Lecture 3: Extremal Solutions

Least squares solutions

Let $A \in \mathbb{C}^{m \times n}$, $\mathbf{b} \in \mathbb{C}^m$, and let $\| \cdot \|$ denote the Euclidean norm. Consider the equations

$$A\,\mathbf{x} = \mathbf{b} \tag{1}$$

A **least–squares solution** (**LSS**) of (1) is any $\mathbf{x}$ in

$$\arg\min \|A\mathbf{x} - \mathbf{b}\| \ .$$

Least squares solutions

Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$, and let $\|\cdot\|$ denote the Euclidean norm. Consider the equations

$$A\,\mathbf{x} = \mathbf{b} \tag{1}$$

A **least–squares solution** (**LSS**) of (1) is any $\mathbf{x}$ in

$$\arg\min \|A\mathbf{x} - \mathbf{b}\| \ .$$

Writing

$$\mathbf{b} = P_{R(A)}\,\mathbf{b} + P_{N(A^*)}\,\mathbf{b}$$

It follows that

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\| = \min_{\mathbf{x}} \|A\mathbf{x} - P_{R(A)}\,\mathbf{b} - P_{N(A^*)}\,\mathbf{b}\| = \|P_{N(A^*)}\,\mathbf{b}\|$$

and the LSS's are the ordinary solutions of

$$A\,\mathbf{x} = P_{R(A)}\,\mathbf{b}$$

Normal equations

The LSS's of
$$A\,\mathbf{x} = \mathbf{b} \tag{1}$$

are the solutions of the **normal equation**

$$A^* A\,\mathbf{x} = A^*\mathbf{b}\ . \tag{2}$$

Proof. $\nabla\,\|A\mathbf{x} - \mathbf{b}\|^2 = 2\,A^*(A\mathbf{x} - \mathbf{b}) = \mathbf{0}$.

Let $A_i \in \mathbb{C}^{m\times n}$, $\mathbf{b}_i \in \mathbb{C}^m$, $(i \in \overline{1,k})$. The minimizers of

$$\sum_{i=1}^{k} \|A_i\mathbf{x} - \mathbf{b}_i\|^2 \tag{3}$$

are the solutions of

$$\left(\sum_{i=1}^{k} A_i^* A_i\right)\mathbf{x} = \sum_{i=1}^{k} A_i^*\mathbf{b}_i\ . \tag{4}$$

$\{1,3\}$–inverses and least squares solutions

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

Theorem. Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. Then $\|A\mathbf{x} - \mathbf{b}\|$ is smallest when $\mathbf{x} = A^{(1,3)}\mathbf{b}$, where $A^{(1,3)} \in A\{1,3\}$.

Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $\|A\mathbf{x} - \mathbf{b}\|$ is smallest when $\mathbf{x} = X\mathbf{b}$, then $X \in A\{1,3\}$.

$\{1,3\}$–inverses and least squares solutions

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

Theorem. Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. Then $\|A\mathbf{x} - \mathbf{b}\|$ is smallest when $\mathbf{x} = A^{(1,3)}\mathbf{b}$, where $A^{(1,3)} \in A\{1,3\}$.

Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $\|A\mathbf{x} - \mathbf{b}\|$ is smallest when $\mathbf{x} = X\mathbf{b}$, then $X \in A\{1,3\}$.

Proof. From

$$\arg\min \|A\mathbf{x} - \mathbf{b}\| = \{\mathbf{x} : A\,\mathbf{x} = P_{R(A)}\mathbf{b}\}$$

and

$$AX = P_{R(A)} \iff X \in A\{1,3\}$$

$\{1,3\}$–inverses and least squares solutions

$$A\mathbf{x} = \mathbf{b} \tag{1}$$

Theorem. Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. Then $\|A\mathbf{x} - \mathbf{b}\|$ is smallest when $\mathbf{x} = A^{(1,3)}\mathbf{b}$, where $A^{(1,3)} \in A\{1,3\}$.

Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $\|A\mathbf{x} - \mathbf{b}\|$ is smallest when $\mathbf{x} = X\mathbf{b}$, then $X \in A\{1,3\}$.

Proof. From

$$\arg\min \|A\mathbf{x} - \mathbf{b}\| = \{\mathbf{x} : A\,\mathbf{x} = P_{R(A)}\mathbf{b}\}$$

and

$$AX = P_{R(A)} \iff X \in A\{1,3\}$$

$$X \in A\{1,3\} \iff \forall\ \mathbf{b} \in \mathbb{C}^m,\ X\mathbf{b} \text{ is a LSS of (1).}$$

$\{1,4\}$–inverses and minimum norm solutions

$$A\mathbf{x} = \mathbf{b}\ ,\ \ \mathbf{b} \in R(A)\ . \tag{1}$$

The **minimum norm solution** (**MNS**) of (1) is

$$\arg\min\{\|\mathbf{x}\| : A\,\mathbf{x} = \mathbf{b}\}\ ,\ \text{the unique solution in } R(A^*)\ .$$

Theorem. Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. If $A\mathbf{x} = \mathbf{b}$ has a solution for $\mathbf{x}$, the minimum norm solution is

$$\mathbf{x} = A^{(1,4)}\mathbf{b}\ ,$$

where $A^{(1,4)} \in A\{1,4\}$. Conversely, if $X \in \mathbb{C}^{n\times m}$ is such that, whenever $A\mathbf{x} = \mathbf{b}$ has a solution, $\mathbf{x} = X\mathbf{b}$ is the MNS, then $X \in A\{1,4\}$.

$\{1,4\}$–inverses and minimum norm solutions

$$A\mathbf{x} = \mathbf{b}\ ,\ \ \mathbf{b} \in R(A)\ . \tag{1}$$

The **minimum norm solution** (**MNS**) of (1) is

$$\arg\min\{\|\mathbf{x}\| : A\,\mathbf{x} = \mathbf{b}\}\ ,\ \text{the unique solution in } R(A^*)\ .$$

Theorem. Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. If $A\mathbf{x} = \mathbf{b}$ has a solution for $\mathbf{x}$, the minimum norm solution is

$$\mathbf{x} = A^{(1,4)}\mathbf{b}\ ,$$

where $A^{(1,4)} \in A\{1,4\}$. Conversely, if $X \in \mathbb{C}^{n\times m}$ is such that, whenever $A\mathbf{x} = \mathbf{b}$ has a solution, $\mathbf{x} = X\mathbf{b}$ is the MNS, then $X \in A\{1,4\}$.

$X \in A\{1,4\} \Longleftrightarrow$ for any $\mathbf{b} \in R(A)$, $X\mathbf{b}$ is the MNS of (1).

Minimum norm least squares solutions (MNLSS)

$$A\mathbf{x} = \mathbf{b} \ . \tag{1}$$

Corollary (**Penrose**). Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. Then $A^\dagger \mathbf{b}$ is the MNLSS of (1). Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $X\mathbf{b}$ is the MNLSS of (1), then $X = A^\dagger$.

Proof. The LSS's of (1) are the solutions of

$$A\mathbf{x} = AA^{(1,3)}\mathbf{b} \ . \tag{2}$$

Therefore the MNLSS of $A\mathbf{x} = \mathbf{b}$ is the MNS of (2),

$$\begin{aligned} \mathbf{x} &= A^{(1,4)}AA^{(1,3)}\mathbf{b} \\ &= A^\dagger \mathbf{b} \end{aligned}$$

Minimum norm least squares solutions (MNLSS)

$$A\mathbf{x} = \mathbf{b}\ . \tag{1}$$

Corollary (**Penrose**). Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$. Then $A^\dagger \mathbf{b}$ is the MNLSS of (1). Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $X\mathbf{b}$ is the MNLSS of (1), then $X = A^\dagger$.

Proof. The LSS's of (1) are the solutions of

$$A\mathbf{x} = AA^{(1,3)}\mathbf{b}\ . \tag{2}$$

Therefore the MNLSS of $A\mathbf{x} = \mathbf{b}$ is the MNS of (2),

$$\begin{aligned} \mathbf{x} &= A^{(1,4)}AA^{(1,3)}\mathbf{b} \\ &= A^\dagger \mathbf{b} \end{aligned}$$

$X = A^\dagger \iff$ for any $\mathbf{b} \in \mathbb{C}^m$, $X\mathbf{b}$ is the MNLSS of (1).

An approximation of the MNLSS

The MNLSS of $A\mathbf{x} = \mathbf{b}$ can be computed in 2 stages

Stage 1: Minimize $\|A\mathbf{x} - \mathbf{b}\|$,

Stage 2: Minimize $\|\mathbf{x}\|$ among all solutions of Stage 1 .

That can be combined,

$$\min_{\mathbf{x}\in\mathbb{C}^n} f_{\alpha^2}(\mathbf{x}) \ , \ \text{where } f_{\alpha^2}(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|^2 + \alpha^2\|\mathbf{x}\|^2 \qquad (2)$$

Let $\mathbf{x}_{\alpha^2}$ be the minimizer of f_{α^2}.

Questions:

(a) Does $\mathbf{x}_{\alpha^2} \to A^\dagger\,\mathbf{b} \quad \text{as} \quad \alpha \to 0$?

(b) What is the error of approximation $\|\mathbf{x}_{\alpha^2} - A^\dagger\,\mathbf{b}\|$?

Before answering these questions, a detour.

A limit for $A^\dagger$

Theorem (den Broeder and Charnes). For any $A \in \mathbb{C}^{m\times n}$, as $\lambda \to 0$ (in $\mathbb{C}$), the following limit exists and

$$\lim_{\lambda\to 0}(A^*A + \lambda I)^{-1}A^* = A^\dagger \tag{1}$$

Proof. We must show that

$$\lim_{\lambda\to 0}(A^*A + \lambda I)^{-1}A^*\mathbf{y} = A^\dagger\mathbf{y}\ ,\ \forall\ \mathbf{y} \in \mathbb{C}^m\ . \tag{2}$$

True for $\mathbf{y} \in N(A^*) = N(A^\dagger)$. Therefore, let $\mathbf{y} \in N(A^*)^\perp = R(A)$, i.e., $\mathbf{y} = A\mathbf{x}$ for unique $\mathbf{x} \in R(A^*)$. We must prove $\forall\ \mathbf{x} \in R(A^*)$

$$\lim_{\lambda\to 0}(A^*A + \lambda I)^{-1}A^*A\mathbf{x} = A^\dagger A\mathbf{x}$$

or equivalently, since $A^\dagger A = P_{R(A^*)}$,

$$\lim_{\lambda\to 0}(A^*A + \lambda I_n)^{-1}A^*A = P_{R(A^*)}\ . \tag{3}$$

Proof (cont'd)

$$\text{To prove:} \quad \lim_{\lambda \to 0} (A^*A + \lambda I_n)^{-1} A^*A = P_{R(A^*)} \ . \tag{3}$$

Let $A^*A = FF^*$, $F \in \mathbb{C}_r^{n \times r}$ be a full–rank factorization. Then

$$(A^*A + \lambda I_n)^{-1} A^*A = (FF^* + \lambda I_n)^{-1} FF^*$$

for any λ for which the inverses exist. Now use the identity

$$(FF^* + \lambda I_n)^{-1} FF^* = F(F^*F + \lambda I_r)^{-1} F^*$$

where

$$\lim_{\lambda \to 0} (F^*F + \lambda I_r)^{-1} = (F^*F)^{-1}$$

since F^*F is nonsingular. Collecting these facts we conclude that

$$\begin{aligned} \lim_{\lambda \to 0} (A^*A + \lambda I_n)^{-1} A^*A &= F(F^*F)^{-1} F^* \\ &= FF^\dagger = P_{R(A^*)} \end{aligned}$$

since the columns of F are a basis for $R(A^*A) = R(A^*)$. $\square$

Alternative proof

It suffices to prove

$$\lim_{\lambda \to 0} (A^* A + \lambda I)^{-1} A^* A \mathbf{x} = \mathbf{x} \ , \ \forall \ \mathbf{x} \in R(A^*) \tag{1}$$

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ be eigenvectors of $A^* A$, with positive eigenvalues,

$$A^* A \mathbf{v}_j = \sigma_j^2 \mathbf{v}_j \ , \quad (\sigma_j > 0 \ , \ j \in \overline{1, r}) \ . \tag{2}$$

that form a basis for $R(A^*)$. Writing $\mathbf{x} \in R(A^*)$ as

$$\mathbf{x} = \sum_{j=1}^{r} \xi_j \mathbf{v}_j \ ,$$

it follows that

$$(A^* A + \lambda I)^{-1} A^* A \mathbf{x} = \sum_{j=1}^{r} \frac{\sigma_j^2 \, \xi_j}{\sigma_j^2 + \lambda} \mathbf{v}_j \ \longrightarrow \ \sum_{j=1}^{r} \xi_j \mathbf{v}_j = \mathbf{x} \ .$$

as $\lambda \longrightarrow 0$, avoiding $\{-\sigma_1^2, -\sigma_2^2, \ldots, -\sigma_r^2\}$. $\square$

Consistent norms

A vector norm $\| \ \|$ and a matrix norm $\| \ \|$ are called **consistent** if

$$\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\| \ , \ \forall \ A \, , \mathbf{x} \tag{1}$$

Given a vector norm $\| \ \|_*$, the **corresponding matrix norm** is

$$\|A\|_* = \sup_{\mathbf{x} \neq \mathbf{0}} \frac{\|A\mathbf{x}\|_*}{\|\mathbf{x}\|_*} \tag{2}$$

a multiplicative matrix norm consistent with $\|\mathbf{x}\|_*$, and for any other matrix norm $\| \ \|$ consistent with $\|\mathbf{x}\|_*$,

$$\|A\| \geq \|A\|_* \ , \ \text{for all } A \ . \tag{1}$$

The matrix norm corresponding to the Euclidean norm $\|\mathbf{x}\|_2$ is the **spectral norm**

$$\|A\|_2 = \max\{\sqrt{\lambda} : \ \lambda \text{ an eigenvalue of } A^* A\}$$

An error bound for $\lim_{\lambda \to 0} (A^*A + \lambda I)^{-1} A^* = A^\dagger$

Theorem (**Boyarintsev**).

$$\|A^\dagger - (A^*A + \lambda I)^{-1} A^*\|_2 \;\leq\; \lambda \, \|A^\dagger\|_2^3 \tag{1}$$

where $\|\cdot\|_2$ is the spectral norm.

Proof.

$$\begin{aligned} A^\dagger - (A^*A + \lambda I)^{-1} A^* &= (A^*A + \lambda I)^{-1}((A^*A + \lambda I)A^\dagger - A^*) \\ &= \lambda \, (A^*A + \lambda I)^{-1} A^\dagger \,, \quad \because \; A^* = A^*AA^\dagger \,, \\ &= \lambda \, (A^*A + \lambda I)^{-1} A^*AA^\dagger A^{\dagger *} A^\dagger \,, \\ &\qquad \because \; A^\dagger = A^\dagger AA^\dagger = A^*A^{\dagger *}A^\dagger = A^*AA^\dagger A^{\dagger *} A^\dagger \,. \end{aligned}$$

If $\lambda > 0$ then $\|(A^*A + \lambda I)^{-1} A^*A\|_2 \leq 1$, and therefore

$$\|A^\dagger - (A^*A + \lambda I)^{-1} A^*\|_2 \leq \lambda \, \|A^\dagger A^{\dagger *} A^\dagger\|_2 \leq \lambda \, \|A^\dagger\|_2^3$$

A limit for the MNLSS

The MNLSS of $A\mathbf{x} = \mathbf{b}$ can be computed in 2 stages

Stage 1: Minimize $\|A\mathbf{x} - \mathbf{b}\|$,

Stage 2: Minimize $\|\mathbf{x}\|$ among all solutions of Stage 1 .

That can be combined,

$$\min_{\mathbf{x}\in\mathbb{C}^n} f_{\alpha^2}(\mathbf{x}) \ , \ \text{where } f_{\alpha^2}(\mathbf{x}) = \|A\mathbf{x} - \mathbf{b}\|^2 + \alpha^2\|\mathbf{x}\|^2 \qquad (2)$$

Let $\mathbf{x}_{\alpha^2}$ be the minimizer of f_{α^2}. Then

$$\mathbf{x}_{\alpha^2} \to A^\dagger\,\mathbf{b} \quad \text{as} \quad \alpha \to 0 \ .$$

Theorem. The function $f_{\alpha^2}(\mathbf{x})$ has a unique minimizer $\mathbf{x}_{\alpha^2}$ given by

$$\mathbf{x}_{\alpha^2} = (A^*A + \alpha^2 I)^{-1} A^*\mathbf{b} \qquad (2)$$

whose norm $\|\mathbf{x}_{\alpha^2}\|$ is a monotone decreasing function of α^2.

Proof

The normal equation of

$$\|A\mathbf{x} - \mathbf{b}\|^2 + \alpha^2\|\mathbf{x}\|^2$$

is

$$(A^*A + \alpha^2 I)\mathbf{x} = A^*\mathbf{b} \ ,$$

with solution

$$\begin{aligned}\mathbf{x}_{\alpha^2} &= (A^*A + \alpha^2 I)^{-1}A^*\mathbf{b} \\ &= (A^*A + \alpha^2 I)^{-1}A^*A\mathbf{v}\end{aligned}$$

where

$$\mathbf{b} = A\mathbf{v} + \mathbf{u} \ , \quad \mathbf{v} \in R(A^*) \ , \quad \mathbf{u} \in N(A) \ .$$

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ be an o.n. basis of $R(A^*)$,

$$A^*A\mathbf{v}_j = \sigma_j^2\mathbf{v}_j \quad (\sigma_j > 0, \, j \in \overline{1, r}) \ .$$

eigenvectors of A^*A corresponding to nonzero eigenvalues.

Proof (cont'd)

Write $\mathbf{v} \in R(A^*)$ as

$$\mathbf{v} = \sum_{j=1}^{r} \beta_j \, \mathbf{v}_j$$

Then

$$\begin{aligned}
\mathbf{x}_{\alpha^2} &= (A^*A + \alpha^2 I)^{-1} A^* A \mathbf{v} \\
&= \sum_{j=1}^{r} \frac{\sigma_j^2 \, \beta_j}{\sigma_j^2 + \alpha^2} \, \mathbf{v}_j \\
\therefore \; \|\mathbf{x}_{\alpha^2}\|^2 &= \sum_{j=1}^{r} \left(\frac{\sigma_j^2}{\sigma_j^2 + \alpha^2} \right)^2 |\beta_j|^2 \, ,
\end{aligned}$$

a monotone decreasing function of α^2. $\square$

The Frobenius norm

In the space $\mathbb{C}^{m\times n}$ define inner product of $X = (x_{ij})$, $Y = (y_{ij})$,

$$\langle X, Y\rangle_F = \sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}\overline{y_{ij}} = \operatorname{trace} Y^* X \tag{1}$$

and the associated **Frobenius norm**

$$\|X\|_F = \langle X, X\rangle_F^{1/2} = \sum_{i=1}^{m}\sum_{j=1}^{n} |x_{ij}|^2 = (\operatorname{trace} X^* X)^{1/2}$$

The Frobenius norm of X is the Euclidean norm of $\operatorname{vec} X$.

The Frobenius norm is unitarily invariant

$$\|UAV\|_F = \|A\|_F \ , \text{ for any unitary matrices } U, V$$

Matrix equations

For any $A, B \in \mathbb{C}^{m\times n}$ define:

range of (A, B): $\quad R(A, B) := \{Y = AXB \in \mathbb{C}^{m\times n} : X \in \mathbb{C}^{n\times m}\}$,

null space of (A, B): $\quad N(A, B) = \{X \in \mathbb{C}^{n\times m} : AXB = O\}$.

Let $\mathbb{C}^{m\times n}$ have the inner product $\langle X, Y\rangle_F = \operatorname{trace} Y^* X$. Then

$$\mathbb{C}^{m\times n} = R(A, B) \overset{\perp}{\oplus} N(A^*, B^*) \ , \ \forall\ A, B \in \mathbb{C}^{m\times n} \ .$$

Matrix equations

For any $A, B \in \mathbb{C}^{m\times n}$ define:

range of (A, B): $\quad R(A, B) := \{Y = AXB \in \mathbb{C}^{m\times n} : X \in \mathbb{C}^{n\times m}\}$,

null space of (A, B): $\quad N(A, B) = \{X \in \mathbb{C}^{n\times m} : AXB = O\}$.

Let $\mathbb{C}^{m\times n}$ have the inner product $\langle X, Y\rangle_F = \operatorname{trace} Y^* X$. Then

$$\mathbb{C}^{m\times n} = R(A, B) \overset{\perp}{\oplus} N(A^*, B^*) \ , \ \forall\ A, B \in \mathbb{C}^{m\times n} \ .$$

Let $A \in \mathbb{C}^{m\times n}$ and let X be a $\{1\}$–inverse of A. Then the following are equivalent:

(a) $X = A^\dagger$,

(b) $X \in R(A^*, A^*)$,

(c) X is the minimum–norm solution of $AXA = A$.

Weighted inner products and norms

The Q–**inner product** in $\mathbb{C}^n$ induced by a PD matrix $Q \in \mathbb{C}^{n\times n}$ is

$$\langle \mathbf{x}, \mathbf{y}\rangle_Q := \langle \mathbf{x}, Q\mathbf{y}\rangle$$

and the corresponding Q–**norm**

$$\|\mathbf{x}\|_Q := \langle \mathbf{x}, \mathbf{x}\rangle_Q^{1/2}$$

The transformation

$$\widetilde{\mathbf{x}} := Q^{1/2}\mathbf{x}$$

brings back the standard inner product and Euclidean norm

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y}\rangle_Q &= \langle \mathbf{x}, Q\mathbf{y}\rangle = \langle Q^{1/2}\mathbf{x}, Q^{1/2}\mathbf{y}\rangle = \langle \widetilde{\mathbf{x}}, \widetilde{\mathbf{y}}\rangle \\ \|\mathbf{x}\|_Q &= \|\widetilde{\mathbf{x}}\|\end{aligned}$$

Weighted least squares

Let W, Q be PD matrices, and consider the equation

$$A\,\mathbf{x} = \mathbf{b} \tag{1}$$

A **weighted LSS** of (1) is a minimizer of

$$\|A\mathbf{x} - \mathbf{b}\|_W^2 = (A\mathbf{x} - \mathbf{b})^* W (A\mathbf{x} - \mathbf{b}) \ , \tag{2}$$

and a **weighted MNS** of (1), if consistent, is the solution $\mathbf{x}$ minimizing $\|\mathbf{x}\|_Q = (\mathbf{x}^* Q\,\mathbf{x})^{1/2}$.

The change of basis,

$$\widetilde{\mathbf{x}} := Q^{1/2}\mathbf{x} \ , \ \widetilde{A} := W^{1/2}\,A\,Q^{-1/2} \ , \ \widetilde{\mathbf{b}} := W^{1/2}\,\mathbf{b} \tag{3}$$

gives back

$$\|A\mathbf{x} - \mathbf{b}\|_W = \|\widetilde{A}\widetilde{\mathbf{x}} - \widetilde{\mathbf{b}}\| \ , \ \|\mathbf{x}\|_Q = \|\widetilde{\mathbf{x}}\| \tag{4}$$

The transformation

$$\widetilde{\mathbf{x}} := Q^{1/2}\mathbf{x}\ ,\ \ \widetilde{A} := W^{1/2}\,A\,Q^{-1/2}\ ,\ \ \widetilde{\mathbf{b}} := W^{1/2}\,\mathbf{b}$$

results in the correspondence

$$\|A\,\mathbf{x}-\mathbf{b}\|_W = (A\,\mathbf{x}-\mathbf{b})^* W (A\,\mathbf{x}-\mathbf{b})^{1/2} = \|\widetilde{A}\,\widetilde{\mathbf{x}}-\widetilde{\mathbf{b}}\|$$

$$\|\mathbf{x}\|_Q = (\mathbf{x}^* Q\,\mathbf{x})^{1/2} = \|\widetilde{\mathbf{x}}\|$$

Define

$$\widetilde{X} := Q^{1/2} X W^{-1/2}$$

$$\therefore\ \widetilde{X} \in \widetilde{A}\{1\} \Longleftrightarrow X \in A\{1\}\ , \qquad \text{(a)}$$

$$\widetilde{X} \in \widetilde{A}\{2\} \Longleftrightarrow X \in A\{2\}\ , \qquad \text{(b)}$$

$$\widetilde{X} \in \widetilde{A}\{3\} \Longleftrightarrow\ WAX = (WAX)^*\ , \qquad \text{(c)}$$

$$\widetilde{X} \in \widetilde{A}\{4\} \Longleftrightarrow\ QXA = (QXA)^*\ . \qquad \text{(d)}$$

Proof of (a): $\widetilde{A}\widetilde{X}\widetilde{A} = W^{1/2}AXAQ^{-1/2} = \widetilde{A} = W^{1/2}AQ^{-1/2}$

Inverses for weighted LSS's and MNS's

Theorem. Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$, and let $W \in \mathbb{C}^{m\times m}$ be PD. Then $\|A\mathbf{x} - \mathbf{b}\|_W$ is smallest when $\mathbf{x} = X\,\mathbf{b}$, where X satisfies

$$AXA = A\,, \tag{1}$$

$$(WAX)^* = WAX\,. \tag{3.W}$$

Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $\|A\mathbf{x} - \mathbf{b}\|_W$ is smallest when $\mathbf{x} = X\,\mathbf{b}$, then X satisfies (1)–(3.W).

Theorem. Let $Q \in \mathbb{C}^{n\times n}$ be PD. If $A\mathbf{x} = \mathbf{b}$ has a solution for $\mathbf{x}$, the unique solution minimizing $\|\mathbf{x}\|_Q$ is $\mathbf{x} = X\mathbf{b}$, where X satisfies

$$AXA = A\,, \tag{1}$$

$$(QXA)^* = QXA\,. \tag{4.Q}$$

Conversely, if $X \in \mathbb{C}^{n\times m}$ is such that, whenever $\mathbf{b} \in R(A)$, $\mathbf{x} = X\mathbf{b}$ is the minimizing $\|\mathbf{x}\|_Q$, then X satisfies (1)–(4.Q).

$A^{(1,2)}_{(W,Q)}$ and weighted MNLSS's

Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$, and let $W \in \mathbb{C}^{m\times m}$, $Q \in \mathbb{C}^{n\times n}$ be PD. Then, there is a unique matrix

$$X = A^{(1,2)}_{(W,Q)} \in A\{1,2\}$$

satisfying

$$(WAX)^* = WAX\ , \quad (QXA)^* = QXA\ .$$

Moreover, $\|A\mathbf{x} - \mathbf{b}\|_W$ is minimized for $\mathbf{x} = X\mathbf{b}$, and among such vectors, $\mathbf{x} = X\mathbf{b}$ is the one for which $\|\mathbf{x}\|_Q$ is minimum.

$A^{(1,2)}_{(W,Q)}$ and weighted MNLSS's

Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$, and let $W \in \mathbb{C}^{m\times m}$, $Q \in \mathbb{C}^{n\times n}$ be PD. Then, there is a unique matrix

$$X = A^{(1,2)}_{(W,Q)} \in A\{1,2\}$$

satisfying

$$(WAX)^* = WAX\ , \quad (QXA)^* = QXA\ .$$

Moreover, $\|A\mathbf{x} - \mathbf{b}\|_W$ is minimized for $\mathbf{x} = X\mathbf{b}$, and among such vectors, $\mathbf{x} = X\mathbf{b}$ is the one for which $\|\mathbf{x}\|_Q$ is minimum.

Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $\mathbf{x} = X\mathbf{b}$ is the vector of $\mathbb{C}^m$ for which $\|\mathbf{x}\|_Q$ is smallest among those for which $\|A\mathbf{x} - \mathbf{b}\|_W$ assumes its minimum value, then $X = A^{(1,2)}_{(W,Q)}$.

$A^{(1,2)}_{(W,Q)}$ and weighted MNLSS's

Let $A \in \mathbb{C}^{m\times n}$, $\mathbf{b} \in \mathbb{C}^m$, and let $W \in \mathbb{C}^{m\times m}$, $Q \in \mathbb{C}^{n\times n}$ be PD. Then, there is a unique matrix

$$X = A^{(1,2)}_{(W,Q)} \in A\{1,2\}$$

satisfying

$$(WAX)^* = WAX\ , \quad (QXA)^* = QXA\ .$$

Moreover, $\|A\mathbf{x} - \mathbf{b}\|_W$ is minimized for $\mathbf{x} = X\mathbf{b}$, and among such vectors, $\mathbf{x} = X\mathbf{b}$ is the one for which $\|\mathbf{x}\|_Q$ is minimum.

Conversely, if $X \in \mathbb{C}^{n\times m}$ has the property that, for all $\mathbf{b}$, $\mathbf{x} = X\mathbf{b}$ is the vector of $\mathbb{C}^m$ for which $\|\mathbf{x}\|_Q$ is smallest among those for which $\|A\mathbf{x} - \mathbf{b}\|_W$ assumes its minimum value, then $X = A^{(1,2)}_{(W,Q)}$.

For $W = I_m, Q = I_n$, $A^{(1,2)}_{(W,Q)} = A^\dagger$

Application of the Bott–Duffin inverse

A resistance network satisfies **Ohm's law**,

$$A\mathbf{x} + \mathbf{y} = A\mathbf{v} + \mathbf{w}\ ,\ \ \mathbf{x} \in R(M^T),\ \mathbf{y} \in N(M) \tag{1}$$

and the **Kelvin minimum principle**: The **rate of energy loss**

$$q(\mathbf{x}) = \tfrac{1}{2}\,(\mathbf{x} - \mathbf{v})^* A(\mathbf{x} - \mathbf{v}) - \mathbf{w}^*\mathbf{x} \tag{2}$$

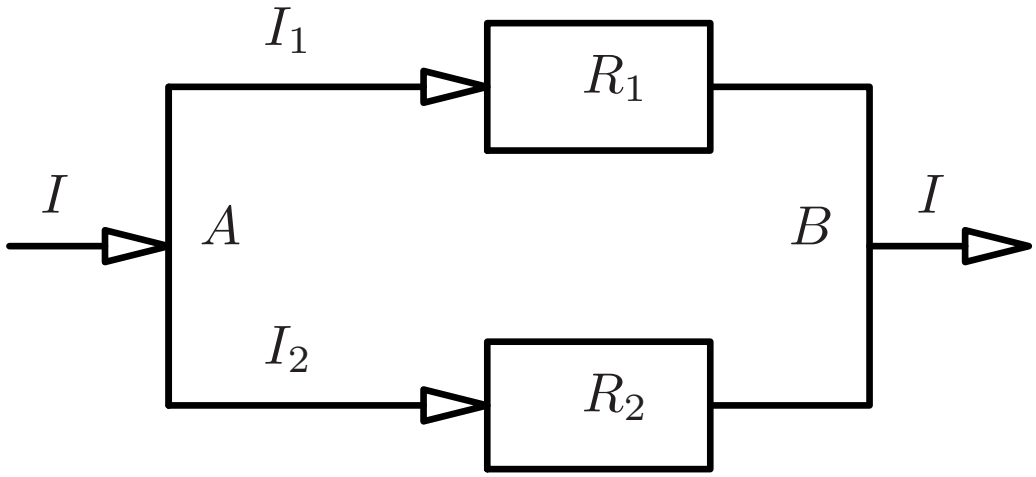
is **minimized**.

Theorem (**Bott and Duffin**). Let $A \in \mathbb{C}^{n\times n}$ be **Hermitian**, L a subspace of $\mathbb{C}^n$ and $A^{(-1)}_{(L)}$ exists. Then for any $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$, the quadratic function (2) has a **unique stationary value** in L, when

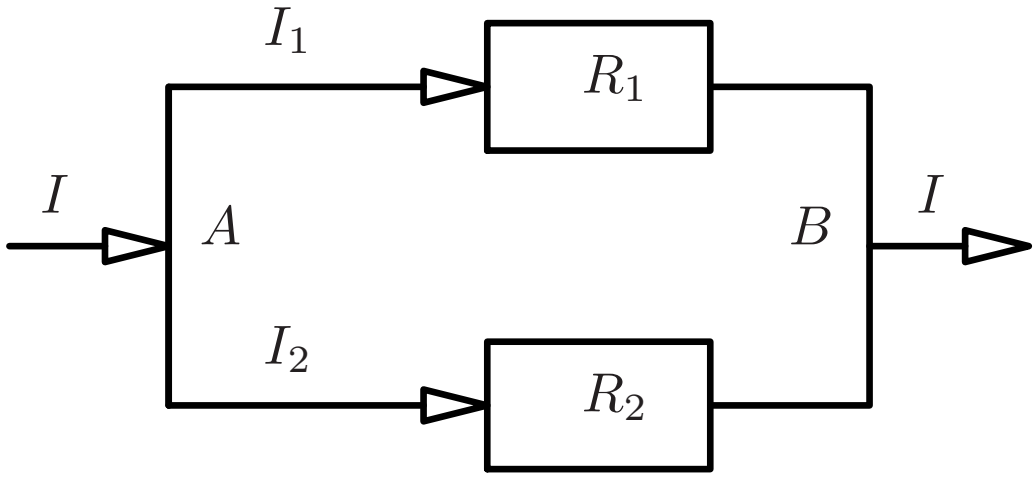
$$\mathbf{x} = A^{(-1)}_{(L)}(A\mathbf{v} + \mathbf{w})\ . \tag{3}$$

Conversely, if A and L are such that for any $\mathbf{v}, \mathbf{w} \in \mathbb{C}^n$, the quadratic function (2) has a stationary value in L, then $A^{(-1)}_{(L)}$ exists and the stationary point is unique for any $\mathbf{v}$, $\mathbf{w}$ and given by (3).

Resistances in parallel



Resistances in parallel



Given I, R_1, R_2, the currents I_1, I_2 solve:

$$\min\{I_1^2 R_1 + I_2^2 R_2 : I_1 + I_2 = I\} = I^2\, R$$

$$R = \frac{1}{1/R_1 + 1/R_2} = \frac{R_1 R_2}{R_1 + R_2}$$

$$\therefore\ I_1 R_1 = I_2 R_2 = I R\,.$$

Parallel sums (the work of Anderson & Duffin)

Let A, B be PSD. Their **parallel sum** is

$$A : B := A(A+B)^{\dagger} B \tag{1}$$

$$= (A^{-1} + B^{-1})^{-1} \ , \ \text{if } A, B \text{ are invertible.} \tag{2}$$

Theorem (**Duffin**, **Anderson**, **Morley**). If A, B, C, D are PSD, then:

(a) $A : B$ is PSD

(b) $A : B = B : A$

(c) $A : B \ \leq \ A$

(c) for all $\mathbf{z}$,

$$\langle \mathbf{z}, A : B\mathbf{z} \rangle = \inf_{\mathbf{x},\mathbf{y}} \left\{ \langle \mathbf{x}, A\mathbf{x} \rangle + \langle \mathbf{y}, B\mathbf{y} \rangle : \ \mathbf{x} + \mathbf{y} = \mathbf{z} \right\} .$$

(d) $(A+B) : (C+D) \ \geq \ A : C + B : D$

What can you do with generalized inverses, that you could not do without?

Theorem (**von Neumann**). If P_L, P_M are the orthogonal projectors on the subspaces L, M, then

$$P_{L\cap M} = \lim_{n\to\infty} (P_L P_M)^n \ . \tag{1}$$

What can you do with generalized inverses, that you could not do without?

Theorem (**von Neumann**). If P_L, P_M are the orthogonal projectors on the subspaces L, M, then

$$P_{L\cap M} = \lim_{n\to\infty} (P_L P_M)^n \ . \tag{1}$$

Theorem (**Anderson and Duffin**).

$$P_{L\cap M} = 2\, P_L (P_L + P_M)^\dagger P_M \ . \tag{2}$$

Proof. Let $H = 2P_L(P_L + P_M)^\dagger P_M = 2P_M(P_L + P_M)^\dagger P_L$.

$$\begin{aligned}
\therefore\ H &= P_{L\cap M} H = P_{L\cap M}\left(P_L(P_L + P_M)^\dagger P_M + P_M(P_L + P_M)^\dagger P_L\right) \\
&= P_{L\cap M}(P_L + P_M)^\dagger (P_L + P_M) \\
&= P_{L\cap M} P_{L+M} \\
&= P_{L\cap M} \ . \qquad \square
\end{aligned}$$