mathematics for signal processing
(signal processing for communication)

2007

spatial processing / beam forming

reader: sections 3.1-3.3

5me00

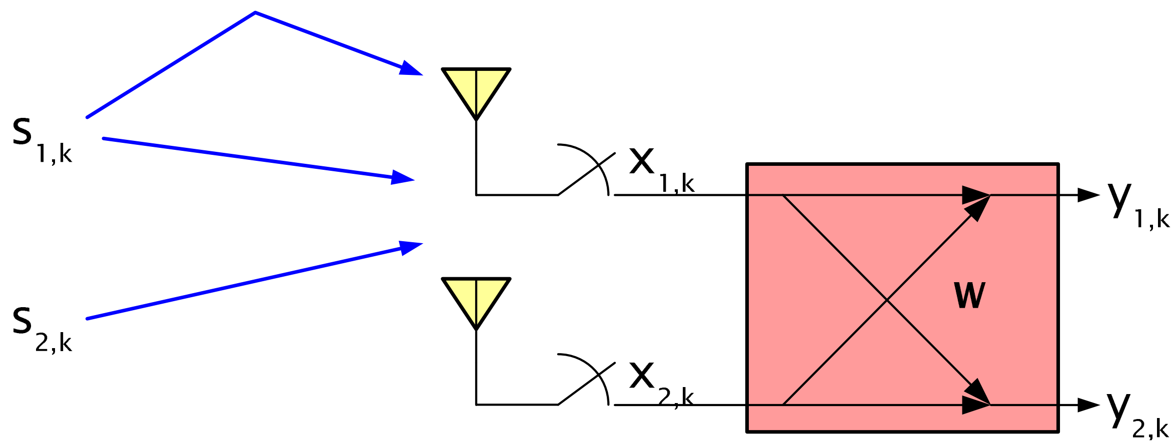
Marc Geilen
PT 9.35
m.c.w.geilen@tue.nl
http://www.es.ele.tue.nl/education/5ME00/

Eindhoven University of Technology

overview

- optimal beam formers
  - deterministic approach
  - stochastic approach
- colored noise
- the matched filter

data model



assume we receive d (narrow band) signals on an antenna array:

$$\mathbf{x}_k = \sum_{i=1}^{d} \mathbf{a}_i s_{i,k} + \mathbf{n}_k = \mathbf{A}\mathbf{s}_k + \mathbf{n}_k$$

objective: construct a receiver weight vector **w** such that

$$y_k = \mathbf{w}^H \mathbf{x}_k$$

is an estimate of one of the sources, or all sources:

$$\mathbf{y}_k = \mathbf{W}^H \mathbf{x}_k$$

deterministic approach

noiseless case $\mathbf{x}_k = \mathbf{A}\mathbf{s}_k \quad \Leftrightarrow \quad \mathbf{X} = \mathbf{A}\mathbf{S}$

MxN *Mxd* *dxN*

objective: find **W** such that $\mathbf{W}^H \mathbf{X} = \mathbf{S}$

minimize $\left\| \mathbf{W}^H \mathbf{X} - \mathbf{S} \right\|$

we consider two scenarios:

- **A** *is known*
- **S** *is known*

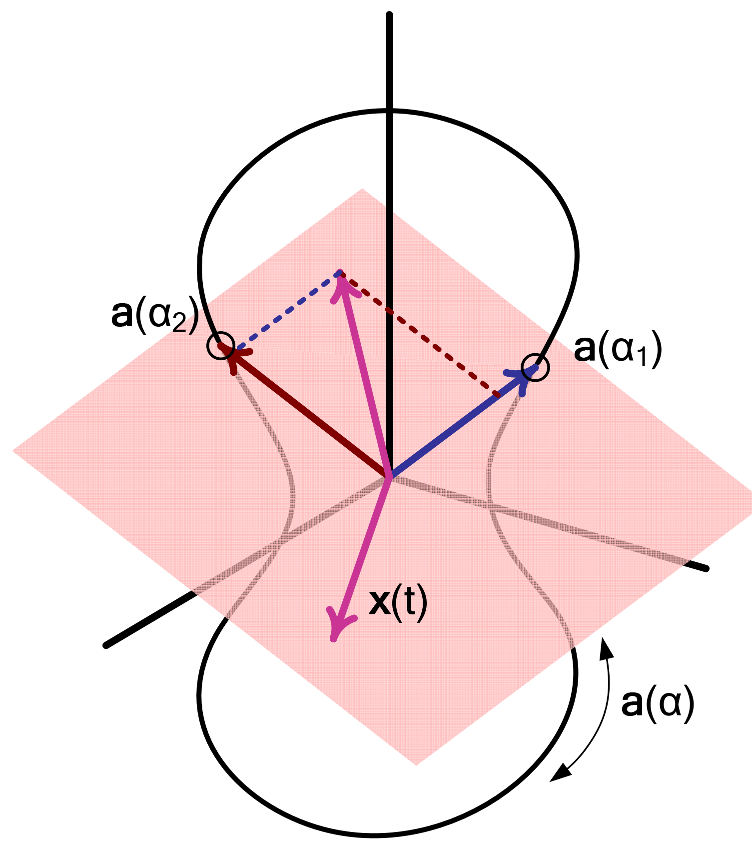
deterministic approach

scenario 1

use *direction finding methods* *to determine how many signals there are and what their response vectors are*

A is known:

$$\mathbf{A} = [\mathbf{a}_1 \ldots \mathbf{a}_d]$$



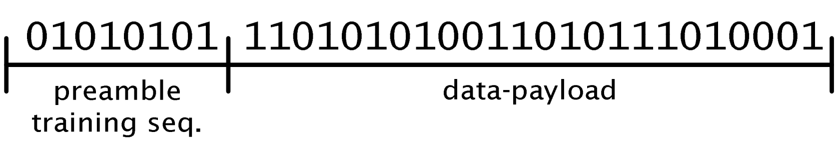
deterministic approach

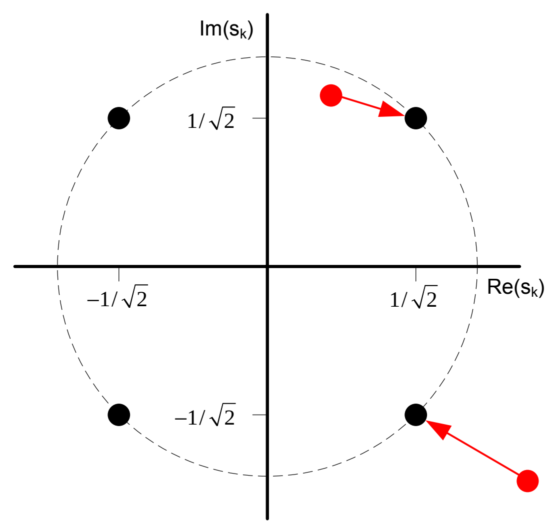
scenario 2

the sending signal makes use of a known training sequence, agreed in the protocol

S *is known*, goal: select **W**

alternatively, this could be the case via decision feedback





deterministic approach

noiseless case $\mathbf{x}_k = \mathbf{A}\mathbf{s}_k \quad \Leftrightarrow \quad \mathbf{X} = \mathbf{A}\mathbf{S}$

objective: find **W** such that $\mathbf{W}^H \mathbf{X} = \mathbf{S}$

*with **A** known*

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad \Leftrightarrow \quad \mathbf{A}^{\dagger}\mathbf{X} = \mathbf{S}, \qquad \mathbf{A}^{\dagger} = (\mathbf{A}^H \mathbf{A})^{-1}\mathbf{A}^H$$

hence, we set $\mathbf{W}^H = \mathbf{A}^{\dagger}$

all interference is cancelled (if $M \geq d$):

$$\mathbf{W}^H \mathbf{A} = \mathbf{I}_d$$

deterministic approach

noiseless case $\mathbf{x}_k = \mathbf{A}\mathbf{s}_k \quad \Leftrightarrow \quad \mathbf{X} = \mathbf{A}\mathbf{S}$

objective: find **W** such that $\mathbf{W}^H\mathbf{X} = \mathbf{S}$

with S known

$$\mathbf{W}^H = \mathbf{S}\mathbf{X}^\dagger = \mathbf{S}\mathbf{X}^H(\mathbf{X}\mathbf{X}^H)^{-1}, \qquad \mathbf{A} = \mathbf{X}\mathbf{S}^\dagger = (\mathbf{W}^H)^\dagger$$

(after training, **W** *is used to estimate the unknown* **S***)*

again, all interference is cancelled ($M \geq d$, $N \geq d$):

$$\mathbf{W}^H\mathbf{A} = \mathbf{I}$$

deterministic approach

noisy case $\mathbf{X} = \mathbf{AS} + \mathbf{N}$

two possible optimization criteria

model matching: adapting the model of **A** or **S** to minimize residual:

$$\left\| \mathbf{X} - \mathbf{AS} \right\|_F^2$$

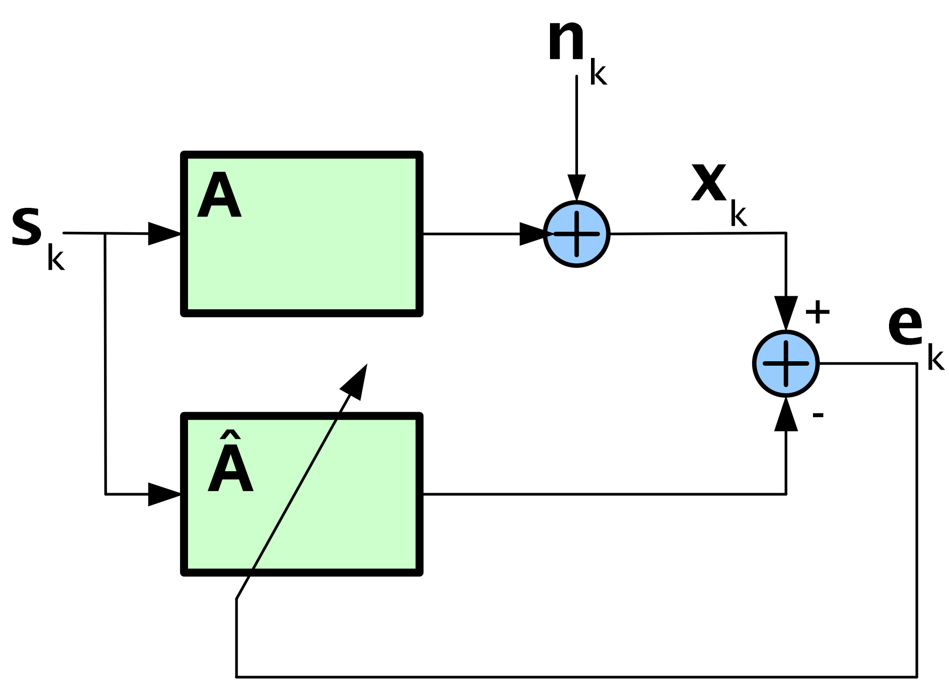
output error minimization

$$\left\| \mathbf{W}^H \mathbf{X} - \mathbf{S} \right\|_F^2$$

deterministic approach

noisy case $\mathbf{X} = \mathbf{AS} + \mathbf{N}$

model matching: adapting the model of **A** (or **S)** to minimize residual: $\|\mathbf{X} - \mathbf{AS}\|_F^2$

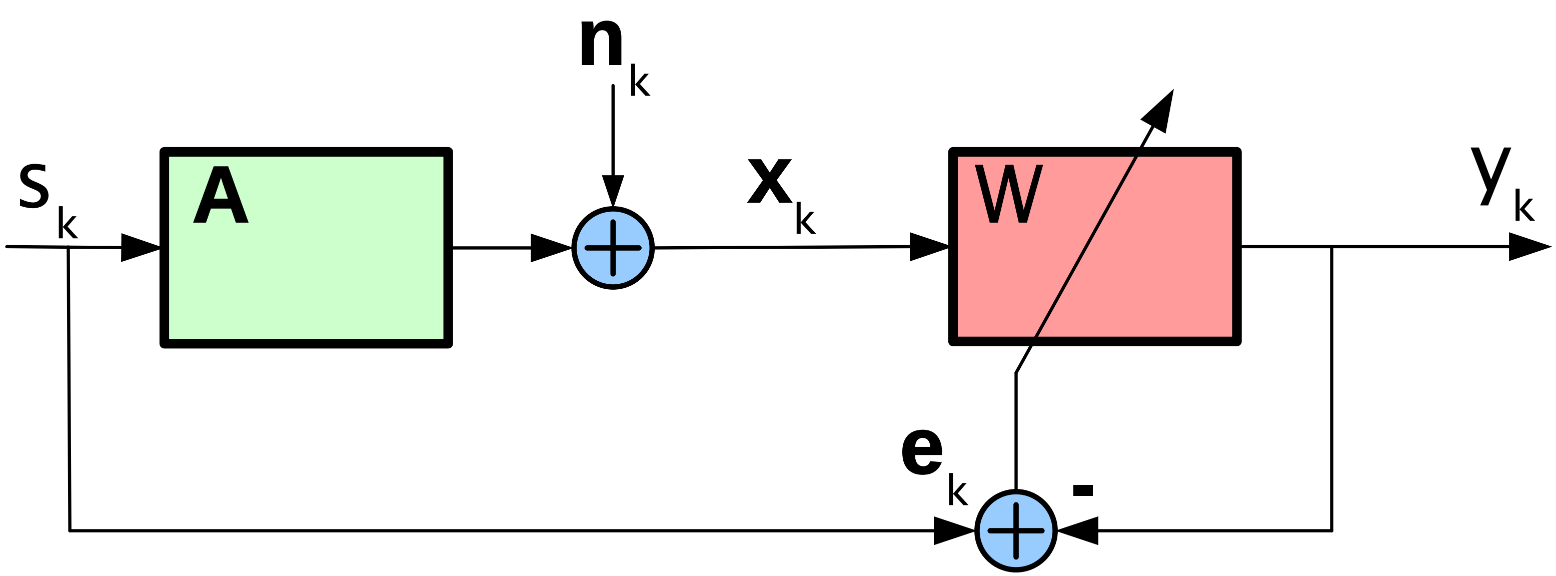


deterministic approach

noisy case $\mathbf{X} = \mathbf{AS} + \mathbf{N}$

output error minimization

$$\left\| \mathbf{W}^H \mathbf{X} - \mathbf{S} \right\|_F^2$$



model matching

with **A** *known*: *zero-forcing* solution

$$\hat{\mathbf{S}} \text{ such that } \left\|\mathbf{X} - \mathbf{A}\hat{\mathbf{S}}\right\|_F^2 \text{ is minimal}$$

$$\hat{\mathbf{S}} = \mathbf{A}^{\dagger}\mathbf{X} \quad \Rightarrow \quad \mathbf{W}^H = \mathbf{A}^{\dagger}$$

$\mathbf{W}^H\mathbf{A}=\mathbf{I}$: all interference is cancelled (hence zero-forcing)
(under what conditions?)

the ZF beamformer maximizes the output Signal-to-Interference Ratio (SIR)

example model matching

the ZF beamformer satisfies $\mathbf{W}^H\mathbf{A}=\mathbf{I}$

let $\mathbf{w}_1$ be the first column of $\mathbf{W}$, the beamformer of the first signal.

$$\mathbf{W}^H\mathbf{A}=\mathbf{I} \quad \Rightarrow \quad \mathbf{w}_1^H[\mathbf{a}_2 \ldots \mathbf{a}_d]=[\mathbf{0} \ldots \mathbf{0}]$$

$$\mathbf{w}_1 \perp \{\mathbf{a}_2,\ldots,\mathbf{a}_d\}$$

thus, $\mathbf{w}_1$ projects out all other sources, except source 1.
but what about noise?

$$y_1(t)=\mathbf{w}_1^H\mathbf{x}(t)=\sum_{i=1}^{d}\mathbf{w}_1^H\mathbf{a}_i s_i(t)+\mathbf{w}_1^H\mathbf{n}(t)$$

$$=s_1(t)+\mathbf{w}_1^H\mathbf{n}(t)$$

the effect on the noise is not considered!

model matching

zero-forcing solution

$$\mathbf{W}^H\mathbf{X} = \mathbf{S} + (\mathbf{A}^\dagger)^H\mathbf{N}$$

the output noise depends on $\mathbf{A}^+$ and can be large, since

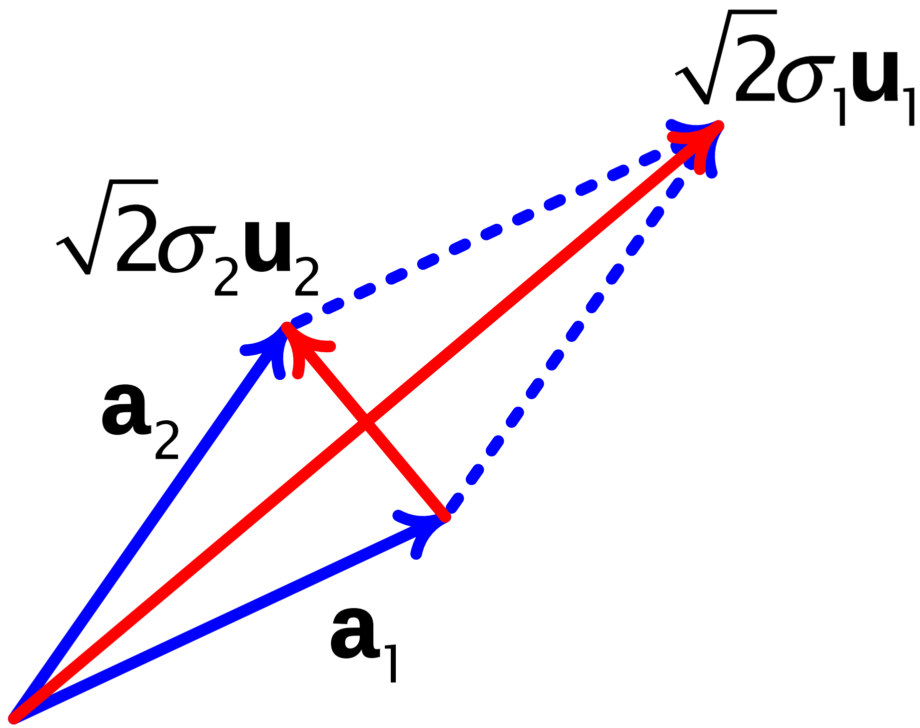
$$\mathbf{A} = \mathbf{U_A}\Sigma_\mathbf{A}\mathbf{V_A} \quad \Rightarrow \quad (\mathbf{A}^\dagger)^H = \mathbf{U_A}\Sigma_\mathbf{A}^{-1}\mathbf{V_A}$$

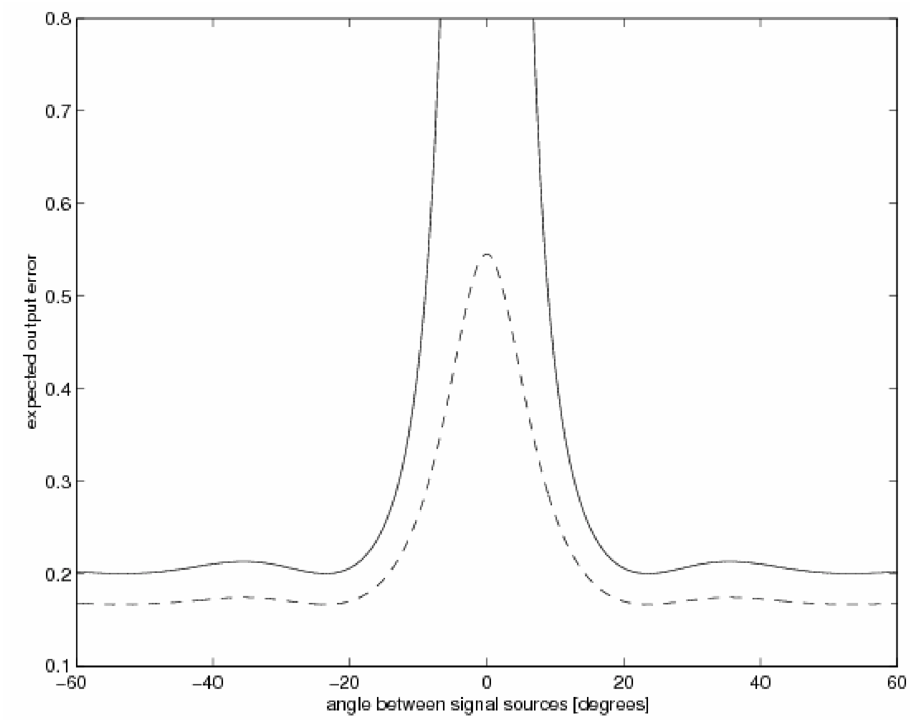
this happens if $\mathbf{\Sigma_A}^{-1}$ is large, i.e. if **A** is *ill-conditioned*

$$\frac{\sigma_1}{\sigma_d} \gg 1$$

e.g. if directions are very close

model matching





to discriminate $\mathbf{a}_1$ *and* $\mathbf{a}_2$*, the ZF beamformer amplifies noise in the direction of* $\mathbf{u}_2$

model matching

with **S** *known*:

$$\hat{\mathbf{A}} \text{ such that } \left\| \mathbf{X} - \hat{\mathbf{A}}\mathbf{S} \right\|_F^2 \text{ is minimal}$$

$$\hat{\mathbf{A}} = \mathbf{X}\mathbf{S}^\dagger = \mathbf{X}\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}$$

this does not specify the beamformer, but it is natural to set

$$\mathbf{W}^H = \hat{\mathbf{A}}^\dagger$$

output error minimization

objective: minimize the output error
*with **S** known*

$$\mathbf{W}^H \text{ such that } \left\| \mathbf{W}^H \mathbf{X} - \mathbf{S} \right\|_F^2 \text{ is minimal}$$

$$\mathbf{W}^H = \mathbf{S}\mathbf{X}^\dagger$$

note that $\mathbf{X}^\dagger = \mathbf{X}^H (\mathbf{X}\mathbf{X}^H)^{-1}$, so that

$$\mathbf{W}^H = \frac{1}{N}\mathbf{S}\mathbf{X}^H \left(\frac{1}{N}\mathbf{X}\mathbf{X}^H\right)^{-1} = \hat{\mathbf{R}}_{\mathbf{XS}}^H \hat{\mathbf{R}}_{\mathbf{X}}^{-1}, \qquad \mathbf{W} = \hat{\mathbf{R}}_{\mathbf{X}}^{-1} \hat{\mathbf{R}}_{\mathbf{XS}}$$

$\hat{\mathbf{R}}_{\mathbf{X}} = \frac{1}{N}\mathbf{X}\mathbf{X}^H$: sample data *covariance matrix*

$\hat{\mathbf{R}}_{\mathbf{XS}} = \frac{1}{N}\mathbf{X}\mathbf{S}^H$: sample correlation between the sources and the received data

output error minimization

objective: minimize the output error $\|\mathbf{W}^H\mathbf{X}-\mathbf{S}\|_F^2$
with **A** *known* and assuming independent sources:

observe that

$$\frac{1}{N}\mathbf{S}\mathbf{S}^H \rightarrow \mathbf{I}, \qquad \frac{1}{N}\mathbf{S}\mathbf{N}^H \rightarrow 0$$

$$\hat{\mathbf{R}}_{\mathbf{XS}} = \frac{1}{N}\mathbf{X}\mathbf{S}^H = \frac{1}{N}\mathbf{A}\mathbf{S}\mathbf{S}^H + \frac{1}{N}\mathbf{N}\mathbf{S}^H \rightarrow \mathbf{A}$$

$$\mathbf{W} = \hat{\mathbf{R}}_{\mathbf{X}}^{-1}\hat{\mathbf{R}}_{\mathbf{XS}} \rightarrow \mathbf{R}_{\mathbf{X}}^{-1}\mathbf{A}$$

$\mathbf{R}_{\mathbf{X}} = E[\mathbf{x}\mathbf{x}^H]$ is the true data covariance matrix.
with *finite samples*, we take the *estimate* from $\mathbf{X}\mathbf{X}^H$:

$$\mathbf{W} = \hat{\mathbf{R}}_{\mathbf{X}}^{-1}\mathbf{A}$$

deterministic approach

$$\mathbf{W} = \hat{\mathbf{R}}_{\mathbf{X}}^{-1}\mathbf{A}$$

this is the Linear Minimum Mean Square Error (LMMSE) or Wiener receiver.

- it maximizes the Signal-to-Interference-plus-Noise Ratio (SINR) at the output.
- it does not cancel all interference:

$$\mathbf{W}^H \mathbf{A} \neq \mathbf{I}$$

(because the loss in interference is compensated by the gain in removing noise)

Wiener filtering

remember that for a convolution filter $h(t)$ the optimal receiver in terms of output signal-to-noise ratio is given by:

$$G^*(f) = \frac{H(f)S(f)}{|H(f)|^2 S(f) + N(f)}$$

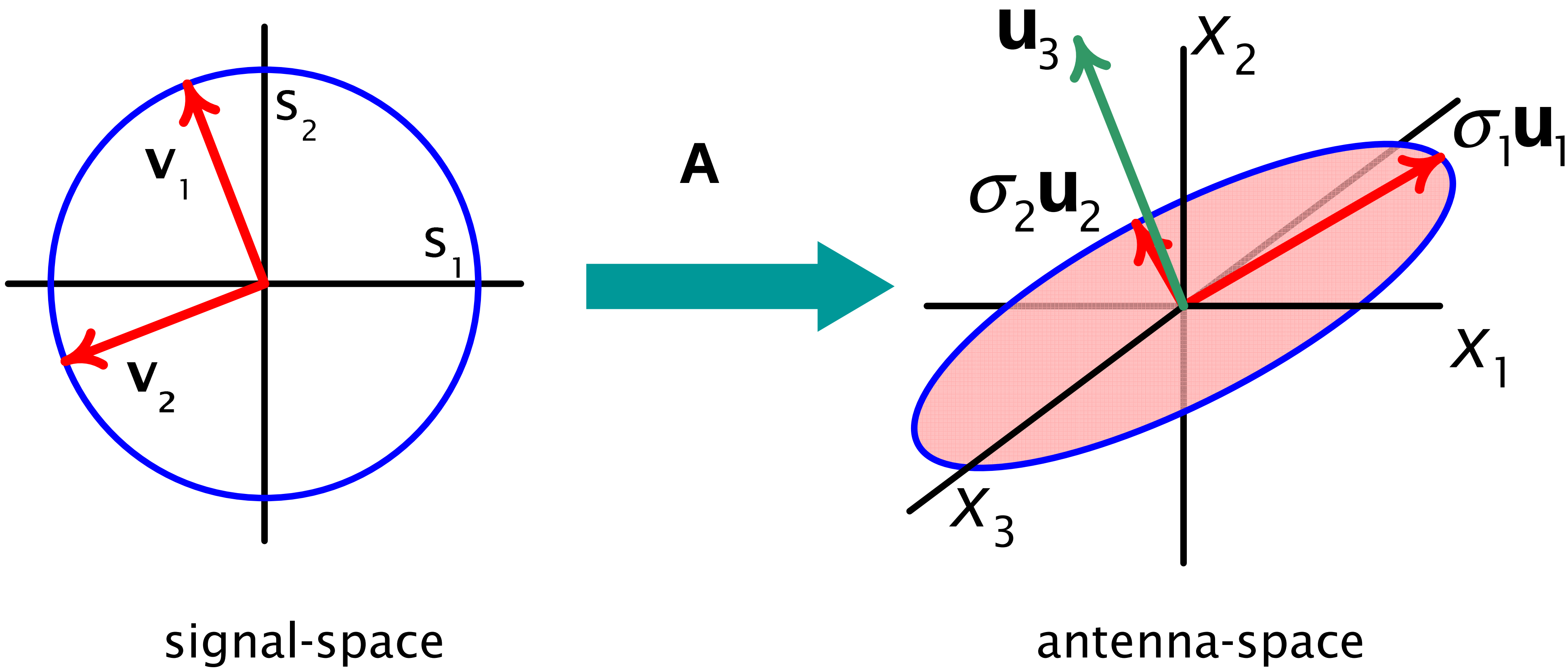
spectral decomposition of array receiver

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \qquad \mathbf{A} = \mathbf{U\Sigma V}^H$$

- columns of **V** are vectors in ‘signal-space’
- that obtain the same attenuation (singular values of $\mathbf{\Sigma}$)
- their impact in ‘antenna–space’ are the columns of U

spectral decomposition

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$$



Wiener filtering

spectral decomposition of array receiver

$$\mathbf{X} = \mathbf{AS} + \mathbf{N} \qquad \mathbf{A} = \mathbf{U\Sigma V}^H$$

define: $\mathbf{X}^* = \mathbf{U}^H\mathbf{X} \quad \mathbf{S}^* = \mathbf{V}^H\mathbf{S} \quad \mathbf{N}^* = \mathbf{U}^H\mathbf{N}$

then

$$\begin{aligned}\mathbf{X}^* &= \mathbf{U}^H(\mathbf{AS} + \mathbf{N}) \\ &= \mathbf{U}^H\mathbf{U\Sigma V}^H\mathbf{S} + \mathbf{U}^H\mathbf{N} \\ &= \mathbf{\Sigma S}^* + \mathbf{N}^*\end{aligned}$$

Wiener filtering

signal and noise are 'spatially' white
in (transformed) signal-space and antenna-space respectively

$$\mathbf{R}_{\mathbf{n}^*} = \mathrm{E}\left[\mathbf{n}^*\mathbf{n}^{*H}\right] = \mathrm{E}\left[(\mathbf{U}^H\mathbf{n})(\mathbf{U}^H\mathbf{n})^H\right] = \mathbf{U}^H\mathrm{E}\left[\mathbf{n}\mathbf{n}^H\right]\mathbf{U} = \mathbf{U}^H\sigma_n^2\mathbf{I}\mathbf{U} = \sigma_n^2\mathbf{I}$$

$$\mathbf{R}_{\mathbf{s}^*} = \mathrm{E}\left[\mathbf{s}^*\mathbf{s}^{*H}\right] = \mathrm{E}\left[(\mathbf{V}^H\mathbf{s})(\mathbf{V}^H\mathbf{s})^H\right] = \mathbf{V}^H\mathrm{E}\left[\mathbf{s}\mathbf{s}^H\right]\mathbf{V} = \mathbf{V}^H\mathbf{I}\mathbf{V} = \mathbf{I}$$

Wiener filtering

covariance of the transformed antenna samples:

$$\mathbf{R}_{\mathbf{x}^*} = \mathrm{E}\left[\mathbf{x}^*\mathbf{x}^{*H}\right] = \mathrm{E}\left[(\boldsymbol{\Sigma}\mathbf{s}^* + \mathbf{n}^*)(\boldsymbol{\Sigma}\mathbf{s}^* + \mathbf{n}^*)^H\right] = \mathrm{E}\left[\boldsymbol{\Sigma}\mathbf{s}^*\mathbf{s}^{*H}\boldsymbol{\Sigma} + \mathbf{n}^*\mathbf{n}^{*H}\right]$$

$$= \boldsymbol{\Sigma}\mathbf{I}\boldsymbol{\Sigma} + \sigma_n^2\mathbf{I} = \boldsymbol{\Sigma}^2 + \sigma_n^2\mathbf{I}$$

a diagonal matrix! in this space, ‘virtual’ antenna elements are independent!

$$\mathbf{R}_{\mathbf{x}^*\mathbf{s}^*} = \mathrm{E}\left[\mathbf{x}^*\mathbf{s}^{*H}\right] = \mathrm{E}\left[(\boldsymbol{\Sigma}\mathbf{s}^* + \mathbf{n}^*)\mathbf{s}^{*H}\right] = \mathrm{E}\left[\boldsymbol{\Sigma}\mathbf{s}^*\mathbf{s}^{*H} + \mathbf{n}^*\mathbf{s}^{*H}\right]$$

$$= \boldsymbol{\Sigma}\mathrm{E}\left[\mathbf{s}^*\mathbf{s}^{*H}\right] + \mathrm{E}\left[\mathbf{n}^*\mathbf{s}^{*H}\right] = \boldsymbol{\Sigma} + \mathbf{0} = \boldsymbol{\Sigma}$$

Wiener filtering

thus, the Wiener receiver in the transformed space is

$$\mathbf{W}^{*H}\mathbf{X}^{*}=\mathbf{S}^{*} \quad \Rightarrow \quad \mathbf{W}^{*H}=\mathbf{S}^{*}\mathbf{X}^{*\dagger}$$

$$\mathbf{W}^{*H}=\frac{1}{N}\mathbf{S}^{*}\mathbf{X}^{*H}\left(\frac{1}{N}\mathbf{X}^{*}\mathbf{X}^{*H}\right)^{-1}=\hat{\mathbf{R}}^{H}_{\mathbf{X}^{*}\mathbf{S}^{*}}\hat{\mathbf{R}}^{-1}_{\mathbf{X}^{*}}, \qquad \mathbf{W}^{*}=\hat{\mathbf{R}}^{-1}_{\mathbf{X}^{*}}\hat{\mathbf{R}}_{\mathbf{X}^{*}\mathbf{S}^{*}}$$

$$\mathbf{W}^{*} \rightarrow \left(\boldsymbol{\Sigma}^{2}+\sigma_{n}^{2}\mathbf{I}\right)^{-1}\boldsymbol{\Sigma}$$

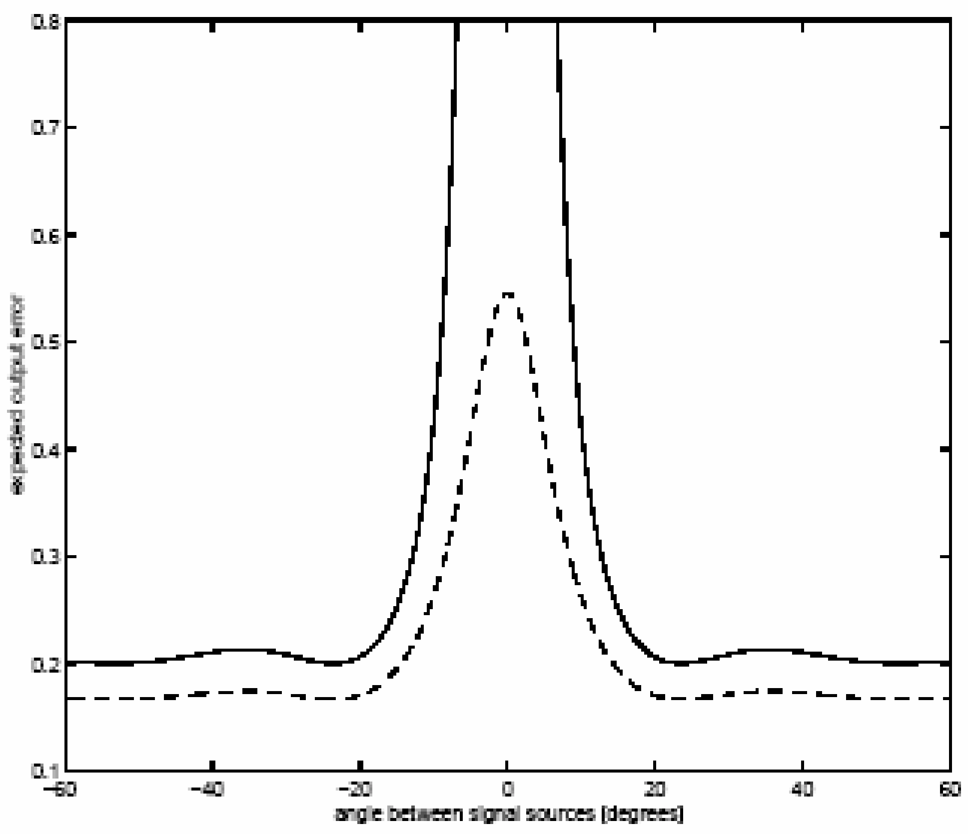
a diagonal matrix (!) with

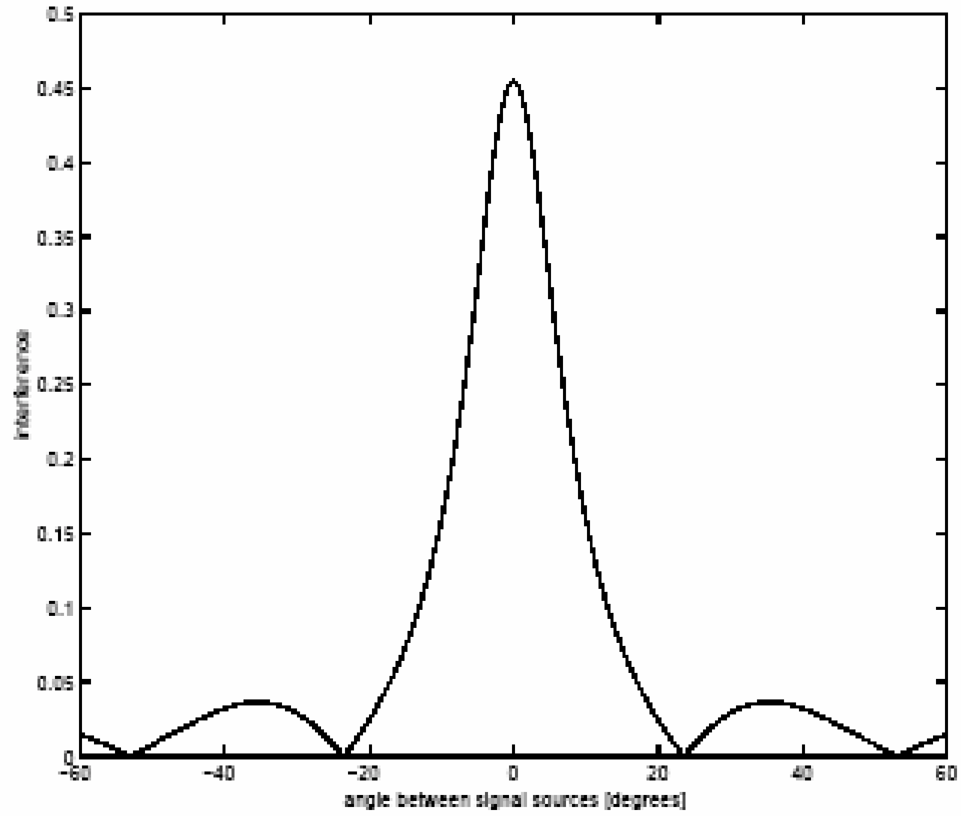
$$w^{*}_{k,k}=\frac{\sigma_k}{\sigma_k^2+\sigma_n^2} \qquad \text{compare: } G^{*}(f)=\frac{H(f)S(f)}{\left|H(f)\right|^{2}S(f)+N(f)}$$

(in this space, beam forming is trivial)

$$\mathbf{Y}=\mathbf{V}\mathbf{Y}^{*}=\mathbf{V}\mathbf{W}^{*H}\mathbf{X}^{*}=\mathbf{V}\mathbf{W}^{*H}\mathbf{U}^{H}\mathbf{X} \qquad \mathbf{W}=\mathbf{U}\mathbf{W}^{*}\mathbf{V}^{H}$$

comparing receivers





output error

zero-forcing receiver (solid)
Wiener receiver (dashed)

interference

Wiener receiver
(what about ZF?)

comparing receivers

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H, \qquad \boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots \sigma_d)$$

- zero forcing

$$\mathbf{W} = \mathbf{U}\mathbf{W}_{ZF}\mathbf{V}^H$$

$$\mathbf{W}_{ZF} \text{ diagonal with } w_{k,k}^{ZF} = \frac{1}{\sigma_k}$$

- Wiener receiver

$$\mathbf{W} = \mathbf{U}\mathbf{W}_{Wiener}\mathbf{V}^H$$

$$\mathbf{W}_{Wiener} \text{ diagonal with } w_{k,k}^{Wiener} = \frac{\sigma_k}{\sigma_k^2 + \sigma_n^2}$$

stochastic approach

assume a model with 1 source

$$\mathbf{x}_k = \mathbf{a}s_k + \mathbf{n}_k, \qquad y_k = \mathbf{w}^H\mathbf{x}_k = (\mathbf{w}^H\mathbf{a})s_k + (\mathbf{w}^H\mathbf{n}_k)$$

we make the following assumptions:

$$\mathrm{E}[|s_k|^2] = 1, \qquad \mathrm{E}[s_k\mathbf{n}_k^H] = \mathbf{0}$$

and define (spatial noise ‘color’):

$$\mathbf{R_n} := \mathrm{E}[\mathbf{n}_k\mathbf{n}_k^H]$$

so that

$$\begin{aligned}
\mathrm{E}[|y|^2] &= \mathrm{E}[(\mathbf{w}^H\mathbf{a}s_k + \mathbf{w}^H\mathbf{n}_k)(\mathbf{w}^H\mathbf{a}s_k + \mathbf{w}^H\mathbf{n}_k)^H] \\
&= \mathrm{E}[\mathbf{w}^H\mathbf{a}\,|s_k|^2\,\mathbf{a}^H\mathbf{w} + \mathbf{w}^H\mathbf{a}s_k\mathbf{n}_k^H\mathbf{w} + \mathbf{w}^H\mathbf{n}_k\bar{s}_k\mathbf{a}^H\mathbf{w} + \mathbf{w}^H\mathbf{n}_k\mathbf{n}_k^H\mathbf{w}] \\
&= \mathrm{E}[\mathbf{w}^H\mathbf{a}\mathbf{a}^H\mathbf{w} + 0 + 0 + \mathbf{w}^H\mathbf{n}_k\mathbf{n}_k^H\mathbf{w}] \\
&= (\mathbf{w}^H\mathbf{a})(\mathbf{a}^H\mathbf{w}) + \mathbf{w}^H\mathbf{R_n}\mathbf{w}
\end{aligned}$$

stochastic approach

$$E[|y|^2] = \mathbf{w}^H \mathbf{a}\mathbf{a}^H \mathbf{w} + \mathbf{w}^H \mathbf{R_n} \mathbf{w}$$

signal noise

so the Signal to Noise Ratio (SNR) at the output of the Wiener receiver is:

$$SNR_{out}(\mathbf{w}) = \frac{E[|(\mathbf{w}^H \mathbf{a}) s_k|^2]}{E[|(\mathbf{w}^H \mathbf{n}_k)|^2]} = \frac{\mathbf{w}^H \mathbf{a}\mathbf{a}^H \mathbf{w}}{\mathbf{w}^H \mathbf{R_n} \mathbf{w}}$$

stochastic approach

two stochastic optimization criteria

- maximum likelihood

  the likelihood of a set of data **X** *is the probability (density) of obtaining that particular set of data, given the transmitted signal* **S**. *The Maximum Likelihood Estimate of* **S** *is the value of* **S** *that maximizes the probability of receiving that particular* **X**. (compare deterministic model matching)

- stochastic output error minimization

  the error at the output y of the receiver is a stochastic process, which depends on the weight vector ***w***. *Choose the weight vector* ***w***, *which minimizes the variance of this process.* (compare deterministic output error minimization)

stochastic model matching

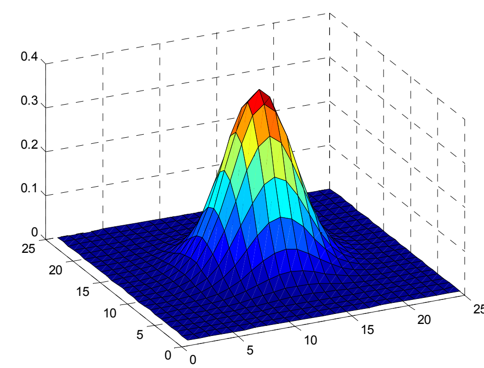
assume a model with d sources

$$\mathbf{x}_k = \mathbf{A}\mathbf{s}_k + \mathbf{n}_k \quad (k = 1,\ldots,N) \qquad \Leftrightarrow \qquad \mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N}$$

assume s_k to be deterministic

noise i.i.d. in time (temporally white), and spatially white ($\mathbf{R}_n=\sigma^2\mathbf{I}$) and jointly complex Gaussian distributed

$$\mathbf{n}_k \sim CN(\mathbf{0},\sigma^2\mathbf{I}) \quad \Leftrightarrow \quad p(\mathbf{n}_k) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{\|\mathbf{n}_k\|^2}{\sigma^2}}$$



stochastic model matching

$\mathbf{n}_k = \mathbf{x}_k - \mathbf{A}\mathbf{s}_k$ so the probability (density) to receive a certain vector $\mathbf{x}_k$, when $\mathbf{s}_k$ has been transmitted is:

$$p(\mathbf{x}_k \mid \mathbf{s}_k) = p(\mathbf{n}_k = \mathbf{x}_k - \mathbf{A}\mathbf{s}_k) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{\|\mathbf{x}_k - \mathbf{A}\mathbf{s}_k\|^2}{\sigma^2}}$$

$$p(\mathbf{X} \mid \mathbf{S}) = \prod_{k=1}^{N} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{\|\mathbf{x}_k - \mathbf{A}\mathbf{s}_k\|^2}{\sigma^2}} = \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^N e^{-\frac{\sum \|\mathbf{x}_k - \mathbf{A}\mathbf{s}_k\|^2}{\sigma^2}}$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma}}\right)^N e^{-\frac{\|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2}{\sigma^2}}$$

$$= const \cdot e^{-\frac{\|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2}{\sigma^2}}$$

*(p(**X**|**S**) is the likelihood of receiving a certain data matrix **X**, for a given transmitted data matrix **S**)*

stochastic model matching

$$p(\mathbf{X}\,|\,\mathbf{S}) = const \cdot e^{-\frac{\|\mathbf{X}-\mathbf{AS}\|_F^2}{\sigma^2}}$$

deterministic maximum likelihood technique: *estimate* ***S*** *as the one which maximizes the likelihood of the actually received* ***X***

$$\hat{\mathbf{S}} \text{ such that } p(\mathbf{X}\,|\,\mathbf{S}) \text{ is maximal}$$

$$\Leftrightarrow \quad \hat{\mathbf{S}} \text{ such that } \mathrm{e}^{-\frac{\|\mathbf{X}-\mathbf{AS}\|_F^2}{\sigma^2}} \text{ is maximal}$$

$$\Leftrightarrow \quad \hat{\mathbf{S}} \text{ such that } \|\mathbf{X}-\mathbf{AS}\|_F^2 \text{ is minimal}$$

for white Gaussian noise, Maximum Likelihood is equivalent to deterministic model matching (hence, the solutions are also the same)

stochastic derivation

stochastic output error minimization

minimize the Linear Minimum Mean Square Error cost:

$$J(w) = E\left[\left|\mathbf{w}^H\mathbf{x}_k - \mathbf{s}_k\right|^2\right]$$

it can be worked out as follows:

$$J(w) = E\left[\left|\mathbf{w}^H\mathbf{x}_k - s_k\right|^2\right] = E\left[(\mathbf{w}^H\mathbf{x}_k - s_k)(\mathbf{w}^H\mathbf{x}_k - s_k)^H\right]$$

$$= \mathbf{w}^H E\left[\mathbf{x}_k\mathbf{x}_k^H\right]\mathbf{w} - \mathbf{w}^H E\left[\mathbf{x}_k\bar{s}_k\right] - E\left[s_k\mathbf{x}_k^H\right]\mathbf{w} + E\left[\left|s_k\right|^2\right]$$

$$= \mathbf{w}^H\mathbf{R}_\mathbf{x}\mathbf{w} - \mathbf{w}^H E\left[\mathbf{a}s_k\bar{s}_k\right] - E\left[s_k\bar{s}_k\mathbf{a}_k^H\right]\mathbf{w} + E\left[\left|s_k\right|^2\right]$$

if s_k is regarded stochastic with $E[|s_k|^2]=1$, then

$$J(w) = \mathbf{w}^H\mathbf{R}_\mathbf{x}\mathbf{w} - \mathbf{w}^H\mathbf{a} - \mathbf{a}^H\mathbf{w} + 1$$

stochastic output error minimization

$$J(\mathbf{w}) = \mathbf{w}^H \mathbf{R_x} \mathbf{w} - \mathbf{w}^H \mathbf{a} - \mathbf{a}^H \mathbf{w} + 1$$

differentiate with respect to $\overline{\mathbf{w}}$: let $\overline{\mathbf{w}}=\mathbf{u}$-j$\mathbf{v}$ wit $\mathbf{u}$ and $\mathbf{v}$ real valued, then the gradient is

$$\nabla_{\overline{\mathbf{w}}} J = \frac{1}{2}\nabla_{\mathbf{u}} J - \frac{1}{2} j \nabla_{\mathbf{v}} J = \frac{1}{2}\begin{bmatrix} \frac{\partial}{\partial u_1} J \\ \vdots \\ \frac{\partial}{\partial u_d} J \end{bmatrix} - \frac{1}{2} j \begin{bmatrix} \frac{\partial}{\partial v_1} J \\ \vdots \\ \frac{\partial}{\partial v_d} J \end{bmatrix}$$

with properties $\nabla_{\overline{\mathbf{w}}} \mathbf{w}^H \mathbf{a} = \mathbf{a}, \quad \nabla_{\overline{\mathbf{w}}} \mathbf{a}^H \mathbf{w} = 0 \quad \nabla_{\overline{\mathbf{w}}} \mathbf{w}^H \mathbf{R_x} \mathbf{w} = \mathbf{R_x} \mathbf{w}$

$$\Rightarrow \quad \nabla_{\overline{\mathbf{w}}} J = \mathbf{R_x} \mathbf{w} - \mathbf{a}$$

the minimum $J(\mathbf{w})$ is attained for

$$\nabla_{\overline{\mathbf{w}}} J = \mathbf{0} \quad \Rightarrow \quad \mathbf{w} = \mathbf{R_x}^{-1} \mathbf{a}$$

we thus obtain the Wiener receiver

stochastic output error minimization

the minimum $J(\mathbf{w})$ is attained for the Wiener receiver

$$\mathbf{w} = \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{a}$$

the expected output error becomes:

$$\begin{aligned} J_{\min} &= \mathbf{a}^H\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{R}_{\mathbf{x}}\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{a} - \mathbf{a}^H\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{a} - \mathbf{a}^H\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{a} + 1 \\ &= 1 - \mathbf{a}^H\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{a} \\ &= 1 - \mathbf{a}^H\left(\mathbf{a}\sigma_s^2\mathbf{a}^H + \mathbf{R}_{\mathbf{n}}\right)^{-1}\mathbf{a} \\ &= 1 - \mathbf{a}^H\left(\mathbf{a}\mathbf{a}^H + \sigma_n^2\mathbf{I}\right)^{-1}\mathbf{a} \end{aligned}$$

stochastic output error minimization

the expected output error becomes:

$$
\begin{aligned}
J_{\min} &= 1-\mathbf{a}^H\left(\mathbf{a}\mathbf{a}^H+\sigma_n^2\mathbf{I}\right)^{-1}\mathbf{a} \\
&= 1-\mathbf{a}^H\left(\mathbf{a}\mathbf{a}^H+\sigma_n^2\mathbf{I}\right)^{-1}\mathbf{a}\left(\mathbf{a}^H\mathbf{a}+\sigma_n^2\right)\left(\mathbf{a}^H\mathbf{a}+\sigma_n^2\right)^{-1} \\
&= 1-\mathbf{a}^H\left(\mathbf{a}\mathbf{a}^H+\sigma_n^2\mathbf{I}\right)^{-1}\left(\mathbf{a}\mathbf{a}^H\mathbf{a}+\sigma_n^2\mathbf{a}\right)\left(\mathbf{a}^H\mathbf{a}+\sigma_n^2\right)^{-1} \\
&= 1-\mathbf{a}^H\left(\mathbf{a}\mathbf{a}^H+\sigma_n^2\mathbf{I}\right)^{-1}\left(\mathbf{a}\mathbf{a}^H+\sigma_n^2\mathbf{I}\right)\mathbf{a}\left(\mathbf{a}^H\mathbf{a}+\sigma_n^2\right)^{-1} \\
&= 1-\mathbf{a}^H\mathbf{a}\left(\mathbf{a}^H\mathbf{a}+\sigma_n^2\right)^{-1}
\end{aligned}
$$

$J_{\min} \to 1$ if $\sigma_n^2 \gg \|\mathbf{a}\|^2$ $\qquad$ $J_{\min} \to 0$ if $\sigma_n^2 \to 0$

colored noise

what if noise vector is not spatially white?

assume noise has a known covariance E[$\mathbf{n}\mathbf{n}^H$]=$\mathbf{R}_n \neq \sigma_2 \mathbf{I}$

'prewhiten' the data with a *square root factor* $\mathbf{R}_n^{-1/2}$

$$\mathbf{x}_k = \mathbf{A}\mathbf{s}_k + \mathbf{n}_k \quad \Rightarrow \quad \mathbf{R}_n^{-1/2}\mathbf{x}_k = \mathbf{R}_n^{-1/2}\mathbf{A}\mathbf{s}_k + \mathbf{R}_n^{-1/2}\mathbf{n}_k$$

$$\underline{\mathbf{x}}_k = \underline{\mathbf{A}}\mathbf{s}_k + \underline{\mathbf{n}}_k$$

note that $\mathbf{R}_{\underline{n}} = E\left[\underline{\mathbf{n}}_k \underline{\mathbf{n}}_k^H\right] = \mathbf{R}_n^{-1/2}\mathbf{R}_n\mathbf{R}_n^{-1/2} = \mathbf{I}$

so that the noise $\underline{\mathbf{n}}_k$ is white, with variance σ^2=1

colored noise

covariance matrices have a square root:

$$\mathbf{R_X} = \mathbf{XX}^H \qquad \mathbf{X} = \mathbf{U\Sigma V}^H$$

$$\begin{aligned}\mathbf{R_X} &= \mathbf{U\Sigma V}^H\mathbf{V\Sigma U}^H \\ &= \mathbf{U\Sigma}^2\mathbf{U}^H\end{aligned}$$

$$\mathbf{R_X}^{1/2} = \mathbf{U\Sigma U}^H \qquad \mathbf{R_X}^{-1/2} = \mathbf{U\Sigma}^{-1}\mathbf{U}^H$$

check that

$$\mathbf{R_X}^{1/2}\mathbf{R_X}^{1/2} = \mathbf{R_X} \qquad \mathbf{R_X}^{1/2}\mathbf{R_X}^{-1/2} = \mathbf{I}$$

colored noise

we use the known results on the new variables:

$$\underline{\mathbf{A}} = \mathbf{R}_n^{-1/2}\mathbf{A} \qquad \underline{\mathbf{x}}_k = \mathbf{R}_n^{-1/2}\mathbf{x}_k \qquad \underline{\mathbf{n}}_k = \mathbf{R}_n^{-1/2}\mathbf{n}_k$$

the ZF equalizer becomes

$$\begin{aligned} \mathbf{y}_k &= \underline{\mathbf{A}}^{\dagger}\underline{\mathbf{x}}_k = (\underline{\mathbf{A}}^H\underline{\mathbf{A}})^{-1}\underline{\mathbf{A}}^H\underline{\mathbf{x}}_k \\ &= (\mathbf{A}^H(\mathbf{R}_n^{-1/2})^H\mathbf{R}_n^{-1/2}\mathbf{A})^{-1}\mathbf{A}^H(\mathbf{R}_n^{-1/2})^H\mathbf{R}_n^{-1/2}\mathbf{x}_k \\ &= (\mathbf{A}^H\mathbf{R}_n^{-1}\mathbf{A})^{-1}\mathbf{A}^H\mathbf{R}_n^{-1}\mathbf{x}_k = \mathbf{W}^H\mathbf{x}_k \\ &\Rightarrow \quad \mathbf{W} = \mathbf{R}_n^{-1}\mathbf{A}(\mathbf{A}^H\mathbf{R}_n^{-1}\mathbf{A})^{-1} \end{aligned}$$

the Wiener receiver is the same as before, since $\mathbf{R}_\text{n}$ is not used in the derivation *check:*

$$\underline{\mathbf{W}} = \underline{\mathbf{R}}_{\mathbf{x}}^{-1}\underline{\mathbf{A}} = (\mathbf{R}_n^{-1/2}\mathbf{R}_{\mathbf{x}}\mathbf{R}_n^{-1/2})^{-1}\mathbf{R}_n^{-1/2}\mathbf{A} = \mathbf{R}_n^{1/2}\mathbf{R}_{\mathbf{x}}^{-1}\mathbf{A}$$

$$\Rightarrow \quad \mathbf{Y} = \underline{\mathbf{W}}^H\underline{\mathbf{X}} = \underline{\mathbf{W}}^H\mathbf{R}_n^{-1/2}\mathbf{X} = \mathbf{W}^H\mathbf{X}$$

$$\Rightarrow \quad \mathbf{W} = \mathbf{R}_n^{-1/2}\underline{\mathbf{W}} = \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{A}$$

the matched filter

consider a *single* signal in white noise,

$$\mathbf{x}_k = \mathbf{a}s_k + \mathbf{n}_k, \qquad E\left[\mathbf{n}_k\mathbf{n}_k^H\right] = \sigma^2\mathbf{I}$$

the ZF beamformer is given by

$$\mathbf{w} = (\mathbf{a}^\dagger)^H = \mathbf{a}(\mathbf{a}^H\mathbf{a})^{-1} = \gamma\mathbf{a}$$

note: a scalar multiplication does not change the output SNR

w=a is known as a matched filter, classical beamformer and Maximum Ratio Combining (MRC)

the matched filter

with non-white noise,

$$\mathbf{x}_k = \mathbf{a}s_k + \mathbf{n}_k, \qquad E\left[\mathbf{n}_k\mathbf{n}_k^H\right] = \mathbf{R}_\mathbf{n}$$

we have seen that the zero forcing beam former in non-white noise equals:

$$\mathbf{W} = \mathbf{R}_n^{-1}\mathbf{A}(\mathbf{A}^H\mathbf{R}_n^{-1}\mathbf{A})^{-1}$$

hence, in case of a single signal we get:

$$\mathbf{w} = \mathbf{R}_\mathbf{n}^{-1}\mathbf{a}(\mathbf{a}^H\mathbf{R}_\mathbf{n}^{-1}\mathbf{a})^{-1} = \gamma\mathbf{R}_\mathbf{n}^{-1}\mathbf{a}$$

thus, the *matched filter in non-white noise* is

$$\mathbf{w} = \mathbf{R}_\mathbf{n}^{-1}\mathbf{a}$$

the matched filter

similarly the Wiener filter, in white noise,

$$
\begin{aligned}
\mathbf{w} = \mathbf{R}_{\mathbf{x}}^{-1}\mathbf{a} &= (\mathbf{a}\mathbf{a}^H + \sigma^2\mathbf{I})^{-1}\mathbf{a} \\
&= (\mathbf{a}\mathbf{a}^H + \sigma^2\mathbf{I})^{-1}\mathbf{a}(\mathbf{a}^H\mathbf{a} + \sigma^2)(\mathbf{a}^H\mathbf{a} + \sigma^2)^{-1} \\
&= (\mathbf{a}\mathbf{a}^H + \sigma^2\mathbf{I})^{-1}(\mathbf{a}\mathbf{a}^H\mathbf{a} + \sigma^2\mathbf{a})(\mathbf{a}^H\mathbf{a} + \sigma^2)^{-1} \\
&= (\mathbf{a}\mathbf{a}^H + \sigma^2\mathbf{I})^{-1}(\mathbf{a}\mathbf{a}^H\mathbf{a} + \sigma^2\mathbf{I}\mathbf{a})(\mathbf{a}^H\mathbf{a} + \sigma^2)^{-1} \\
&= (\mathbf{a}\mathbf{a}^H + \sigma^2\mathbf{I})^{-1}(\mathbf{a}\mathbf{a}^H + \sigma^2\mathbf{I})\mathbf{a}(\mathbf{a}^H\mathbf{a} + \sigma^2)^{-1} \\
&= \mathbf{a}(\mathbf{a}^H\mathbf{a} + \sigma^2)^{-1} \\
&= \gamma\mathbf{a}
\end{aligned}
$$

it is also equal to a multiple of the matched filter!

the matched filter

Wiener Filter in colored noise, (using whitening)

$$\underline{\mathbf{a}} = \mathbf{R}_n^{-1/2}\mathbf{a}$$

$$\begin{aligned}\mathbf{w} &= \mathbf{R}_\mathbf{x}^{-1}\mathbf{a} = (\mathbf{a}\mathbf{a}^H + \mathbf{R}_\mathbf{n})^{-1}\mathbf{a} \\ &= \mathbf{R}_\mathbf{n}^{-1/2}(\underline{\mathbf{a}}\underline{\mathbf{a}}^H + \mathbf{I})^{-1}\underline{\mathbf{a}} \\ &= \mathbf{R}_\mathbf{n}^{-1/2}\underline{\mathbf{a}}(\underline{\mathbf{a}}^H\underline{\mathbf{a}} + 1)^{-1} \\ &= \gamma\mathbf{R}_\mathbf{n}^{-1}\mathbf{a}\end{aligned}$$

this is again equal to a multiple of the matched filter for colored noise

the matched filter

the colored noise case is relevant also for the following reason:
with more than one signal, we can write the model as

$$\mathbf{x}_k = \mathbf{A}\mathbf{s}_k + \mathbf{n}_k = \mathbf{a}_1 s_k + (\mathbf{A}'\mathbf{s}'_k + \mathbf{n}_k)$$

this is of the form

$$\mathbf{x}_k = \mathbf{a} s_k + \mathbf{n}_k, \qquad \mathbf{R}_\mathbf{n} = \mathbf{A}'\mathbf{A}'^H + \sigma^2 \mathbf{I}$$

where the “noise” is colored due to the contribution of the interfering sources.

matched filter

in summary,

- the matched filter is optimal with only one signal
- equal to both zero-forcing and Wiener filter
  (with one signal there is no interference)
- with colored noise, they also coincide
- noise color can be used to capture interference of other signals

summary

- criteria: model matching vs. output error minimization
- solutions: zero-forcing, Wiener filter
- zero forcing optimizes: model matching, interference
- Wiener filter optimizes: output error, SNR
- matched filter: for a single signal both coincide and the beam former equal the response vector
- both deterministic and stochastic analysis
- colored noise