

Incoming signal with two wave fronts:

$$\underline{\boldsymbol{x}}[n] = \underline{\boldsymbol{a}}(\Theta_1)s_1[n] + \underline{\boldsymbol{a}}(\Theta_2)s_2[n] + \underline{\boldsymbol{\nu}}[n] = \underbrace{\begin{bmatrix}\underline{\boldsymbol{a}}(\Theta_1) & \underline{\boldsymbol{a}}(\Theta_2)\end{bmatrix}}_{\underset{\sim}{\boldsymbol{A}}} \underbrace{\begin{bmatrix} s_1[n] \\ s_1[n] \end{bmatrix}}_{\underline{\boldsymbol{s}}} + \underline{\boldsymbol{\nu}}[n] = \underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{s}} + \underline{\boldsymbol{\nu}}[n]$$

with $\quad \Theta_1 = \frac{\pi}{2} - \alpha; \qquad \Theta_2 = \Theta_1 + \varphi = \frac{\pi}{2} - \alpha + \varphi$

Least Square: $\underset{\sim}{\boldsymbol{W}}^H = \underset{\sim}{\boldsymbol{A}}^+$

$\hat{\underline{\boldsymbol{s}}} = \underset{\sim}{\boldsymbol{A}}^+\underline{\boldsymbol{x}}[n] = \underset{\sim}{\boldsymbol{A}}^+\underset{\sim}{\boldsymbol{A}}\underline{\boldsymbol{s}}[n] + \underset{\sim}{\boldsymbol{A}}^+\underline{\boldsymbol{\nu}}$

Because $\underline{\boldsymbol{a}}(\Theta_1), \underline{\boldsymbol{a}}(\Theta_2)$ ale LID: $\underset{\sim}{\boldsymbol{A}}^+\underset{\sim}{\boldsymbol{A}} = \underset{\sim}{\boldsymbol{I}}$

$\Rightarrow \quad \hat{\underline{\boldsymbol{s}}} = \underline{\boldsymbol{s}}[n] + \underbrace{\underset{\sim}{\boldsymbol{A}}^+\underline{\boldsymbol{\nu}}}_{\underline{\boldsymbol{\nu}}'}$

Observation Noise: $\underline{\boldsymbol{\nu}} \quad \Rightarrow \quad E[\underline{\boldsymbol{\nu}}\underline{\boldsymbol{\nu}}^H] = \sigma_\nu^2 \underset{\sim}{\boldsymbol{I}}$

Estimation Noise: $\underline{\boldsymbol{\nu}}' \quad \Rightarrow \quad E[||\underline{\boldsymbol{\nu}}||_2^2] = \sigma_\nu^2 \,\mathrm{tr}\left(\left(\underset{\sim}{\boldsymbol{A}}^H\underset{\sim}{\boldsymbol{A}}\right)^{-1}\right)$

Signal to Noise Ratio:

$$SNR = \frac{E[||\underline{s}||_2^2]}{\sigma_\nu^2 \operatorname{tr}\left(\left(\underset{\sim}{\boldsymbol{A}}^H \underset{\sim}{\boldsymbol{A}}\right)^{-1}\right)}$$

Example with 2 antennas and 2 wave fronts:

$$\underset{\sim}{\boldsymbol{A}} = \begin{bmatrix} 1 & 1 \\ e^{-j\pi\cos\Theta_1} & e^{-j\pi\cos\Theta_1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ e^{-j\pi\sin\alpha} & e^{-j\pi\sin(\alpha-\varphi)} \end{bmatrix}$$

$$SNR = \frac{E[||\underline{s}||_2^2]}{2\sigma_\nu^2} \cdot \left(1 - \underbrace{\cos\left(\pi\sin\left(\alpha-\varphi\right) - \sin\alpha\right)}_{-1...1}\right) \leq \frac{E[||\underline{s}||_2^2]}{\sigma_\nu^2}$$

Find optimal tilde angle α for the ULA:

$\alpha_{opt} = \arg\max_\alpha SNR = \frac{\varphi}{2}$

Best Result: Symmetric situation: $\quad \Theta_1 = \frac{\pi}{2} - \frac{\varphi}{2} \qquad \Theta_1 = \frac{\pi}{2} + \frac{\varphi}{2}$

$SNR_{opt} = SNR\big|_{\alpha=\frac{\varphi}{2}} = \frac{E[||\underline{s}||_2^2]}{\sigma_\nu^2} \cdot \sin^2\left(\pi\sin\frac{\varphi}{2}\right)$

Example: $\varphi = 22° \quad \Rightarrow \quad SNR_{opt} = 0.3183 \cdot SNR_{max}$

References

[1] R. Schmidt, “Multiple emitter location and signal parameter estimation”, *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986, ISSN: 0018-926X. DOI: `10.1109/TAP.1986.1143830`.