Optimization and data assimilation

- Linear problem without error

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

A: m×n matrix, m>n: measurements>unknown

rank of matrx: p, p<n: under-determined problem

- Linear problem with error

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \delta$$

$$y_1 = x_1 + 0.99x_2 + \delta_1 \qquad x_1 = \frac{y_1 - \delta_1 - 0.99(y_2 - \delta_2)}{1 - 0.99^2}$$

$$y_2 = 0.99x_1 + x_2 + \delta_2$$

Linear estimator

- Linear forward model

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \delta = \mathbf{A}\mathbf{x} + \delta$$

$$p(\mathbf{y} \mid \mathbf{x}) = \frac{1}{(2\pi)^{n/2} \mid \mathbf{S}_\delta \mid^{1/2}} \exp\left[-\tfrac{1}{2}(\mathbf{y} - \mathbf{A}\mathbf{x})^T \mathbf{S}_\delta^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x})\right]$$

$$-2\ln p(\mathbf{y} \mid \mathbf{x}) = (\mathbf{y} - \mathbf{A}\mathbf{x})^T \mathbf{S}_\delta^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x}) + c_1$$

$$-2\ln p(\mathbf{x}) = (\mathbf{x} - \mathbf{x}_a)^T \mathbf{S}_a^{-1} (\mathbf{x} - \mathbf{x}_a) + c_2$$

$$-2\ln p(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} - \mathbf{A}\mathbf{x})^T \mathbf{S}_\delta^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x}) + (\mathbf{x} - \mathbf{x}_\mathbf{a})^T \mathbf{S}_\mathbf{a}^{-1} (\mathbf{x} - \mathbf{x}_\mathbf{a}) + c_3$$

- Estimator

$$\hat{\mathbf{x}} = \mathbf{x}_a + (\mathbf{A}\mathbf{S}_\delta^{-1}\mathbf{A}^\mathbf{T} + \mathbf{S}_\mathbf{a}^{-1})^{-1} \mathbf{A}^T \mathbf{S}_\delta^{-1} (\mathbf{y} - \mathbf{A}\mathbf{x}_\mathbf{a})$$

Nonlinear inversion

- Nonlinear forward problem

$$\mathbf{y} = \mathbf{F}(\mathbf{x})$$

- Not exactly a problem of solving a set of nonlinear eqs.
  - Prior information
  - Error characterization
  - Cost function, not quadratic

Degree of nonlinearity

- Linearization

$$\hat{\mathbf{y}} - \mathbf{y}_\mathbf{a} = \mathbf{F}(\hat{\mathbf{x}}) - \mathbf{F}(\mathbf{x}_\mathbf{a}) \approx \mathbf{A}(\hat{\mathbf{x}} - \mathbf{x}_\mathbf{a}) + \Delta_\mathbf{y}$$

$$\Delta_\mathbf{y} = \mathbf{F}(\hat{\mathbf{x}}) - \mathbf{F}(\mathbf{x}_\mathbf{a}) - \mathbf{A}(\hat{\mathbf{x}} - \mathbf{x}_\mathbf{a})$$

- Nonlinearity

$$c^2 = \Delta_\mathbf{y}^T \mathbf{S}_\delta^{-1} \Delta_\mathbf{y}$$

Newton iteration method

- Extension from linear solution

$$\hat{\mathbf{x}} = \mathbf{x}_a + (\mathbf{A}\mathbf{S}_\delta^{-1}\mathbf{A}^{\mathbf{T}} + \mathbf{S}_{\mathbf{a}}^{-1})^{-1}\mathbf{A}^T\mathbf{S}_\delta^{-1}(\mathbf{y} - \mathbf{A}\mathbf{x}_{\mathbf{a}})$$

$$\mathbf{y} - \mathbf{A}\mathbf{x}_{\mathbf{a}} \Rightarrow y - \mathbf{F}(\hat{\mathbf{x}}) - \mathbf{A}(\hat{\mathbf{x}} - \mathbf{x}_{\mathbf{a}})$$

$$\hat{\mathbf{x}}_{i+1} = \mathbf{x}_a + (\mathbf{A}\mathbf{S}_\delta^{-1}\mathbf{A}^{\mathbf{T}} + \mathbf{S}_{\mathbf{a}}^{-1})^{-1}\mathbf{A}^T\mathbf{S}_\delta^{-1}(\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}}_i) - \mathbf{A}(\hat{\mathbf{x}}_i - \mathbf{x}_{\mathbf{a}}))$$

- Simplified form (Levenberg-Marquardt method)

$$\hat{\mathbf{x}}_{i+1} = \hat{\mathbf{x}}_i + (\mathbf{A}\mathbf{A}^{\mathbf{T}} + \gamma_i\mathbf{I})^{-1}\mathbf{A}^T(\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}}_i))$$

Data assimilation as inversion

- Numerical weather prediction: Initialization with observations
  - Not straightforward
  - Observation error
  - Compatibility problem
- As inverse problem, taking both prior information and measurement error into account in retrieval process

Assimilation methods

$$J = [\mathbf{x} - \mathbf{x_a}]^{\mathbf{T}} \mathbf{S_a^{-1}} [\mathbf{x} - \mathbf{x_a}] + [\mathbf{y} - \mathbf{F(x)}]^{\mathbf{T}} \mathbf{S_\delta^{-1}} [\mathbf{y} - \mathbf{F(x)}]$$

- Successive correction

$$\hat{x}_i - x_{ai} = \sum_k \frac{\sigma_a^2 w_{ik}}{\sigma_a^2 w_{ik} + \sigma_k^2} (y_k - y_{ak})$$

- Variational analysis: adjoint method

- Ensemble Kalman filtering

Adjoint method

- Iterative solution

$$\mathbf{x}_{i+1} = \mathbf{x}_i - [\nabla_{\mathbf{x}} (\nabla_{\mathbf{x}} J]^{-1} \nabla_{\mathbf{x}} J$$

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \gamma_i \nabla_{\mathbf{x}_i} J$$

- Adjoint solution

$$\mathbf{x}_t = \mathbf{M}_t \mathbf{x}_{t-1}$$

$$\mathbf{y}_t = \mathbf{F}(\mathbf{x}_t) + \delta_t = \mathbf{F}(\mathbf{M_t} \ldots \mathbf{M_1} \mathbf{x}_0) + \delta_t$$

$$J = \sum_{t=0}^{T} [(\mathbf{y} - \mathbf{F}(\mathbf{x}_i)]^T \mathbf{S}_\delta^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}_i)]$$

$$\nabla_{\mathbf{x}_0} J = -\sum_{t=0}^{T} [\partial \mathbf{F_t}(\mathbf{x_t}) / \partial \mathbf{x}_0]^T \mathbf{S}_{\delta,t}^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}_i)] = 0$$