Remote sensing is a quasi-linear estimation problem.

Equation of radiative transfer:

\[ T_B(\circ K) = T_{B_0} e^{-\tau_0} + \int_0^L T(z)\alpha(z)e^{-\tau(z)}dz \]

\[ \tau(z) = \int_0^z \alpha(z)dz \]

\[ \tau_0 = \tau(0) \]
Radiation from Sky-Illuminated Reflective Surfaces

\[ R = 1 - \varepsilon \]

\[ T_B (\text{oK}) = RT_s e^{-2\tau_o} + \varepsilon T_G e^{-\tau_o} + \int_0^L L(z) e^{-\int_0^Z \alpha(z) dz} dz \]

\[ + \int_0^L T(z)\alpha(z) e^{-\int_0^Z \alpha(z)dz} dz \]

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

reflectivity
emissivity (specular surface)
Temperature Weighting Function $W(z,f,T)$

Terms:

1. $T_B(f)$
2. $T_o$
3. $z$
4. $T(z)$

For $\tau_o >> 1$:

$$T_B(f) \cong T_o + \int_0^L T(z)W(z,f,T(z))\,dz$$

Alternatively,

$$T_B(f) \cong T'_o + \int_0^L (T(z) - T_o(z))W'(z,f,T_o(z))\,dz$$

Incremental weighting function:

$$W'(z,f,T_o(z)) = \frac{\partial T_B}{\partial T(z)}T_o(z)$$

Note: we have ~linear relation:

$$T_B(f) \leftrightarrow T(z)$$

(not Fourier)
Atmospheric Temperature $T(z)$ Retrievals from Space

Therefore $\alpha_0 \neq f(P)$ for P-broadening trace constituents

$$T_B \approx \int_0^L T(z) \left[ \alpha(z)e^{-\tau(z)} \right] dz \text{ for } \tau_0 \gg 1$$
Atmospheric Temperature $T(z)$ Retrievals from Space

Pressure dominates $\Delta \omega$

$$\omega \approx \omega_0$$

Increasing $P$

$$\Delta \omega$$

$\omega_0$ to $\omega$

Increasing $P$

$$\tau \approx 1$$

Scale height

$$\tau > > 1$$
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 $\omega_o$

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

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

transparency, frequency dependent
Atmospheric Composition Profiles

\[ T_B \approx \int_0^L \rho(z) \left[ \frac{\alpha(z)}{\rho(z)} \cdot T(z) e^{-\int_0^L \alpha(z) dz} \right] dz = T_{B_0} + \int_0^L \left[ \rho(z) - \rho_0(z) \right] W'(z,f) dz \]

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 \( \hat{\rho} = \bar{D} \bar{d} \) (\( \bar{d} = [d_1, \ldots, d_N] \))

“determination matrix” data vector

Choose \( \bar{D} \) to minimize \( E[(\hat{\rho} - \rho)^t(\hat{\rho} - \rho)] \)

Derive \( \bar{D} \):

\[
\frac{\partial}{\partial D_{ij}} E[(\hat{\rho} - \rho)^t(\hat{\rho} - \rho)] = 0 \quad \text{row } i \text{ of } \bar{D}
\]

\[
= \frac{\partial}{\partial D_{ij}} E[(d^t \bar{D} - \rho^t)(\bar{D} \bar{d} - \rho)] = E[2d_j \bar{D} \bar{d} - 2d_j \rho_i]
\]

Therefore

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

\[
\bar{D} E[\bar{d}d^t] = E[pd^t] ; \quad \bar{C}_d \bar{D}^t = E[d\rho^t]
\]

\( \Delta \equiv \bar{C}_d \)
The linear regression solution is \( \hat{\rho} = D\hat{d} \) where \( D^t = \left[ C_d \right]^{-1} \mathbb{E} \left[ d\rho^t \right] \).
D is Least-Square-Error Optimum if:

1) Jointly Gaussian process (physics + instrument):

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

\[ \bar{\Lambda} \Delta= \mathbb{E}\left[(\bar{r} - \bar{m})(\bar{r}^t - \bar{m}^t)\right] \]

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

2) Problem is linear:

\[ \text{data } \bar{d} = \bar{M}r + \bar{n} + \bar{d}_o \]

parameter vector noise (JGRVZM)
Examples of Linear Regression

1 - D: \( \hat{\rho} = [D_{11} D_{12}] \begin{bmatrix} 1 \\ d \end{bmatrix} \)

Equivalently:

\[ \hat{\rho} = \langle \rho \rangle + D_{12} (d - \langle d \rangle) \]

\( (E[\bullet] \equiv \langle \bullet \rangle) \)
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; \( \bar{D} \) 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: \( \vec{d} = \overline{W} \overline{T} \)

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

Claim:

If \( \overline{T} = \sum_{i=1}^{N} a_i \overline{W}_i \) and noise \( \vec{n} = 0 \),
then \( \overline{\vec{T}} = \overline{\vec{D}} \vec{d} = \overline{T} \), perfect retrieval (if \( \overline{W} \) not singular)
Proof for Continuous Variables

Claim:

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

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

Let:

\[
\begin{align*}
W_1(h) &\triangleq b_{11}\phi_1(h) \\
W_2(h) &\triangleq b_{21}\phi_1(h) + b_{22}\phi_2(h) \\
W_3(h) &\triangleq b_{31}\phi_1(n) + b_{32}\phi_2(h) + b_{33}\phi_3(h)
\end{align*}
\]

Where:

\[
\int_0^\infty \phi_i(h) \cdot \phi_j(h) dh = \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}
\]

\(\phi_i; b_{ij} \) are known a priori (from physics). Then:

\[
d_j \triangleq \int_0^\infty T(h)W_j(h)dh
\]
Proof for Continuous Variables

Then: \[ d_j \overset{\Delta}{=} \int_0^\infty T(h)W_j(h)\,dh \quad W_1(h) \overset{\Delta}{=} b_{11} \phi_1(h) \]

If we force \[ T(h) \overset{\Delta}{=} \sum_{i=1}^N k_i W_j(h) \]

Then: \[ d_j = \sum_{i=0}^N \int_0^\infty (k_i W_i(h)) W_j(h)\,dh \Rightarrow \bar{d}^t = \bar{k}^t \bar{W} \bar{W}^t = \bar{k}^t \bar{Q} \]

\[ = \sum_{i=1}^N \int_0^\infty k_i \left( \sum_{m=1}^i b_{im} \phi_m(h) \right) \left( \sum_{n=1}^N b_{jn} \phi_n(h) \right)\,(dh) \overset{\Delta}{=} \sum_{i=1}^N k_i Q_{ji} \]

Therefore \[ \bar{d} = \bar{Q} k \text{ where } \bar{Q} \text{ is a known square matrix} \]

So let \[ \bar{k} = \bar{Q}^{-1} \bar{d} = \bar{k} \text{ where } \bar{Q} \text{ is non-singular} \]
Proof for Continuous Variables

Claim: If $\bar{T} = \sum_{i=1}^{N} a_i \bar{W}_i$ and noise $\bar{n} = 0$, then $\hat{\bar{T}} = \bar{D}\bar{d} = \bar{T}$, perfect retrieval (if $\bar{W}$ not singular)

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

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

Equivalently: $\bar{T} \triangleq \bar{W}\bar{k} = \bar{W}\left(\bar{Q}^{-1}\bar{d}\right) = \left(\bar{W}\bar{Q}^{-1}\right)\bar{d} = \bar{D}\bar{d}$

So: $\bar{D} = \bar{W}\bar{Q}^{-1} \triangleq "minimum\ information" \ solution$

Which is exact if $\bar{T} = \bar{W}\bar{k}$, $\bar{n} = 0$
To what is an instrument blind?

\[ W_1(h) \triangleq b_{11}\phi_1(h) \]
\[ W_2(h) \triangleq b_{21}\phi_1(h) + b_{22}\phi_2(h) \]

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, \( T(h) = \sum_{i=1}^{N} k_i W_i(h) + \sum_{N+1}^{\infty} a_i \phi_i(h) \)

seen by N instrument channels  
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: \( \hat{T}(h) = a_1 \left( \phi_1 + \frac{1}{2} \phi_{N+1} \right) + \sum_{i=2}^{N} a_i \phi_i \)

shrinks with decorrelation

Thus hidden components can be “uncovered” if correlated with visible ones.
\[
\hat{T} = \bar{D}d \text{ where } \bar{D}_i = \left[WQ^{-1}\right]_i + \sum_{j=N+1}^{\infty} a_{ij}\phi_j
\]

Thus retrieval can be drawn only from the space spanned by \(\phi_1, \phi_2, \ldots, \phi_N; \beta_1, \beta_2, \ldots, \beta_N\) (dimensionality is \(N\); \(\hat{T} = \sum_{i=1}^{N} d_i\bar{D}_i\))

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 \(N\).
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:

\[
\text{a priori variance} = \text{observed} + \text{uncovered} + \text{variance lost due to noise and decorrelation.}
\]

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 $\sim 10$-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

MID-LATITUDES (TOTAL POWER = 1222 K², 15 LEVELS)

- OBSERVED POWER
  - MSU - 55° INCIDENCE ANGLE, LAND
  - MSU - 0° ANGLE, NADIR
  - AMSU - NADIR
  - AMSU - NADIR

- UNCOVERED POWER
  - 3 km WEIGHTING FUNCTION
  - 10
  - 16
  - 6
  - 12
  - 22
  - 26

- LOST POWER

TROPICS (TOTAL POWER = 184 K²)

- OBSERVED POWER
  - MSU – NADIR, LAND
  - AMSU - NADIR

- UNCOVERED

- LOST

Lec22.5-21
3/27/01