Statistical Signal Processing: Applications to Beamforming, Detection, and Estimation*

Louis L. Scharf+
Electrical and Computer Engineering, and Statistics
Colorado State University, Ft Collins, CO 80523

SPIE Ultrasound Meeting, Feb. 18, 2004

* Supported by the Office of Naval Research and the National Science Foundation

+ Special thanks to Magnus Lundberg & Hongya Ge for insights on beamforming & conjugate gradient filtering.

A serious talk by one of you on imaging would probably address:

- physics of imaging
- biology of systems
- processing of information

It would have a concrete result for:

- nuclear magnetic imaging
- positron emission tomography
- computer axial tomography
- ultrasound

Medical Imaging 2004: Ultrasonic Imaging and Signal Processing,
edited by William F. Walker, Stanislav Y. Emelianov, Proc. of SPIE Vol. 5373
(SPIE, Bellingham, WA, 2004) · 1605-7422/04/$15 · doi: 10.1117/12.541397

This will not be possible this morning. What might be possible is to:

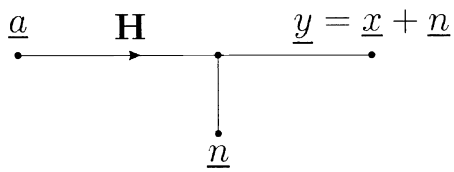
- review some new and old ideas in statistical signal processing,
- bring a little more intuition for what you already know,
- suggest new ways for you to think about what you do, and perhaps suggest new directions you might take.

With this in mind, let's

- Review the geometry of signal processing in low-dimensional subspaces.
- Establish some performance bounds, all of which have a revealing geometry.
- Briefly comment on matched subspace detectors and their application to spectrum analysis.
- Compare time-frequency distributions to scattering functions for active imaging (beamforming).
- Present ongoing work on multi-rank Bartlett and Capon beamforming to manage field mismatches, and connect with recent work on subspace expanding estimators based on conjugate gradients.

Linear Models & Subspace Signal Processing

Apriori Algebra:



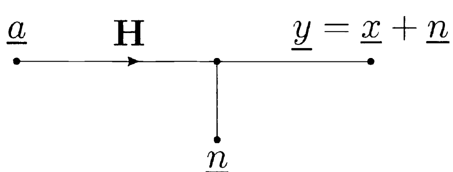
$$\underline{x} = \mathbf{H}\underline{a} = \underline{h}_1 a_1 + \sum_{i=2}^{p} \underline{h}_i a_i$$

$$= [\underline{h}_1, \mathbf{H}_1] \begin{bmatrix} a_1 \\ \mathbf{A}_1 \end{bmatrix}$$

∇SPIE Ultrasound 2004: Statistical Signal Processing: Applications to Beamforming, Detection, and Estimation – p.5-25

Linear Models & Subspace Signal Processing

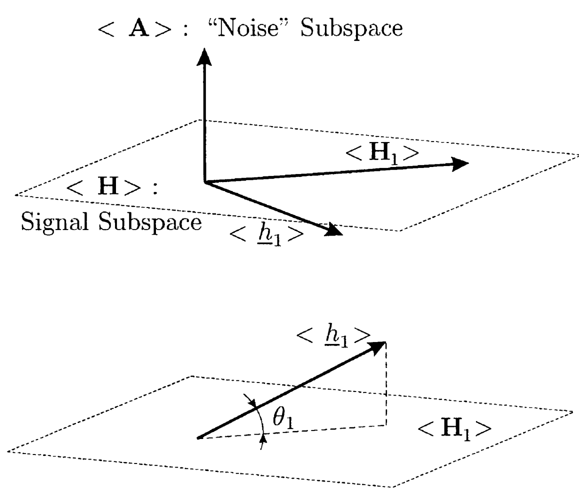
Apriori Algebra:



$$\underline{x} = \mathbf{H}\underline{a} = \underline{h}_1 a_1 + \sum_{i=2}^{p} \underline{h}_i a_i$$

$$= [\underline{h}_1, \mathbf{H}_1] \begin{bmatrix} a_1 \\ \mathbf{A}_1 \end{bmatrix}$$

Apriori Geometry:



SPIE Ultrasound 2004: Statistical Signal Processing: Applications to Beamforming, Detection, and Estimation – p.5-25

Linear Models & Subspace Signal Processing

Aposteriori Algebra:

$$\mathbf{I} = \mathbf{E}_{\underline{h}_1\mathbf{H}_1} + \mathbf{E}_{\mathbf{H}_1\underline{h}_1} + \mathbf{P}_{\mathbf{A}}$$

(*3-way resolution of identity*)

$$\mathbf{E}_{\underline{h}_1\mathbf{H}_1} = \underline{h}_1 \left(\underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{h}_1\right)^{-1} \underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp}$$

$$\mathbf{P}_{\mathbf{H}_1}^{\perp} = \mathbf{I} - \mathbf{H}_1 \left(\mathbf{H}_1^* \mathbf{H}_1\right)^{-1} \mathbf{H}_1^*$$

(*both idempotent*)

$$\mathbf{E}_{\underline{h}_1\mathbf{H}_1}\underline{h}_1 = \underline{h}_1 \quad \& \quad \mathbf{E}_{\underline{h}_1\mathbf{H}_1}\mathbf{H}_1 = \underline{0}$$

(*perfect imaging*)

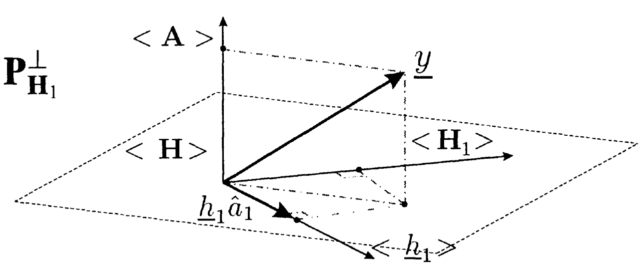
Linear Models & Subspace Signal Processing

Aposteriori Algebra:

Aposteriori Geometry:

$$\mathbf{I} = \mathbf{E}_{\underline{h}_1\mathbf{H}_1} + \mathbf{E}_{\mathbf{H}_1\underline{h}_1} + \mathbf{P}_{\mathbf{A}}$$

(*3-way resolution of identity*)



$$\mathbf{E}_{\underline{h}_1\mathbf{H}_1} = \underline{h}_1 \left(\underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{h}_1\right)^{-1} \underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp}$$

$$\mathbf{P}_{\mathbf{H}_1}^{\perp} = \mathbf{I} - \mathbf{H}_1 \left(\mathbf{H}_1^* \mathbf{H}_1\right)^{-1} \mathbf{H}_1^*$$

(*both idempotent*)

$$\mathbf{E}_{\underline{h}_1\mathbf{H}_1}\underline{h}_1 = \underline{h}_1 \quad \& \quad \mathbf{E}_{\underline{h}_1\mathbf{H}_1}\mathbf{H}_1 = \underline{0}$$

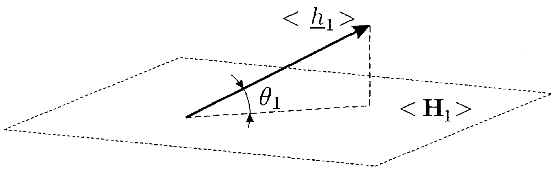
(*perfect imaging*)

Performance: Matched Subspace Filter

$$\hat{a}_1 = \left(\underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{h}_1\right)^{-1} \underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{y} \quad : \quad \left\{ a_1, \frac{1}{\underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{h}_1} \right\}$$

$$\underline{h}_1 \hat{a}_1 = \mathbf{E}_{\underline{h}_1 \mathbf{H}_1} \underline{y} \quad : \quad \left\{ \underline{h}_1 a_1, \underline{h}_1 \left(\underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{h}_1\right)^{-1} \underline{h}_1^* \right\}$$

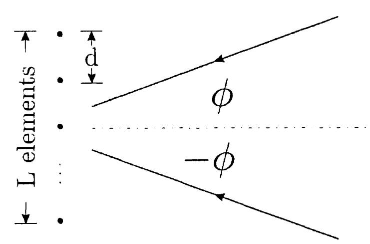
$$\text{MSE} = \text{Tr(error cov)} = \frac{\underline{h}_1^* \underline{h}_1}{\underline{h}_1^* \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{h}_1} = \frac{1}{\sin^2(\theta_1)}$$

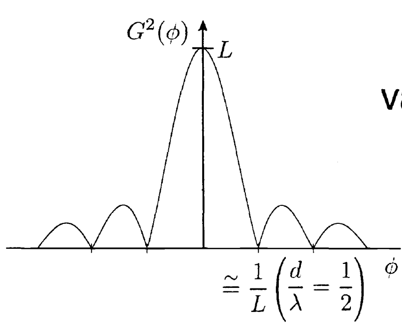


It is the "nearness" of mode $\underline{h}_1$ to interfering modes $\mathbf{H}_1$ that accounts for noise gain!

Example

Super-resolution of plane-waves in a linear array





$$\text{var}(\hat{a}_1) = \frac{\sigma^2/L}{1 - G^2(\phi)} = \frac{1/\text{SNR}}{\sin^2\theta_1}$$

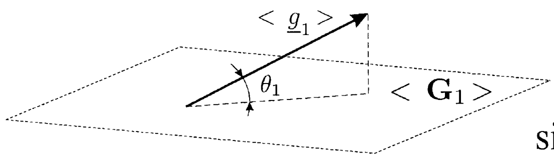
$$G^2(\phi) = \frac{\sin^2\left(2\pi\frac{d}{\lambda}L\sin\phi\right)}{\sin^2\left(2\pi\frac{d}{\lambda}\sin\phi\right)}$$

$$\text{SNR} = L/\sigma^2 \;\&\; \sin^2\theta_1 = 1 - G^2(\phi)$$

Super-resolution does not work except at high SNR.

Cramér-Rao Bound (CRB)

This performance result is exact for a known estimator. Sometimes the computation for a known estimator is elusive, and at other times the estimator is unknown. Then we would like to know how much information the data carries about a parameter, without specifying how we extract this info. If the answer is too pessimistic, we must redesign our experiment.



$$\begin{aligned} \text{var}(\hat{\theta}_1) &\geq 1/\left(\text{SNR}\sin^2\theta\right) \\ \text{SNR} &= \underline{g}_1^* \underline{g}_1/\sigma^2 \\ \sin^2\theta_1 &= \frac{\underline{g}_1^* \mathbf{P}_{\mathbf{G}_1}^{\perp} \underline{g}_1}{\underline{g}_1^* \underline{g}_1} \end{aligned}$$

$$\mathbf{G} = \left[\underline{g}_1, \mathbf{G}_1\right]; \quad \underline{g}_i = \frac{\partial \underline{x}}{\partial \theta_i} : \text{sensitivity}$$

Matched Subspace Detectors

Question: Is there a significant $\underline{h}_1$ effect in the model,

$$\underline{y} = \underline{h}_1 a_1 + \mathbf{H}_1 \mathbf{A}_1 + \underline{n},$$

or are we seeing only

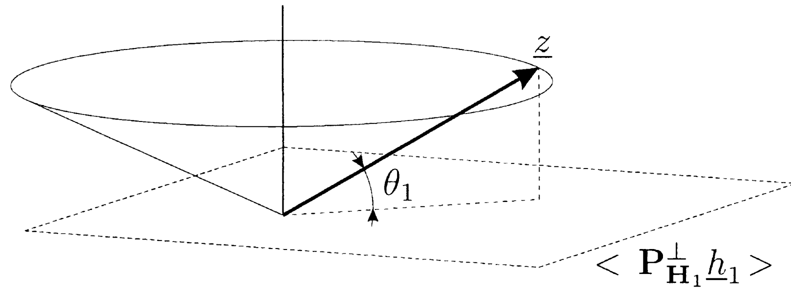
$$\underline{y} = \mathbf{H}_1 \mathbf{A}_1 + \underline{n}?$$

Test: $H_0 : a_1 = 0$ vs $H_1 : a_1 \neq 0$

The uniformly most powerful-invariant, and GLR, test is

$$\begin{aligned} \underline{z} &= \mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{y}; \\ \sin^2\theta &= \frac{\underline{z}^* \mathbf{P}_{\mathbf{P}_{\mathbf{H}_1}^{\perp} \underline{h}_1} \underline{z}}{\underline{z}^* \underline{z}} \overset{H_0}{\underset{H_1}{\gtrless}} \eta. \end{aligned}$$

The geometry and invariances are these

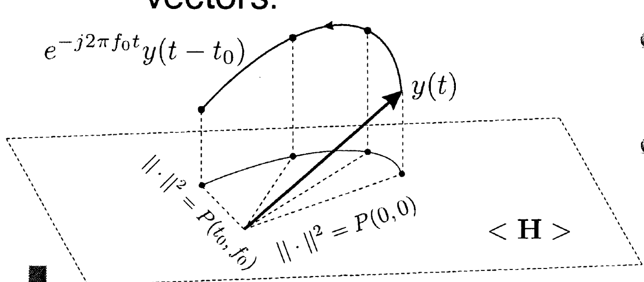


- The detector measures sine^2 of the angle between $\underline{z}$ and $< \cdot >$.
- Any rotation or scaling of $\underline{z}$ leaves $\sin^2\theta_1$ invariant. This is a good thing.
- This result extends in many ways to produce adaptive detectors.

SPIE Ultrasound 2004: Statistical Signal Processing: Applications to Beamforming, Detection, and Estimation – p.11/25

Example: Estimating Time-Frequency Distributions

- There is a version of the matched subspace detector that illuminates much of what is done in smoothed or multi-window spectrum analysis, and Rihaczek or Wigner-Ville time-frequency analysis.
- $< \mathbf{H} >$: Space of time-limited, band-limited signals, approximately spanned by $r = 2TW$ independent vectors.



- Spectral multi-windowism for $t_0 = 0$
- Time-frequency multiwindowism for $t_0 \neq 0$

SPIE Ultrasound 2004: Statistical Signal Processing: Applications to Beamforming, Detection, and Estimation – p.12/25

Intermediate Recap:

- So far everything comes down to sines of angles between subspaces.
- The subspaces change from problem to problem. But the idea, itself, remains unchanged.
- Examples:
  - for estimation, $<\underline{h}_1>$ & $<\mathbf{H}_1>$
  - for bounding, $<\underline{g}_1>$ & $<\mathbf{G}_1>$
  - for detection, $<\underline{z}>$ & $<\mathbf{P}^{\perp}_{\mathbf{H}_1}\underline{h}_1>$
  - for time-freq. analysis, $<$ Slepian $>$

Active Beamforming

- The problem is to transmit a waveform through a randomly time varying medium, and then measure some characteristic, such as the scattering function

$$\text{SF:}\ P_{\sigma\sigma}(\tau,\nu)\delta(\tau)\delta(\nu) = \mathsf{E}|\sigma(\tau,\nu)|^2$$

- The measurement is assumed to be

$$y(t) = \int\int \sigma(\tau,\nu)e^{j2\pi\nu t}x(t-\tau)d\nu d\tau$$

  i.e., a linear combination of delayed, dopplered, & complex scaled signals.
- The problem is to design the signal x so that the SF $P_{\sigma\sigma}$ may be estimated from y.

Active Beamforming

- I will not go into the details of estimation, but instead tell you that the best estimator that is quadratic in y and delay and modulation invariant will be attempting to estimate

$$(\Gamma_{xx} \cdot R_{HH})(\Delta f, \Delta t) \Leftrightarrow (V_{xx} * P_{\sigma\sigma})(\tau, \nu),$$

  where $\Gamma_{xx} \Leftrightarrow V_{xx}$ is a Fourier transform pair of ambiguity (Γ) and Rihaczek time-frequency dist. (V).
- $V_{xx}(t,f)$ is the Rihaczek TF-dist. $X(f)e^{j2\pi ft}x^*(t)$, an instantaneous inner product.
- The problem is to design the signal x for a desired V or Γ.

Active Beamforming

The reason this is interesting is that the story for Scattering Functions (SF), told this way, is dual to the story of Time-Frequency Distributions (TFD):

- SF: $(\Gamma_{xx} \cdot R_{HH})(\Delta f, \Delta t) \Leftrightarrow (V_{xx} * P_{\sigma\sigma})(\tau, \nu)$

  Design V_{xx} (Rihaczek) or Γ_{xx} (ambiguity) for <u>deconvolution</u> of $Vxx * P_{\sigma\sigma}$.

- TFD: $(\Gamma_{xx} \cdot R_{HH})(\Delta f, \Delta t) \Leftrightarrow (V_{xx} * P_{\sigma\sigma})(t, f)$

  Design $P_{\sigma\sigma}$ (time-freq. windows) or R_{HH} (ambiguity) for <u>convolution</u> $V_{xx} * P_{\sigma\sigma}$.

Examples:

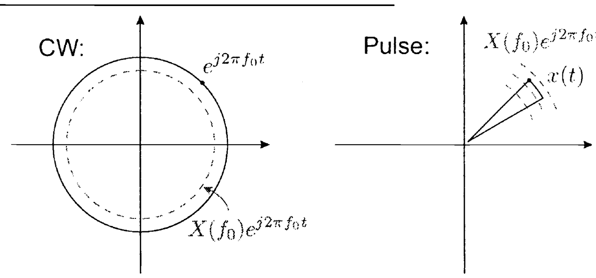
CW: $\Gamma_{xx}(\Delta f, \Delta t) = \delta(\Delta f) \;\Leftrightarrow\; V_{xx}(t,f) = \delta(f)$

: $\hat{P}_{\sigma\sigma}(t,f) = \int P_{\sigma\sigma}(t,f)dt$: frequency marginal

Pulse: $\Gamma_{xx}(\Delta f, \Delta t) = \delta(\Delta t) \;\Leftrightarrow\; V_{xx}(t,f) = \delta(t)$

: $\hat{P}_{\sigma\sigma}(t,f) = \int P_{\sigma\sigma}(t,f)df$: time marginal

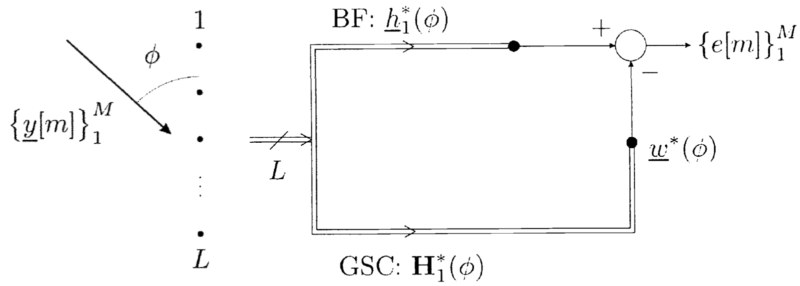
The time-frequency picture is this:



Passive Beamforming

- The problem is to image power as a function of range-doppler-angle. To simplify our arguments, let's image only as a function of angle, ϕ.
- We shall let $\underline{h}_1(\phi)$ stand for the conventional Bartlett beamformer and $\mathbf{H}_1(\phi)$ stand for the a matrix of generalized sidelobe cancellers (GSCs) that are orthogonal to $\underline{h}_1(\phi)$.
- We shall approach the issue as a 2-channel problem.

2-Channel Model



There are two things going on here. We are

1. Imaging with beamformer $\underline{h}_1$ to estimate that part of $\underline{y}$ that looks like $\underline{h}_1 a_1$, originating from angle ϕ.
2. Imaging with Generalized Sidelobe Canceller (GSC) to estimate that part of $\underline{y}$ that looks like $\underline{h}_1 a_1$, originating from angle ϕ.

Two Common Beamformers

- The conventional *Bartlett beamformer* computes the power

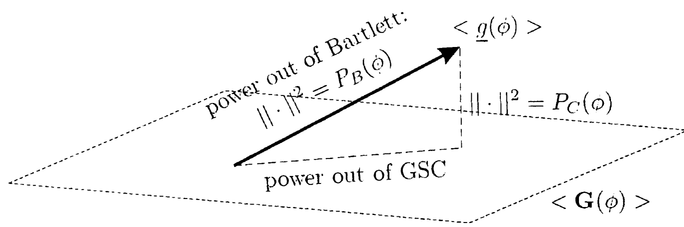
$$P_B(\phi) = \frac{1}{M}\sum_{m=1}^{M} |\underline{h}_1^*(\phi)\underline{y}(m)|^2 = \underline{h}_1^*(\phi)\mathbf{R}\underline{h}_1(\phi) = \underline{g}_1^*(\phi)\underline{g}_1(\phi)$$

- The *Capon beamformer* computes the power

$$P_C(\phi) = \frac{1}{\underline{h}_1^*(\phi)\mathbf{R}^{-1}\underline{h}_1(\phi)} = \underline{g}_1^*(\phi)\mathbf{P}^{\perp}_{\mathbf{G}(\phi)}\underline{g}_1(\phi)$$

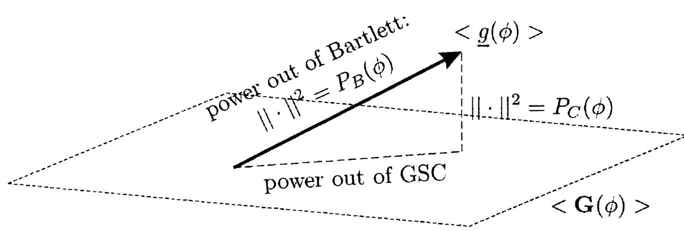
$$\boxed{\begin{aligned} \mathbf{R} &= \frac{1}{M}\sum_{m=1}^{M} \underline{y}(m)\underline{y}^*(m) \ : \ \text{sample covariance} \\ \underline{g}_1(\phi) &= \mathbf{R}^{1/2}\underline{h}_1(\phi) \ \& \ \mathbf{G}(\phi) = \mathbf{R}^{1/2}\mathbf{H}_1(\phi) \end{aligned}}$$

The Geometry

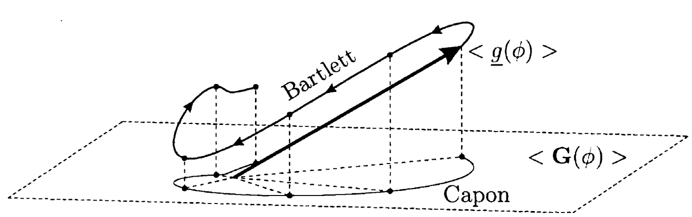


Capon and GSC orthogonally resolve Bartlett. Thus $P_C(\phi) \leq P_B(\phi)$ (Kantorovich ineq.). If this is the picture for a single angle ϕ, then the picture as we steer through angles ϕ is

The Geometry



Capon and GSC orthogonally resolve Bartlett. Thus $P_C(\phi) \leq P_B(\phi)$ (Kantorovich ineq.). If this is the picture for a single angle ϕ, then the picture as we steer through angles ϕ is



Connection to Filtering in Expanding Subspaces

- In the Capon (or MVDL) beamformer, the computation in the denominator is

$$\underline{h}_1^* \mathbf{R}^{-1} \underline{h}_1 = \underline{h}_1^* \underline{w}$$

$$\underline{w} = \mathbf{R}^{-1} \underline{h}_1 \ : \text{ Wiener filter}$$

- But the filter $\underline{w}$ is known to lie in the L-dimensional Krylov subspace

$$< K > = < \underline{h}_1, \mathbf{R}\underline{h}_1, \ldots, \mathbf{R}^{L-1} \underline{h}_1 >$$

which is known to terminate at dimension $r << L$ for many interesting problems.

- Moreover it is known how to use conjugate gradients to expand $< K >$ from $< \underline{h}_1 >$ to $< \underline{h}_1, \mathbf{R}\underline{h}_1 >$ to ...

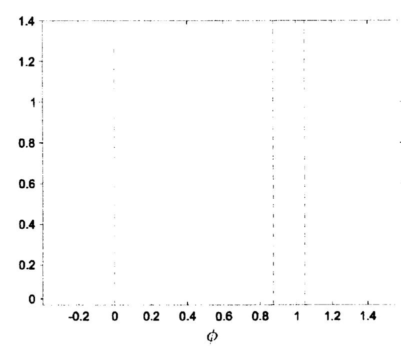
To make a long story short...

- The matrix inversion can be avoided, and the Capon beamformer may be written as

$$P_C^{(r)}(\phi) = \frac{1}{\underline{h}_1^*(\phi) \underline{w}^{(r)}(\phi)}$$

where $\underline{w}^{(r)}$ is computed recursively with CG's.

- Bearing response pattern, showing the evolution of the beamformer with r.



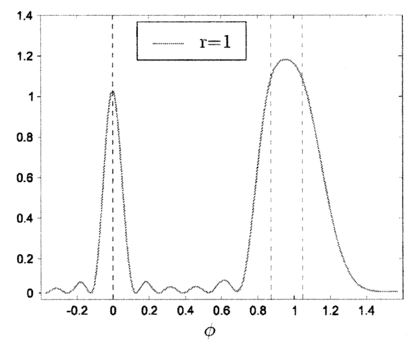
To make a long story short...

- The matrix inversion can be avoided, and the Capon beamformer may be written as

$$P_C^{(r)}(\phi) = \frac{1}{\underline{h}_1^*(\phi)\underline{w}^{(r)}(\phi)}$$

where $\underline{w}^{(r)}$ is computed recursively with CG's.

- Bearing response pattern, showing the evolution of the beamformer with r.



∇SPIE Ultrasound 2004: Statistical Signal Processing: Applications to Beamforming, Detection, and Estimation – p.23/25

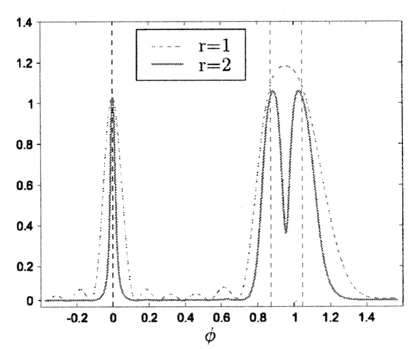
To make a long story short...

- The matrix inversion can be avoided, and the Capon beamformer may be written as

$$P_C^{(r)}(\phi) = \frac{1}{\underline{h}_1^*(\phi)\underline{w}^{(r)}(\phi)}$$

where $\underline{w}^{(r)}$ is computed recursively with CG's.

- Bearing response pattern, showing the evolution of the beamformer with r.



∇SPIE Ultrasound 2004: Statistical Signal Processing: Applications to Beamforming, Detection, and Estimation – p.23/25

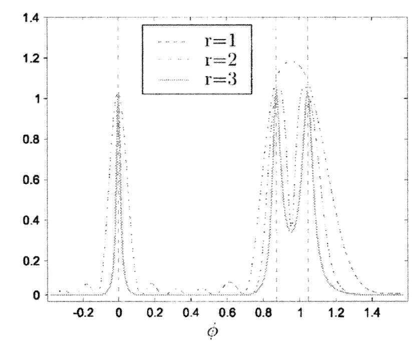
To make a long story short...

- The matrix inversion can be avoided, and the Capon beamformer may be written as

$$P_C^{(r)}(\phi) = \frac{1}{\underline{h}_1^*(\phi)\underline{w}^{(r)}(\phi)}$$

where $\underline{w}^{(r)}$ is computed recursively with CG's.

- Bearing response pattern, showing the evolution of the beamformer with r.



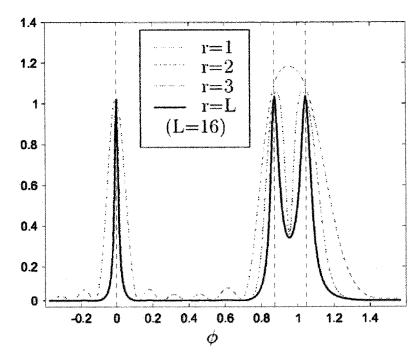
To make a long story short...

- The matrix inversion can be avoided, and the Capon beamformer may be written as

$$P_C^{(r)}(\phi) = \frac{1}{\underline{h}_1^*(\phi)\underline{w}^{(r)}(\phi)}$$

where $\underline{w}^{(r)}$ is computed recursively with CG's.

- Bearing response pattern, showing the evolution of the beamformer with r.



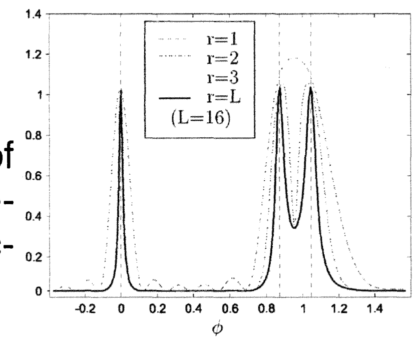
To make a long story short...

- The matrix inversion can be avoided, and the Capon beamformer may be written as

$$P_C^{(r)}(\phi) = \frac{1}{\underline{h}_1^*(\phi)\underline{w}^{(r)}(\phi)}$$

  where $\underline{w}^{(r)}$ is computed recursively with CG's.

- Bearing response pattern, showing the evolution of the beamformer with r.
- Suggests that this way of beamforming, allows for angle-dependent dimension reduction, which is a good thing.



Recap

1. Angles between signal and interfering subspaces determine performance of estimators (and detectors).
2. Matched subspace detectors actually estimate angles between measurements and subspaces.
3. Multi-window or smoothed spectrum analysis and TF analysis can be seen as subspace detection ... or ought to.
4. Active beamforming is dual to TF analysis.

Recap

5. In passive beamforming, the powers out of the Capon and GSC beamformers orthogonally decompose the power out of the Bartlett beamformer. This explains the higher resolution of the Capon.
6. There is a connection between Capon beamforming and conjugate gradient filtering, allowing for reduced-dimensional beamforming with angle-dependent dimensions.
7. Euclid and Pythagoras would be comfortable among us.