Outline

1. Fancier methods
   - SIMEX
   - Kalman filter

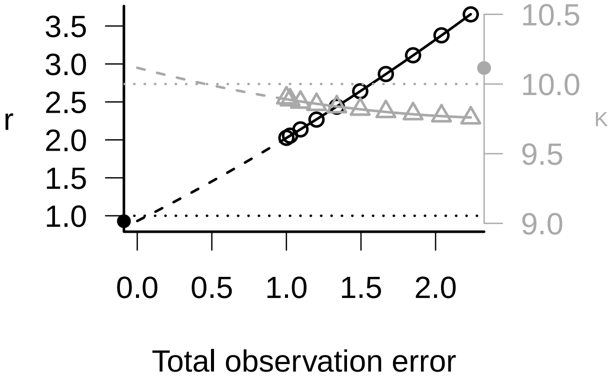
SIMEX

- **SIM**ulation-**EX**trapolation method
- Requires (1) an independent estimate of the observation error; (2) that we can sensibly add **additional** observation error to the data
- Slightly easier for Normal errors
- Probably most sensible for experimental data?
- Examples: Ellner et al. (2002); Melbourne and Chesson (2006)

Procedure

- based on estimated observation error, pick a range of increased error values, e.g. tripling the existing observation variance in 4–8 steps
- for each error magnitude, generate a data set with that increased error (more stable to inflate a single set of errors)
- estimate parameters for each set using gradient matching (i.e. assume $\sigma^2_{\text{obs}} = 0$)
- fit a linear or quadratic regression model for parameter $= f(\text{total error})$
- extrapolate the fit to zero

Logistic fit



Kalman filter

- General approach to account for dynamic variance, expected population state
- Works for **linear** (typically Normal) models; can be extended to nonlinear models
- Natural multivariate extensions: include bias, external shocks, etc. (Schnute, 1994)

Concept and implementation

- **Concept**
  - Variance increases with process error;
    decreases with (accurate) observations
  - Expected population state follows expected dynamics;
    drawn toward (accurate) observations
- **Procedure** (pseudo-pseudo-code)
  - Run KF for specified values of parameters, σ^2_{obs}, σ^2_{proc} to compute $\hat{N}(t)$, $\sigma^2_N(t)$
  - Estimate objective function (SSQ) for $N_{\text{obs}}|\hat{N}, \sigma^2_N$
  - Minimize over $\{\text{parameters}, \sigma^2_{\text{obs}}, \sigma^2_{\text{proc}}\}$

Concept and implementation

- **Concept**
  - Variance increases with process error; decreases with (accurate) observations
  - Expected population state follows expected dynamics; drawn toward (accurate) observations
- **Procedure** (pseudo-pseudo-code)
  - Run KF for specified values of parameters, σ^2_{obs}, σ^2_{proc} to compute $\hat{N}(t)$, $\sigma^2_N(t)$
  - Estimate objective function (SSQ) for $N_{\text{obs}}|\hat{N}, \sigma^2_N$
  - Minimize over $\{\text{parameters}, \sigma^2_{\text{obs}}, \sigma^2_{\text{proc}}\}$

Autoregressive model

$$N(t) \sim \text{Normal}(a + bN(t-1), \sigma^2_{\text{proc}})$$
$$N_{\text{obs}}(t) \sim \text{Normal}(N((t), \sigma^2_{\text{obs}})$$

- $b < 1, a > 0 \rightarrow$ stable dynamics
- $b > 1 \rightarrow$ exponential growth

Procedure

1. Update mean, variance of true density according to previous expected mean and variance:

$$\text{mean}(N(t)|N_{\text{obs}}(t-1)) \equiv \mu_1 = a + b\mu_0$$
$$\text{Var}(N(t)|N_{\text{obs}}(t-1)) \equiv \sigma_1^2 = b^2\sigma_0^2 + \sigma_{\text{proc}}^2$$

2. Now update the mean and variance of the **observed** density at time t:

$$\begin{aligned}\text{mean}(N_{\text{obs}}(t)|N_{\text{obs}}(t-1)) &\equiv \mu_2 = \mu_1 \\ \text{Var}(N_{\text{obs}}(t)|N_{\text{obs}}(t-1)) &\equiv \sigma_2^2 = \sigma_1^2 + \sigma_{\text{obs}}^2\end{aligned}$$

3. Now update true (expected) mean and variance to account for **current** observation:

$$\text{mean}(N|N_{\text{obs}}(t)) \equiv \mu_3 = \mu_1 + \frac{\sigma_1^2}{\sigma_2^2}(N_{\text{obs}}(t) - \mu_2)$$

$$\text{Var}(N(t)|N_{\text{obs}}(t)) \equiv \sigma_3^2 = \sigma_1^2\left(1 - \frac{\sigma_1^2}{\sigma_2^2}\right)$$

Pseudo-code

```
KFpred <- function(params,var_proc,var_obs,init) {
  set_initial_values
  for (i in 2:nt) {
     ## ... calculate mu{1-3}, sigma^2{1-3} as above
     N[i] <- mu_3; Var[i] <- sigmasq_3
  }
  return(list(N=N,Var=Var))
}
KFobj <- function(params,var_proc,var_obs,init,Nobs) {
    pred <- KFpred(params,var_proc,var_obs,init)
    obj_fun(Nobs,mean=pred$N,sd=sqrt(pred$Var))
}
minimize(KFobj,start_values,Nobs)
```

Extended Kalman filter

To fit (mildly) nonlinear models with the deterministic skeleton

$$N(t+1) = f(N(t)),$$

we just replace a and b in the autoregressive model $N(t+1) = a + bN(t)$ with the coefficients of the first two terms of the **Taylor expansion** of $f()$:

$$f(N(\tau)) \approx f(N(t)) + \frac{df}{dN}(N(\tau) - N(t)) + \dots$$

Multivariate extension (Schnute, 1994)

$$\begin{aligned} \text{process: } \boldsymbol{X}_t &= \boldsymbol{A}_t + \boldsymbol{B}_t \boldsymbol{X}_{t-1} + \boldsymbol{\delta}_t \\ \text{observation: } \boldsymbol{Y}_t &= \boldsymbol{C}_t + \boldsymbol{D}_t \boldsymbol{X}_t + \boldsymbol{\epsilon}_t \end{aligned}$$

Allows for bias, cross-species effects in both process and observation, correlation in process and observation noise ...

References

Ellner, S.P., Seifu, Y., and Smith, R.H., 2002. *Ecology*, 83(8):2256–2270.

Melbourne, B.A. and Chesson, P., 2006. *Ecology*, 87:1478–1488.

Schnute, J.T., 1994. *Canadian Journal of Fisheries and Aquatic Sciences*, 51:1676–1688.