Antenna array theory & beamforming

Usual defn of $AF(\psi) = \sum_n a_n e^{jn\psi}$, with the control residing in the excitation coeffs, a_n.

→ We have already seen the case of progressive phase shifts, i.e. $a_n = \exp(j\beta n)$ where $\beta = -kd\cos\theta_0$, leads to b.f. at desired θ_0.

→ Now let's study a more practical version of this for communications. A classic algo from the DSP community:

MVDR (min variance distortionless response) beamformer (Capon)
— Balanis AT 16.4

(will need some pre-reqs, will build it along the way)

① $AF(\psi) = \sum a_n \exp(jn \underbrace{kd\cos\theta_0}_{\psi})$. Rewrite:

$$w = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-1} \end{bmatrix} \qquad a(\theta_0) = \begin{bmatrix} 1 \\ \exp(jkd\cos\theta_0) \\ \vdots \\ \exp(jkd(N-1)\cos\theta_0) \end{bmatrix}$$

$a(\theta_0)$ → steering vector.

why do this?

↳ Thus separate the excitations and geometry

can change ↓ (excitations) ; fixed after fab. ↓ (geometry)

$$AF(\psi) = w^T a(\theta_0).$$

Note: In the signal processing world, the receive beamforming output is written as $y = w^H x$, where the input is as: $x(t) = s(t)\, a(\theta_0)$, i.e. a signal $s(t)$ coming from θ_0 direction. H instead of T.

∴ In this convention, the excitation coefficients are $a_n = w_n^*$, not w_n.

[In either case, the received power $|AF|^2$ is the same]. So we proceed with $y = w^H x$.

Practice Q: How do we write the result of the conventional progressive phase shift result in this steering vector notation?

We want $\max_{w} \left[|w^H a(\theta_0)| \right]$ → objective function.

w ↳ variable of optimization

$$y = w^H a = \sum_n w_n^* e^{j\psi_0 n}. \quad \text{If } w_n^* = e^{-jn\psi_0}$$

then each term has max possible value.

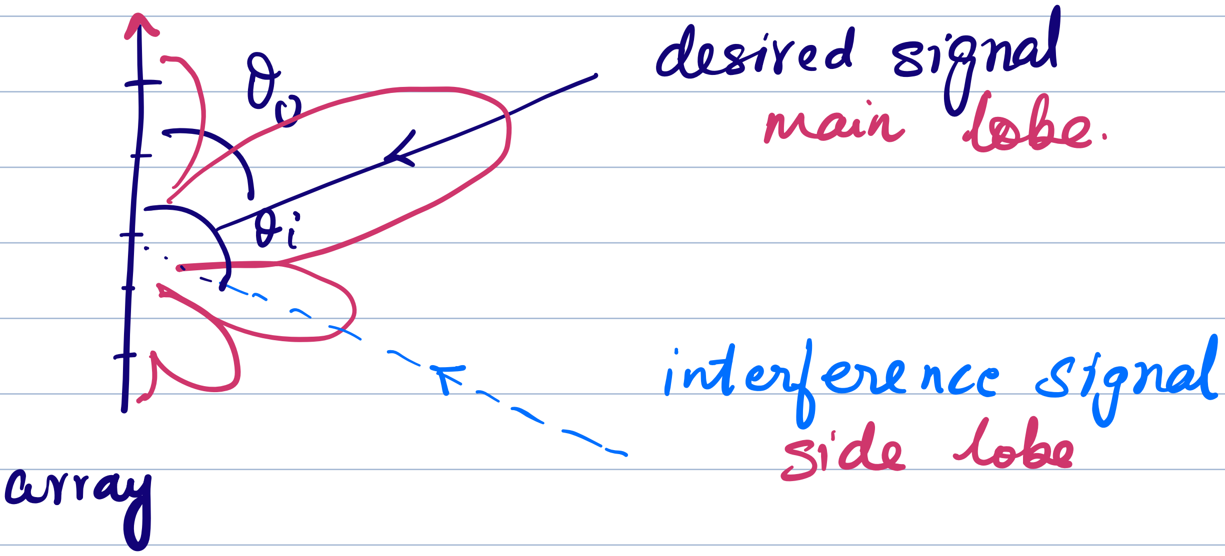
$$\Rightarrow w_{n,opt} = \exp(jkd\cos\theta_0\, n).$$

$\Rightarrow$ excitation coeffs: $a_n = w^*_{n,opt} = \exp(-jn\psi_0)$

i.e. $\beta = -kd\cos\theta_0$.

as before.

② Now a practical situation:



So, the net signal falling on the array will be:

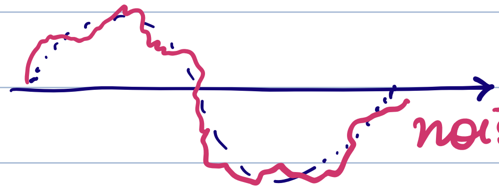
$$x = \underbrace{s(t)\,a(\theta_0)}_{\substack{\text{desired}\\ \text{signal at}\\ \theta_0}} + \underbrace{i(t)\,a(\theta_i)}_{\substack{\text{interference}\\ \text{signal at } \theta_i}}.$$

The received signal:

$$y = w^H x = s(t)\,\underbrace{w^H a(\theta_0)}_{\downarrow\ \text{large}} + i(t)\,\underbrace{w^H a(\theta_i)}_{\downarrow\ \text{not large, but not small}}$$

So, by choosing $a_n = e^{j\beta n}$ we are being naive.
Can we do better?

↳ Before we state the new problem, an aside:
In actual terms we don't simply take $y = w^H x$, because all signals are time varying with noise overriding them. e.g.



noise + signal.

So we don't take just y, but model them as random processes and take the expected values. e.g. if a r.v 'v' takes values $v_1, v_2, \ldots$ with probabilities $p_1, p_2 \ldots$ then $E[v] = \sum p_i v_i$ or $\int v\,p(v)\,dv$.

(discrete rv) (continuous r.v.)

So we do: $P_{out} = E[|y|^2]$ → scalar

Now $E[|y|^2] = E\left[\underbrace{(w^H x)}(w^H x)^H\right]$

The w's are constant, so

$$= E\left[w^H x x^H w\right].$$

$$= w^H E\left[x x^H\right] w. = w^H R w$$

$R = E[x x^H] \in \mathbb{C}^{N \times N}$ is the spatial covariance matrix of the Rx data.

$R_{ij} = E(x_i x_j^*)$, i.e. a correlation between signals at i & j.

R is Hermitian, positive semidefinite.

③ Now the MVDR formulation:

$$\underbrace{\min_{w} \; w^H R w}_{\text{minimize total power (signal + interference)}} \quad \text{subject to} \quad \underbrace{w^H a(\theta_0) = 1.}_{\text{hold unity gain at } \theta_0.}$$

Intuitively → allow desired signal through, minimize interference.

[Our earlier (progressive phase shift) beamformer <u>only</u> did $w^H a(\theta_0) = N$.]

④ Solution via constrained optimization. We use the technique of Lagrange multipliers.

→ Those who know opt, can follow fully.
→ Those who don't, follow like a recipe.

<u>Recap</u>: <u>unconstrained</u>: $\min_x f(x)$

First order stationarity cond: $\nabla f(x^*) = 0$

What is ∇f? If $f = f(x_1, x_2, \ldots, x_N)$ then

$$\nabla f = \begin{pmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_N \end{pmatrix}$$

e.g.s $\nabla(b^T x) = b$

$\nabla(x^T A x) = A + A^T$

Constrained: $\min\limits_{x} f(x)$ s.t. $c(x) = 0$.

Define Lagrangian : $\mathcal{L}(x,\lambda) = f(x) - \lambda c(x)$

1st order Stationarity : $\nabla_x \mathcal{L}(x^*,\lambda^*) = 0$

Here? $\min\limits_{w} w^H R w$ s.t. $w^H a - 1 = 0$.

$\therefore$ Lagrangian $\mathcal{L}(w,\lambda) = w^H R w - \lambda(w^H a - 1)$

<Another aside>: Complex calculus since w is complex. The term $\frac{df}{dx}$ for real x makes sense, but what does $\frac{df}{dz}$ for $z \in \mathbb{C}$ even mean?

The framework is Wirtinger Calculus. Rather than write $f(z)$ as $f(x+jy)$ and separately compute $\frac{\partial f}{\partial x}$ & $\frac{\partial f}{\partial y}$, we define:

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - j\frac{\partial}{\partial y}\right) \text{ and}$$

$$\frac{\partial}{\partial z^*} = \frac{1}{2}\left(\frac{\partial}{\partial x} + j\frac{\partial}{\partial y}\right), \text{ and}$$

treat z, z^* as independent variables just like x, y. i.e. as $\frac{\partial y}{\partial x} = 0$, so also $\frac{\partial z}{\partial z^*} = 0$ and $\frac{\partial z^*}{\partial z} = 0$.

↳ Take an e.g. say $f(z,z^*) = z^* \alpha z$, $\alpha \in \mathbb{R}$.

Then $f = z^* \alpha z = \alpha(x^2+y^2)$.

So $\frac{\partial f}{\partial z^*} = \alpha z$. (simply). OR, using defn,

$$\frac{\partial f}{\partial z^*} = \frac{1}{2}\left(\frac{\partial}{\partial x} + j\frac{\partial}{\partial y}\right)\left(\alpha\left(x^2+y^2\right)\right)$$

$$= \frac{1}{2}\left[2\alpha x + j2\alpha y\right] = \alpha(x+jy) = \alpha z.$$

So both are equivalent. $\partial/\partial z$ or $\partial/\partial z^*$ more convenient.

↳ When f is real, $\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial f}{\partial x} - j\frac{\partial f}{\partial y}\right)$

Take conj → $\left(\frac{\partial f}{\partial z}\right)^* = \frac{1}{2}\left(\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y}\right) = \frac{\partial f}{\partial z^*}$

$\Rightarrow \left(\frac{\partial f}{\partial z}\right)^* = \frac{\partial f}{\partial z^*}$

For first order optimality, use either $\frac{\partial f}{\partial z} = 0$ or $\frac{\partial f}{\partial z^*} = 0$. ⟶ Same!

In the vector case: $\nabla_{w^*} f = 0 = \nabla_w f$

<end aside>

⑤ Soln of our case, $\nabla_{w^*}\Big[\underbrace{w^H R w - \lambda(w^H a - 1)}_{\mathcal{L}(w,\lambda)}\Big] = 0$

Gives: $Rw - \lambda a = 0$

$\Rightarrow \quad w = \lambda R^{-1} a \quad$ — ①

Since $w^H a = 1 \Rightarrow \lambda\left(a^H (R^{-1})^H\right) a = 1$

$\left(\begin{array}{l} R \text{ is Hermitian pos. def.} \\ \text{So is } R^{-1} \end{array}\right)$

$$\Rightarrow \lambda \overbrace{(a^H R^{-1} a)}^{\text{scalar}} = 1$$

$$\Rightarrow \quad w_{MVDR} = \frac{R^{-1} a}{a^H R^{-1} a} = \frac{R^{-1} a(\theta_0)}{a^H(\theta_0) R^{-1} a(\theta_0)}.$$

Corresponding excitation coefs: $a_n = w^*_{MVDR,n}$

A few points:

① The cov matrix $R \triangleq E[x x^H]$. Say that there is one desired signal from θ_0 and K interferers & thermal noise (white) at each receiver, then R is:

$$R = \underbrace{E[S_0 S_0^*]}_{P_S} a(\theta_0) a(\theta_0)^H + \sum_{k=1}^{K} \underbrace{E[S_k S_k^*]}_{P_k} a(\theta_k) a(\theta_k)^H + \sigma^2 I$$

This assumes all sources are uncorrelated.

② Intuitively w_{MVDR} is placing nulls at the interfering directions <u>without</u> explicitly estimating these directions.

③ Output power?

$$P_{MVDR} = w^*_{MVDR} R \, w_{MVDR} = \frac{1}{a^H R^{-1} a}$$

smallest power among all w that satisfies $w^H a(\theta_0) = 1$.

④ Requires R. In practice $\hat{R} = \frac{1}{K} \sum_{k=1}^{K} x_k x_k^H$, i.e. K data snapshots.

⑤ Note: when sources are correlated this method breaks down. Remedy is to go to "spatial smoothening".