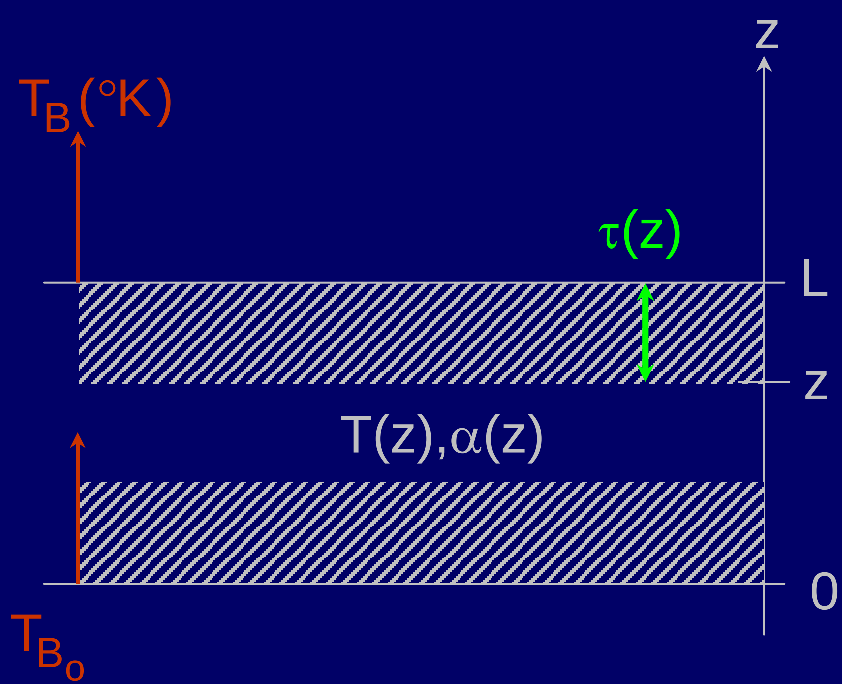
Remote Sensing

Remote sensing is a quasi-linear estimation problem

Equation of radiative transfer:



$$T_B(°K) = T_{B_o} e^{-\tau_o} + \int_0^L T(z)\alpha(z)e^{-\tau(z)}dz$$

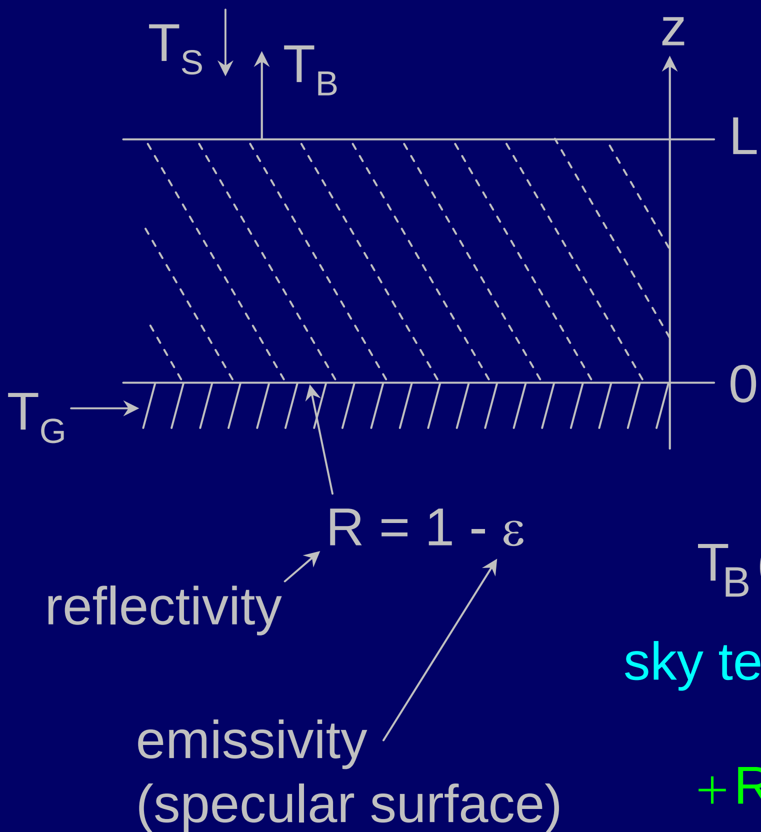
temperature profile

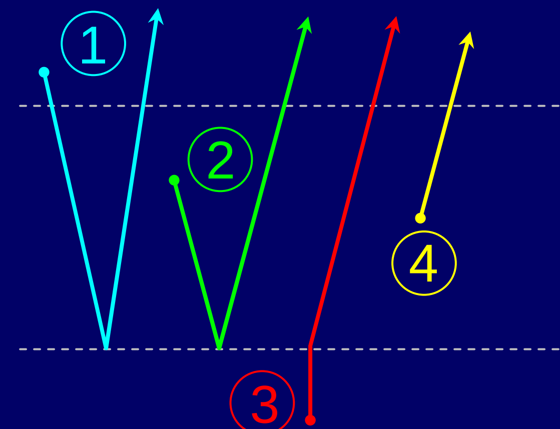
nepers m^{-1}

$$\tau(z) \stackrel{\Delta}{=} \int_z^L \alpha(z)dz$$

$$\tau_o \stackrel{\Delta}{=} \tau(0)$$

Radiation from Sky-Illuminated Reflective Surfaces





terms

$$T_B(^\circ K) = RT_s e^{-2\tau_o} + \varepsilon T_G e^{-\tau_o} + \qquad 1 + 3$$

sky temperature — ground temperature

$$+ R e^{-\tau_o} \int_0^L T(z)\alpha(z) e^{-\int_0^z \alpha(z)dz} dz \qquad 2$$

$$+ \int_0^L T(z)\alpha(z) e^{-\int_z^L \alpha(z)dz} dz \qquad 4$$

$$\cong \int_0^L T(z)\left[\alpha(z) e^{-\tau(z)}\right] dz \quad \text{for } \tau_o >> 1 \qquad 4$$

Temperature Weighting Function W(z,f,T)

Terms:

$$\text{For } \tau_o >> 1: \quad T_B(f) \cong \underbrace{T_o}_{1,\,2,\,3} + \int_0^L T(z) \underbrace{W(z,f,T(z))}_{4}\, dz$$

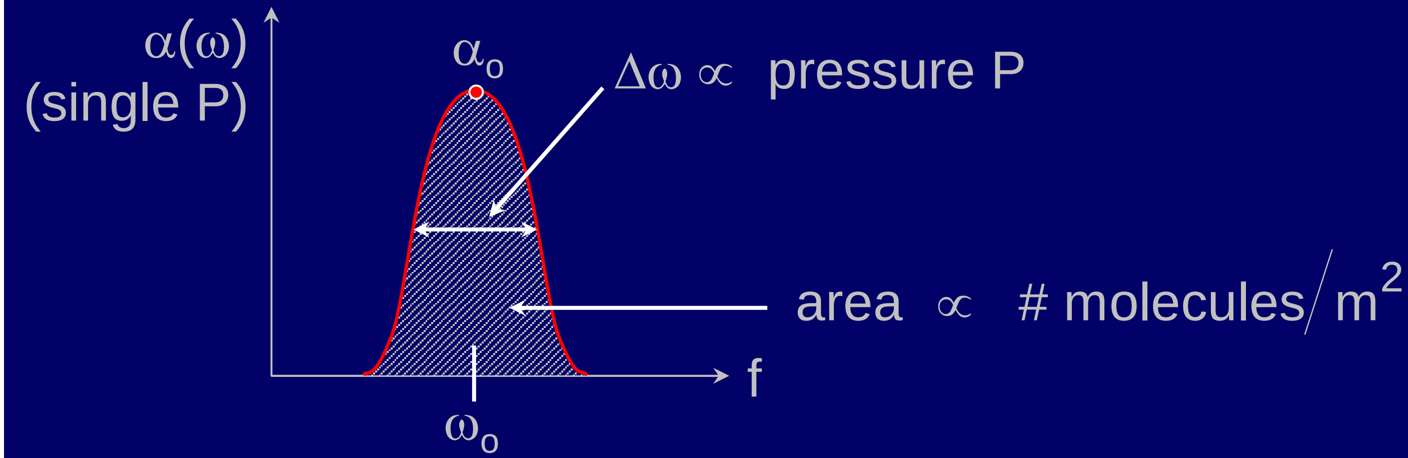
$$\text{Alternatively,} \quad T_B(f) \cong T_o' + \int_0^L (T(z) - T_o(z)) \underbrace{W'(z,f,T_o(z))}\, dz$$

$$\text{Incremental weighting function:} \quad W'(z,f,T_o(z)) = \left.\frac{\partial T_B}{\partial T(z)}\right|_{T_o(z)}$$

$$\text{Note: we have \textasciitilde linear relation:} \quad T_B(f) \overset{\sim}{\leftrightarrow} T(z)$$

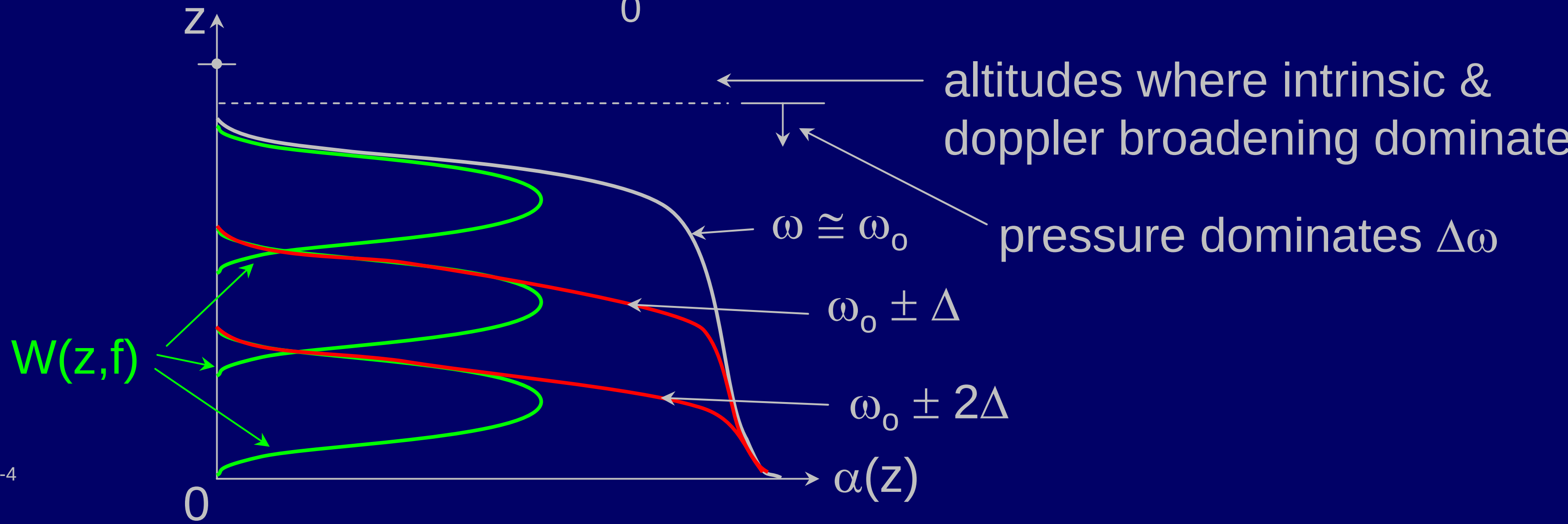
(not Fourier)

Atmospheric Temperature T(z) Retrievals from Space

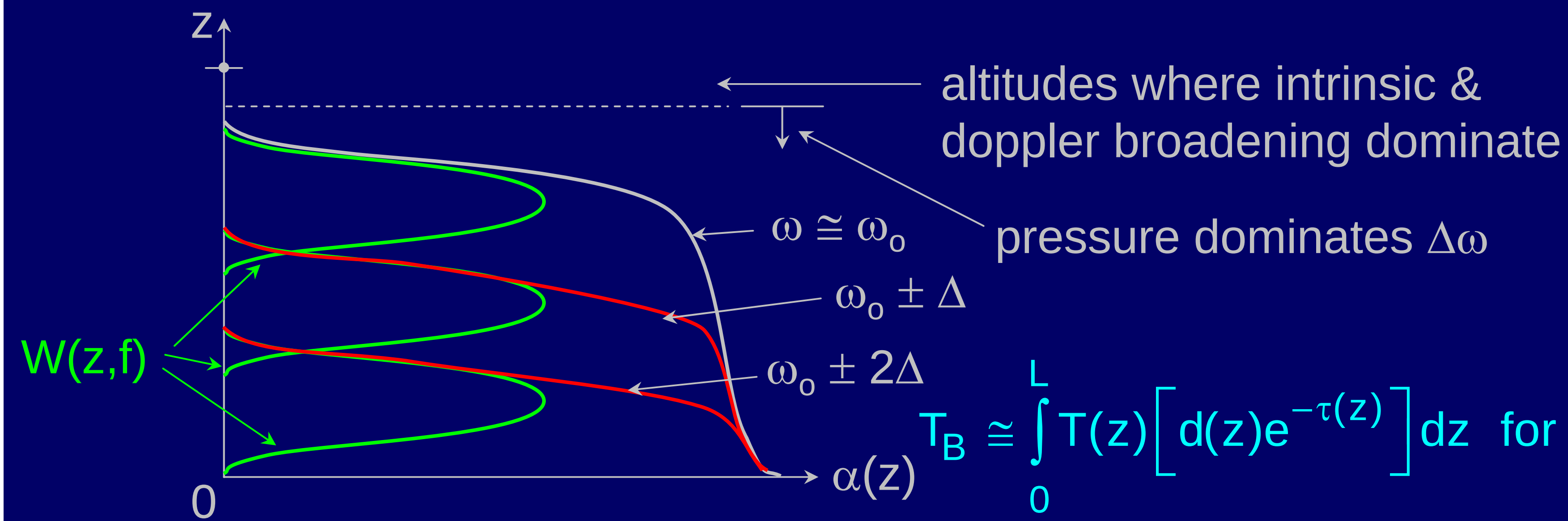


Therefore $\alpha_o \ncong f(P)$ for P-broadening trace constituents

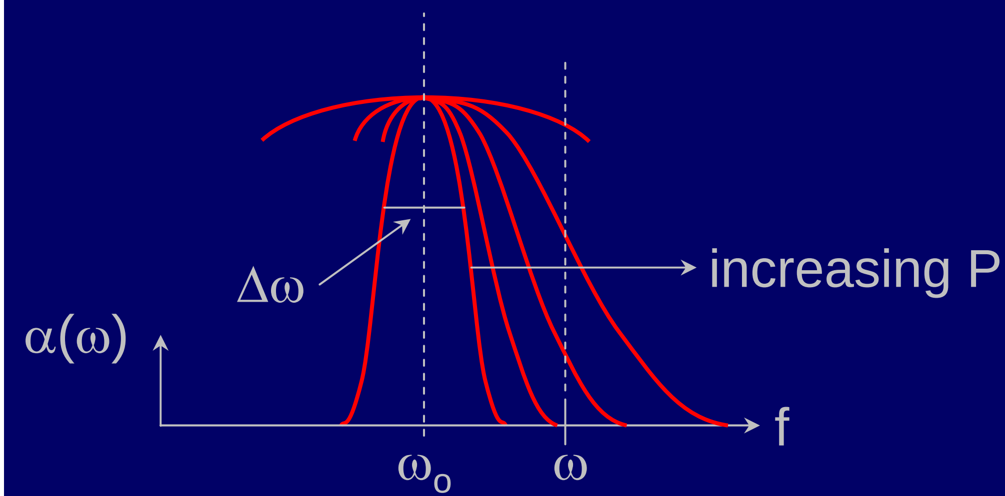
$$T_B \cong \int_0^L T(z)\left[\alpha(z)e^{-\tau(z)}\right]dz \quad \text{for } \tau_o >> 1$$

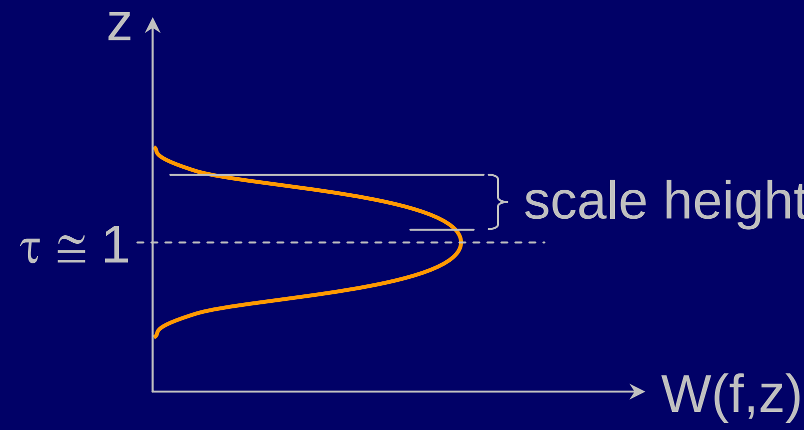


Atmospheric Temperature T(z) Retrievals from Space

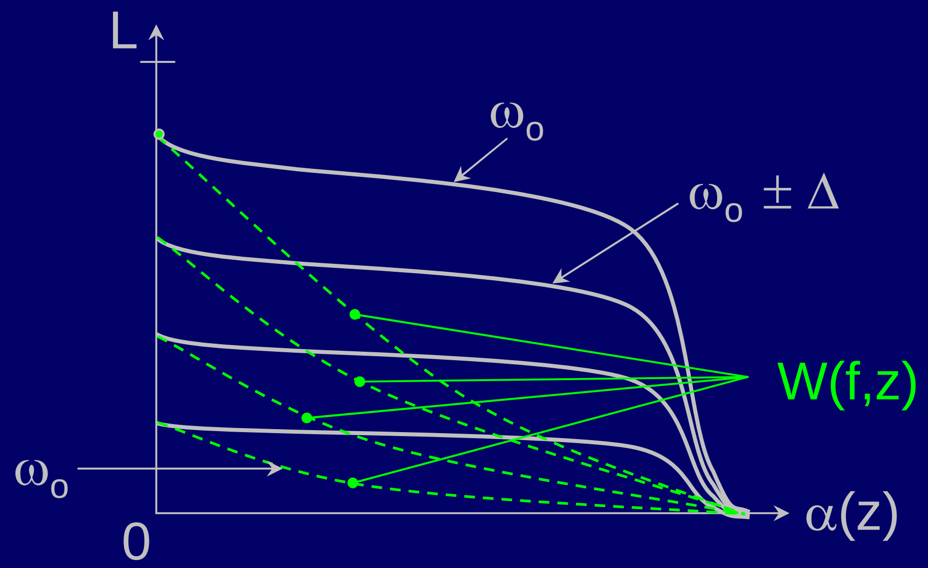


$$T_B \cong \int_0^L T(z)\left[d(z)e^{-\tau(z)}\right]dz \quad \text{for } \tau_o >> 1$$





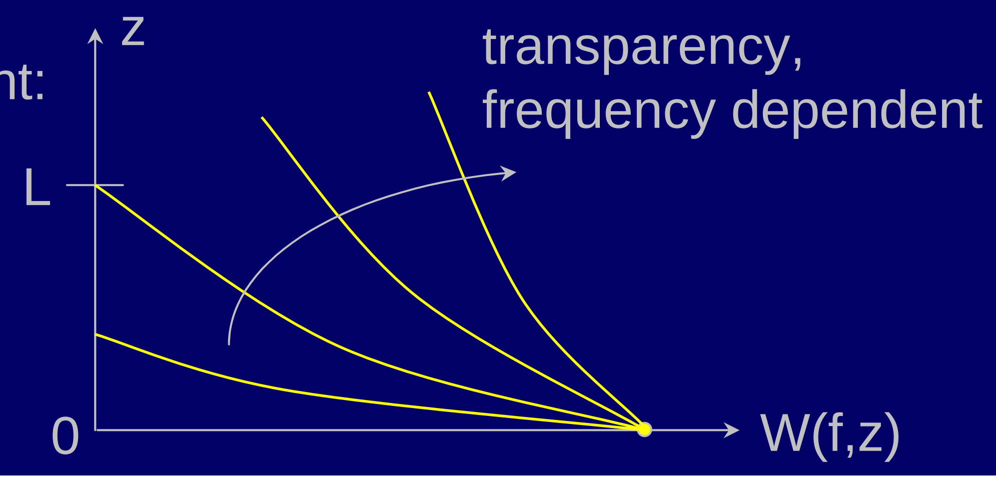
Atmospheric Temperature T(z) Retrievals from Below



$W(f,z) = \alpha(z)e^{-\int_0^L \alpha(z)dz}$ ~decaying exponentials, rate is fastest for ω_o

Temperature profile retrievals in semi-transparent solids or liquids where $(1/\alpha) >> \lambda$:

If $\alpha(z) \cong$ constant:



Atmospheric Composition Profiles

$$T_B \cong \int_0^L \underbrace{\rho(z)\left[\frac{\alpha(z)}{\rho(z)} \bullet T(z) e^{-\int_z^L \alpha(z)dz}\right]}_{W(z,f)} dz = T_{B_o} + \int_0^L \left[\rho(z) - \rho_o(z)\right] W'_\rho(z,f)\, dz$$

if viewed from space

Because $\alpha(z)$ and W(z,f) are strong functions of the unknown $\rho(z)$, this retrieval problem is quite non-linear and can be singular (e.g. if T(z) = constant). In this case, good statistics can be helpful. Incremental weighting functions defined relative to a nearly normal $\rho(z)$ can help linearize the problem.

Optimum Linear Estimates (Linear Regression)

Parameter vector estimate $\hat{\bar{p}} = \underset{\text{“determination matrix”}}{\bar{\bar{D}}}\,\underset{\text{data vector}}{\bar{d}}$ $\qquad \left(\bar{d} \triangleq [1, d_1, \ldots, d_N]\right)$

Choose $\bar{\bar{D}}$ to minimize $E\left[(\hat{\bar{p}} - \bar{p})^t (\hat{\bar{p}} - \bar{p})\right]$

Derive $\bar{\bar{D}}$:

$$\frac{\partial}{\partial D_{ij}} \left\{ E\left[(\hat{\bar{p}} - \bar{p})^t (\hat{\bar{p}} - \bar{p})\right] = 0 = \right.$$

$$= \frac{\partial}{\partial D_{ij}} E\left[\left(\bar{d}^t \bar{\bar{D}}^t - \bar{p}^t\right)\left(\bar{\bar{D}}\bar{d} - \bar{p}\right)\right] = E\left[2 d_j \underset{i^{TH} \text{ row of } \bar{\bar{D}}}{\bar{\bar{D}}_i}\, \bar{d} - 2 d_j p_i\right]$$

Therefore

$$\bar{\bar{D}}_i E\left[\bar{d} d_j\right] = E\left[p_i d_j\right]$$

$$\bar{\bar{D}}\, \underbrace{E\left[\bar{d}\bar{d}^t\right]}_{\triangleq\, \bar{\bar{C}}_d} = E\left[\bar{p}\bar{d}^t\right]; \quad \bar{\bar{C}}_d^t \bar{\bar{D}}^t = E\left[\bar{d}\bar{p}^t\right]$$

Optimum Linear Estimates (Linear Regression)

Therefore

$$\bar{\bar{D}}_i E\left[\bar{d} d_j\right] = E\left[p_i d_j\right]$$

$$\bar{\bar{D}} \underbrace{E\left[\bar{d}\bar{d}^t\right]}_{\stackrel{\Delta}{=}\ \bar{\bar{C}}_d} = E\left[\bar{p}\bar{d}^t\right];\ \bar{\bar{C}}_d^t \bar{\bar{D}}^t = E\left[\bar{d}\bar{p}^t\right]$$

The linear regression solution is $\hat{\bar{p}} = \bar{\bar{D}}\bar{d}$ where $\bar{\bar{D}}^t = \left[\bar{\bar{C}}_d\right]^{-1} E\left[\bar{d}\bar{p}^t\right]$

$\overline{\overline{D}}$ is Least-Square-Error Optimum if:

1) Jointly Gaussian process (physics + instrument):

$$p_r(\bar{r}) = \frac{1}{(2\pi)^{N/2}|\Lambda|^{1/2}} e^{-\frac{1}{2}\left[(\bar{r}-\bar{m})\overline{\overline{\Lambda}}^{-1}(\bar{r}-\bar{m})\right]}$$

$$\overline{\overline{\Lambda}} \triangleq E\left[(\bar{r}-\bar{m})\left(\bar{r}^t-\bar{m}^t\right)\right]$$

$$\bar{m} \triangleq E[\bar{r}], \text{ where } r = r_1, r_2, \ldots, r_N$$

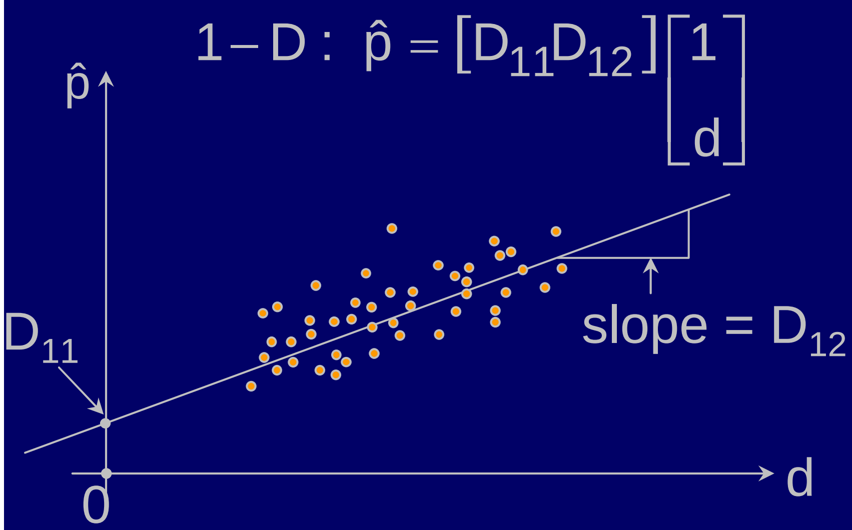
2) Problem is linear:

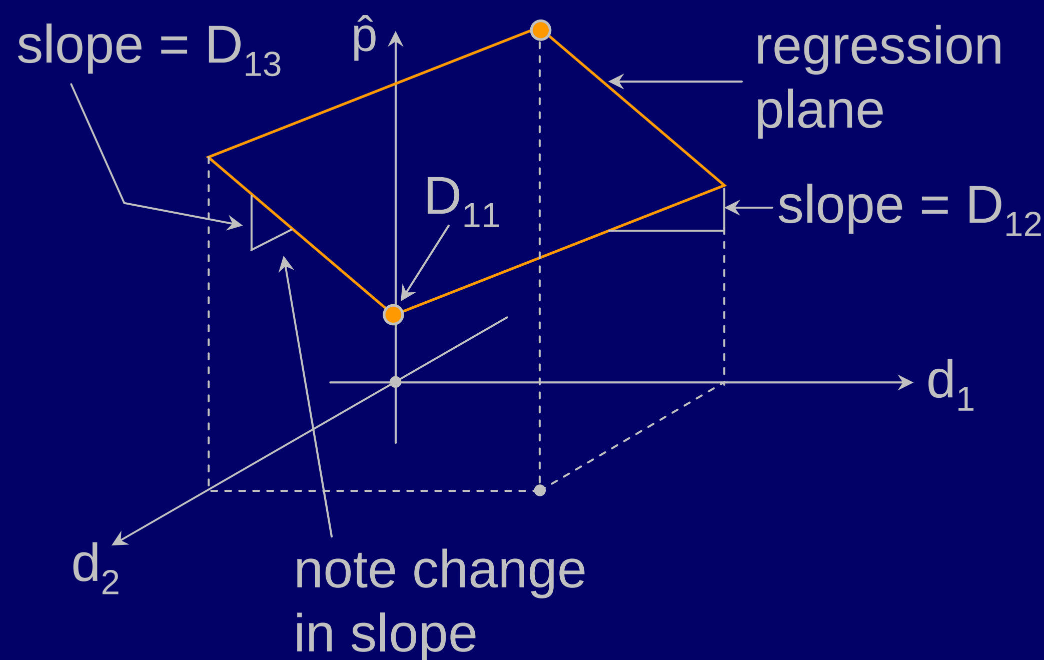
$$\text{data } \bar{d} = \overline{\overline{M}}\bar{r} + \bar{n} + \bar{d}_o$$

$$\uparrow \qquad \uparrow$$

parameter vector   noise (JGRVZM)

Examples of Linear Regression





Equivalently:

$$\hat{p} = \langle p \rangle + D_{12} \left(d - \langle d \rangle \right)$$

$$\left(E[\bullet] \equiv \langle \bullet \rangle \right)$$

Regression Information

1) The instrument alone (via weighting functions)

2) “Uncovered information (to which the instrument is blind, but which is correlated with properties the instrument does see; $\overline{\overline{D}}$ retrieves this too.)

3) “Hidden information (invisible in instrument and uncorrelated with visible information); it is lost.

Nature of Instrument-Provided Information

Assume linear physics: $\bar{d} = \bar{\bar{W}}\,\bar{T}$

Where
$\bar{d} = $ data vector $[1, d_1, d_2, \ldots, d_N]$
$\bar{T} = $ temperature profile $[T_1, T_2, \ldots, T_M]$
$\bar{\bar{W}} = $ weighting function matrix
$\bar{W}_i = i^{th}$ row of $\bar{\bar{W}}$

Claim:

If $\bar{T} = \sum_{i=1}^{N} a_i \bar{W}_i$ and noise $\bar{n} = 0$,

then $\hat{\bar{T}} = \bar{\bar{D}}\bar{d} = \bar{T}$, perfect retrieval $\left(\text{if } \bar{\bar{W}} \text{ not singular}\right)$

Proof for Continuous Variables

Claim:

$$\text{If } \overline{T} = \sum_{i=1}^{N} a_i \overline{W_i} \text{ and noise } \overline{n} = 0,$$

$$\text{then } \hat{\overline{T}} = \overline{\overline{D}}\,\overline{d} = \overline{T}, \text{ perfect retrieval } \left(\text{if } \overline{\overline{W}} \text{ not singular}\right)$$

Let:

$$W_1(h) \overset{\Delta}{=} b_{11}\phi_1(h)$$

$$W_2(h) \overset{\Delta}{=} b_{21}\phi_1(h) + b_{22}\phi_2(h)$$

$$W_3(h) \overset{\Delta}{=} b_{31}\phi_1(n) + b_{32}\phi_2(h) + b_{33}\phi_3(h)$$

Where: $\int_0^\infty \phi_i(h) \bullet \phi_j(h)dh = \delta_{ij} = \begin{cases} 0 \\ 1 \end{cases}$

$\phi_i; b_{ij}$ are known a priori (from physics). Then: $d_j \overset{\Delta}{=} \int_0^\infty T(h)W_j(h)dh$

Proof for Continuous Variables

Then: $d_j \triangleq \int_0^{\infty} T(h) W_j(h) dh$

$W_1(h) \triangleq b_{11}\phi_1(h)$

$W_2(h) \triangleq b_{21}\phi_1(h) + b_{22}\phi_2(h)$

If we force $T(h) \triangleq \sum_{i=1}^{N} k_j W_j(h)$

$$\text{Then:}\quad d_j = \sum_{i=0}^{N} \int_0^{\infty} \left(k_i W_i(h)\right) W_j(h) dh \Rightarrow \bar{d}^t = \bar{k}^t \bar{\bar{W}} \bar{\bar{W}}^t = \bar{k}^t \bar{\bar{Q}}$$

$$= \sum_{i=1}^{N} \int_0^{\infty} k_i \left(\sum_{m=1}^{i} b_{im}\phi_m(h) \right) \left(\sum_{n=1}^{N} b_{jn}\phi_n(h) \right) (dh) \triangleq \sum_{i=1}^{N} k_i Q_{ji}$$

Therefore $\bar{d} = \bar{\bar{Q}}\bar{k}$ where $\bar{\bar{Q}}$ is a known square matrix

So let $\hat{\bar{k}} = \bar{\bar{Q}}^{-1}\bar{d} = \bar{k}$ where $\bar{\bar{Q}}$ is non-singular

Proof for Continuous Variables

Claim: If $\overline{T} = \sum_{i=1}^{N} a_i \overline{W_i}$ and noise $\overline{n} = 0$,

then $\hat{\overline{T}} = \overline{\overline{D}}\,\overline{d} = \overline{T}$, perfect retrieval $\left(\text{if } \overline{\overline{W}} \text{ not singular}\right)$

So let $\hat{\overline{k}} = \overline{\overline{Q}}^{-1}\overline{d} = \overline{k}$ where $\overline{\overline{Q}}$ is non-singular

Then: $\hat{T}(h) \overset{\Delta}{=} \sum_{i=1}^{N} \hat{k}_i W_i(h) = \sum_{i=1}^{N} k_i W_i(h) = T(h)$ (exact) Q.E.D.

Equivalently: $\overline{T} \overset{\Delta}{=} \overline{\overline{W}}\,\overline{k} = \overline{\overline{W}}\left(\overline{\overline{Q}}^{-1}\overline{d}\right) = \left(\overline{\overline{W}}\,\overline{\overline{Q}}^{-1}\right)\overline{d} = \overline{\overline{D}}\,\overline{d}$

So:

$\overline{\overline{D}} = \overline{\overline{W}}\,\overline{\overline{Q}}^{-1} \overset{\Delta}{=}$ "minimum information" solution

Which is exact if $\overline{T} = \overline{\overline{W}}\,\overline{k}$, $\overline{n} = 0$

To what is an instrument blind?

$$W_1(h) \overset{\Delta}{=} b_{11}\phi_1(h)$$

$$W_2(h) \overset{\Delta}{=} b_{21}\phi_1(h) + b_{22}\phi_2(h)$$

An instrument is blind to T(h) components outside the space spanned by $\phi_1, \phi_2, \ldots, \phi_N$ or, equivalently, by its W_1, $W_2, \ldots, W_n$.

By definition, the instrument is blind to any $\phi_j \perp W_i$, for all i.

Statistical Methods Can Reveal "Hidden" Components

In general, $$T(h) = \underbrace{\sum_{i=1}^{N} k_i W_i(h)}_{\text{seen by N instrument channels}} + \underbrace{\sum_{N+1}^{\infty} a_i \phi_i(h)}_{\text{all hidden components}}$$

Extreme case: suppose $\phi_1(h)$ always accompanied by $\frac{1}{2}\phi_{N+1}(h)$.

Then our present solution: $$\hat{T}(h) = \sum_{i=1}^{N} k_i W_i(h) = \sum_{i=1}^{N} a_i \phi_i(h)$$

Would become: $$\hat{T}(h) = a_1\left(\phi_1 + \underbrace{\frac{1}{2}}_{\text{shrinks with decorrelation}}\phi_{N+1}\right) + \sum_{i=2}^{N} a_i \phi_i$$

Thus hidden components can be "uncovered" if correlated with visible ones.

General Linear Estimate

$$\hat{\bar{T}} = \bar{\bar{D}}\bar{d} \text{ where } \bar{D}_i = \underbrace{\left[\bar{\bar{W}}\bar{\bar{Q}}^{-1}\right]_i}_{\substack{\text{minimum}\\\text{information}}} + \underbrace{\overbrace{\sum_{j=N+1}^{\infty} a_{ij}\phi_j}^{\triangleq\ \beta_i}}_{\substack{\text{“uncovered”}\\\text{information}}}$$

Thus retrieval can be drawn only from the space spanned by

$$\phi_1,\phi_2,\ldots,\phi_N;\ \beta_1,\beta_2,\ldots,\beta_N \quad \left(\text{dimensionality is N; } \hat{\bar{T}} = \sum_{i=1}^{N} d_i \bar{D}_i\right)$$

That is, N channels contribute N orthogonal basis functions to the minimum-information solution, plus N more basis functions which are orthogonal but correlated with the first N.

General Linear Estimate

As N increases, the fraction of the hidden space which is "uncovered" by statistics is therefore likely to increase, even as the hidden space shrinks.

In general:

a priori variance = observed + uncovered + variance lost

due to noise and decorrelation.

Example: 8 channels of AMSU versus 4 channels of MSU

AMSU and MSU are passive microwave spectrometers in earth orbit sounding atmospheric temperature profiles from above with ~10-km wide weighting functions peaking at altitudes from 3 to 26 km. Note the larger ratio of uncovered/lost power for AMSU.

Example: 8-Channel AMSU vs 4-Channel MSU