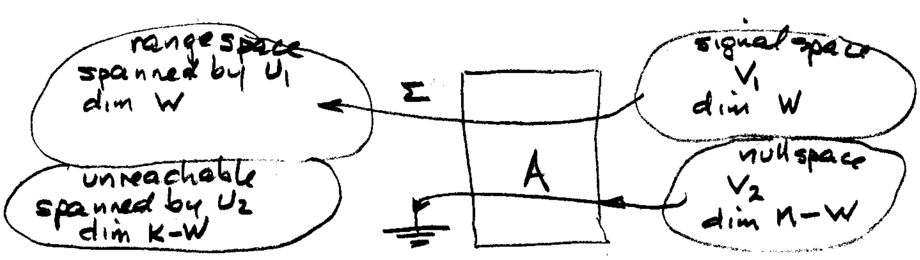
7.8 The Pseudo inverse

- If A is square and full rank (i.e. non singular) we would solve $\underline{d} = A\underline{w}$ by $\underline{w} = A^{-1}\underline{d}$

- In the overdetermined case we must settle for $\underline{d} \approx A\underline{w}$ and minimize $\|\underline{d} - \hat{\underline{d}}\|^2$ (sum of squares). The solution requirement is
$$A^+A\underline{w} = A^+\underline{d}$$
and for full rank A (LI columns) A^+A is invertible:
$$\underline{w} = (A^+A)^{-1}A^+\underline{d}$$
Many texts refer to $A^\# = (A^+A)^{-1}A^+$ as the pseudo inverse of A

- That formulation won't work if cols of A are LD - and they might be, since A is constructed from measurements. The cols are certainly LD in the underdet case. What to do?

- Generalize pseudo inv to
$$A^\# = V\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}U^+ = \begin{bmatrix}V_1 & V_2\end{bmatrix}\begin{bmatrix}\Sigma^{-1} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}U_1^+\\ U_2^+\end{bmatrix} = V_1\Sigma^{-1}U_1^+$$
$$= \sum_{i=1}^{W}\frac{1}{\sigma_i}\underline{v}_i\underline{u}_i^+$$
This simply works backwards on our system



$$A^\#\underline{d} = \sum_{i=1}^{W}\frac{1}{\sigma_i}\underline{v}_i\underbrace{\underline{u}_i^+\underline{d}}_{\delta_i}$$ first projects $\underline{d}$ onto the reachable part of $\mathcal{E}^K$, the range space

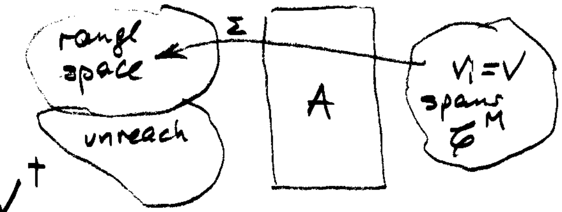
$$= \sum_{i=1}^{W}\frac{\delta_i}{\sigma_i}\underline{v}_i$$ then compensates for the forward gain σ_i

$$= \sum_{i=1}^{W}\omega_i\underline{v}_i = \underline{w}$$ to express $\underline{w}$ using the basis V_1 of the signal space

- Special case: overdet, full (col) rank $W = M$

Check to see if it's $(A^+A)^{-1}A^+$, as ~~before~~.

Null space $= \phi \qquad V_1 = V, \quad M \times M$



we have $A = U_1 \Sigma V_1^+ = U_1 \Sigma V^+$

$$A^+A = V \Sigma U_1^+ U_1 \Sigma V^+ = V \Sigma^2 V^+$$

so $$(A^+A)^{-1} = V \Sigma^{-2} V^+$$

and $$A^+ = V \Sigma U_1^+$$

so $$(A^+A)^{-1}A^+ = V \Sigma^{-2} V^+ V \Sigma U_1^+ = V \Sigma^{-1} U_1^+ = V_1 \Sigma^{-1} U_1^+$$

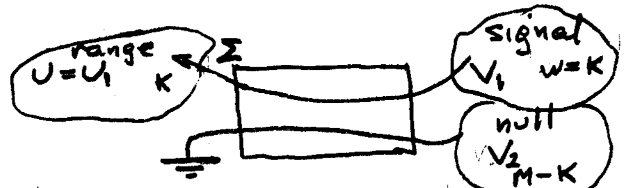
$$= V \left[\Sigma^{-1} \; 0 \right] U^+$$ just like our generalized definition of $A^\#$

(note error in Haykin)

So $A^\# = (A^+A)^{-1}A^+$ if overdet, full rank

- Special case: underdet, full (row) rank $W = K$.

Then $U_1 = U \quad K \times K$, unreachable $= \phi$



We haven't derived it, but check

$$A^+(AA^+)^{-1}$$

$$A = U_1 \Sigma V_1^+ = U \Sigma V_1^+, \quad A^+ = V_1 \Sigma U^+$$

$$AA^+ = U \Sigma V_1^+ V_1 \Sigma U^+ = U \Sigma^2 U^+$$

so $$(AA^+)^{-1} = U \Sigma^{-2} U^+$$

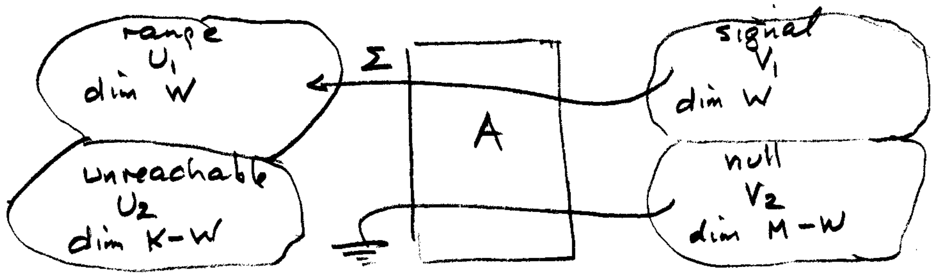
and $$A^+(AA^+)^{-1} = V_1 \Sigma U^+ U \Sigma^{-2} U^+ = V_1 \Sigma^{-1} U^+ = V_1 \Sigma^{-1} U_1^+$$

which also agrees with general def'n $= V \begin{bmatrix} \Sigma^{-1} \\ 0 \end{bmatrix} U^+$ (error in Haykin)

so $A^\# = A^+(AA^+)^{-1}$ for underdet, full row rank

- Complete solution and minimum norm property.

– In general, for LD cols of A (underdet and/or rank defic)



The pseudo inverse $A^{\#} = V_1 \Sigma^{-1} U_1^{+}$ gives us a specific solution

$$\underline{\hat{w}} = A^{\#}\underline{d} \qquad \underline{\hat{d}} = A\,\underline{\hat{w}} = A A^{\#}\underline{d}$$

We can add on any vector $\underline{w}_h$ in the null space and the sum is a solution

$$\underline{w} = \underline{\hat{w}} + \underline{w}_h = \underline{\hat{w}} + \sum_{i=W+1}^{M} \omega_i \underline{v}_i = \underline{\hat{w}} + V_2 \underline{\omega}_h$$

since $\quad \underline{\hat{d}} = A\,\underline{w} = A(\underline{\hat{w}} + \underline{w}_h) = A\underline{\hat{w}} + \underline{0} = A\,\underline{\hat{w}}$

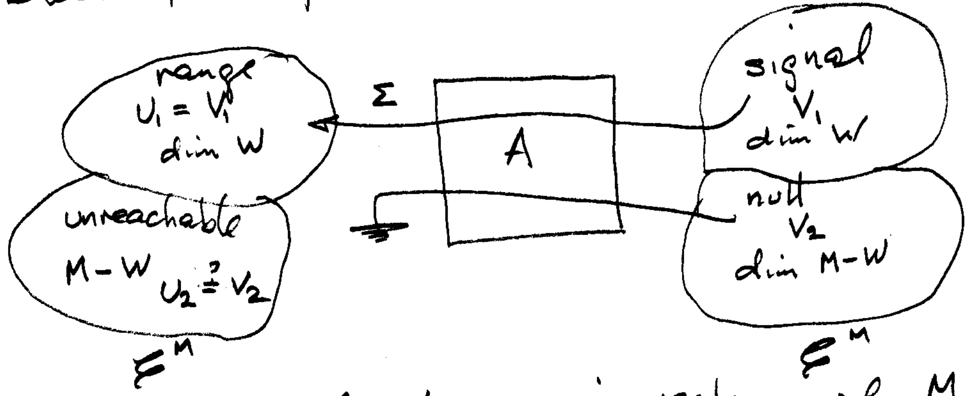
So we have the full set of solutions given by $\underline{\hat{w}}$ + nullspace, dimension M−W, expressed parametrically above $(V_2\,\underline{\omega}_h)$.

– Of all those solutions, $\underline{\hat{w}}$ has minimum norm: since all vectors in the null space are orthog to $\underline{\hat{w}}$, adding any vector from that space increases the length of $\underline{w}$

– $\underline{\hat{w}} = A^{\#}\underline{d}$ is the only vector that both
  – minimizes the sum of error squares $\|\underline{d} - \underline{\hat{d}}\|^2$ in the output space
  – has the minimum norm in the input space

- So far, we have focussed on rectangular matrices. Let's go back to square matrices and establish some of the links.

First, $K = M$ so it is possible to have the same basis for input and output spaces



The columns of V are eigenvectors of $M \times M$ A, and σ_i, $i = 1..W$ are the non zero eigenvals; remaining $\sigma_i = 0$, $i = W+1..M$.

$$AV = A\begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \underbrace{\begin{array}{c|cc} & W & M-W \\ \hline \end{array}\left[\begin{array}{c|c} \Sigma & 0 \\ \hline 0 & 0 \end{array}\right]\begin{array}{l} W \\ M-W \end{array}}_{S}$$

So $U_1 = V_1$, images of eigenvectors are just the same eigenvecs scaled by Σ

V_2 spans null space
unreachable space spanned by vectors orthog to range space — so might as well use $U_2 = V_2$
and treat 0 eigenvals same way as positive ones

so $U = V$, $A = USV^+ = V\left[\begin{array}{c|c} \Sigma & 0 \\ \hline 0 & 0 \end{array}\right]V^+$

- This analysis lets us understand general sets of linear equations.
  - Suppose overdetermined and we want <u>equality</u>

    $$\underline{d} = A\,\underline{w}$$

    No solution unless $\underline{d}$ is in range space, $U_2^T \underline{d} = \underline{0}$
    that is, in the column space of A.
    If it is, then $\underline{w} = A^{\#}\underline{d}$ produces exact solution, unique.
  - If overdetermined, and we <u>relax equality</u> to minimum sum of error squares, then

    $$\hat{\underline{w}} = A^{\#}\underline{d}$$

    both cuts out any part of $\underline{d}$ lying in the unreachable part of output space and finds the $\underline{w}$ that corresponds to what's left of $\underline{d}$, in the range space. unique.
  - If overdet, LD cols of A, then
    - if we want equality, then: no solution if $\underline{d}$ not in col space, infinite number (dim M−W) if it is, and $A^{\#}\underline{d}$ gives min norm sol'n.
    - if min sum of squares OK, then: always inf number (dim M−W) of solutions, $A^{\#}\underline{d}$ gives one with min norm
  - If underdet, then always inf # solutions (dim K−W), with or without full row rank, $A^{\#}\underline{d}$ gives one with min norm

- Check these properties:

$$A A^{\#} = P_{U_1} = P_A \qquad A^{\#}A = P_{V_1} \ (= I \text{ if } W=M)$$

$$A A^{\#} A = A \qquad A^{\#} A A^{\#} = A^{\#}$$