Lecture Notes on

Forecasting

Yakov Ben-Haim
Former Yitzhak Moda'i Chair in Technology and Economics
Faculty of Mechanical Engineering
Technion — Israel Institute of Technology
Haifa 32000 Israel
yakov@technion.ac.il
http:/info-gap.com http:/www.technion.ac.il/yakov

Source material:[1]

• Yakov Ben-Haim, 2010, *Info-Gap Economics: An Operational Introduction,* Chapter 6: Estimation and Forecasting, Palgrave-Macmillan.

• Yakov Ben-Haim, 2009, Info-gap forecasting and the advantage of sub-optimal models, *European Journal of Operational Research,* 197: 203–213. Link to pre-print at:
http:/info-gap.com/content.php?id=22

A Note to the Student: These lecture notes are not a substitute for the thorough study of books. These notes are no more than an aid in following the lectures.

Contents

[0]\lectures\Econ-Dec-Mak\forecasting001.tex 19.6.2022

[1]Additional material in the file: Yakov Ben-Haim, Lecture notes on info-gap estimation and forecasting, \lectures\risk\lectures\estim02.pdf.

1 1-D Dynamic System: European Central Bank Overnight Interest Rates

1.1 The Data and the Questions

Date	Interest rate	Implied λ
1 Jan 1999	4.50	
9 Apr 1999	3.50	0.778
5 Nov 1999	4.00	1.143
4 Feb 2000	4.25	1.063
17 Mar 2000	4.50	1.059
28 Apr 2000	4.75	1.056
9 Jun 2000	5.25	1.105
28 Jun 2000	5.25	1.000
1 Sep 2000	5.50	1.048
6 Oct 2000	5.75	1.045
11 May 2001	5.50	0.957
31 Aug 2001	5.25	0.955

Table 1: Interest rates for overnight loans at the European Central Bank (marginal lending facility). Source: http://www.ecb.int/stats/monetary/rates/html/index.en.html

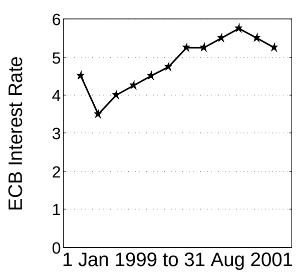

Figure 1: ECB Interest Rates

§ **ECB overnight interest rates: table 1.**
- First loans: 1999.
- Data through August 2001.
- 9 June 2000–31August 2001: $\mu = 5.4\%$, $\sigma = 0.19\%$.
- Typical change: 25 basis points.
- Largest change: 100 basis points.

§ **El-Qaeda attacks in US: 11 Sept 2001.**
- **Predict next interest rate** on 9/12/2001, (1 day after 9/11).
- Asymmetric uncertainty: rate will go down (**Why?**), but **by how much?**

§ **Questions:**
- How to forecast the rate in light of the great uncertainty?
- How to assess confidence in the forecast?

1.2 1-Step Dynamics and Robustness

§ **Uncertain historical model** of dynamical variable y_n, e.g. interest rate:

$$y_1 = \lambda_1 y_0, \quad y_0 > 0, \text{ known}, \quad \lambda_1 \text{ uncertain} \tag{1}$$

§ **Info-gap model. Asymmetric uncertainty:**

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_1 : \ (1-h)\widetilde{\lambda} \le \lambda_1 \le \widetilde{\lambda} \right\}, \quad h \ge 0 \tag{2}$$

- $\widetilde{\lambda}$: known and positive estimate transition coefficient.
- λ_1: unknown true transition coefficient anticipated to be no greater, probably less, than $\widetilde{\lambda}$.

§ **Slope-adjusted forecasting model.** We must choose the "slope" ℓ:

$$y_1^s = \ell y_0 \tag{3}$$

§ **Performance requirement:**
- Absolute error:

$$\varepsilon = |y_1^s - y_1| = |(\ell - \lambda_1) y_0| \tag{4}$$

- Performance requirement:

$$\varepsilon \le \varepsilon_{\rm c} \tag{5}$$

§ **Robustness:**
- Definition:

$$\widehat{h}(\ell, \varepsilon_{\rm c}) = \max \left\{ h : \ \left(\max_{\lambda_1 \in \mathcal{U}(h)} \varepsilon(\lambda_1) \right) \le \varepsilon_{\rm c} \right\} \tag{6}$$

- $m(h)$ is inner maximum in eq.(6): inverse of $\widehat{h}(\varepsilon_{\rm c})$.
- **We will consider a special case:**

$$\ell \le \widetilde{\lambda} \tag{7}$$

- Recall: $y_0 > 0$ and known.
- $m(h)$ occurs for an extremal value of λ_1 at horizon of uncertainty h: either $\widetilde{\lambda}$ or $(1-h)\widetilde{\lambda}$.
- $m(h)$ is the greater of the following:

$$m_1(h) = \left| \widetilde{\lambda} - \ell \right| y_0 \tag{8}$$

$$m_2(h) = \left| \ell - (1-h)\widetilde{\lambda} \right| y_0 \tag{9}$$

- Clearly $m_1(h) > m_2(h)$ for **small** h because $\ell \le \widetilde{\lambda}$. To find the transition:

$$\widetilde{\lambda} - \ell \ \ge \ \ell - (1-h)\widetilde{\lambda} \tag{10}$$

$$\Longleftrightarrow 2\left(\widetilde{\lambda} - \ell\right) \ \ge \ h\widetilde{\lambda} \tag{11}$$

$$\Longleftrightarrow \quad h \ \le \ \frac{2\left(\widetilde{\lambda} - \ell\right)}{\widetilde{\lambda}} \tag{12}$$

- Hence:

$$m(h) = \begin{cases} \left(\widetilde{\lambda} - \ell\right) y_0, & \text{if } h \le \dfrac{2\left(\widetilde{\lambda} - \ell\right)}{\widetilde{\lambda}} \\ \left(\ell - (1-h)\widetilde{\lambda}\right) y_0, & \text{else} \end{cases} \tag{13}$$

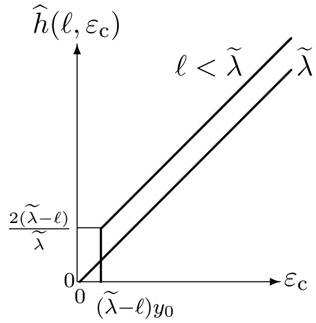

Figure 2: Robustness curve $\widehat{h}(\ell, \varepsilon_c)$, eq.(14).

§ **Robustness function:** equate $m(h)$ to ε_c and solve for h to find the robustness. One finds (fig. 2):

$$\widehat{h}(\ell, \varepsilon_c) = \begin{cases} 0, & \text{if } \varepsilon_c < \left(\widetilde{\lambda} - \ell\right) y_0 \\ \dfrac{\varepsilon_c + \left(\widetilde{\lambda} - \ell\right) y_0}{\widetilde{\lambda} y_0}, & \text{else} \end{cases} \tag{14}$$

- Trade off: robustness $\widehat{h}$ up (good) as critical error ε_c up (not good).
- Zeroing: No robustness at estimated error.
- Discontinuous robustness curve for $\ell < \widetilde{\lambda}$.
- Crossing robustness curves, fig. 3: preference reversal.

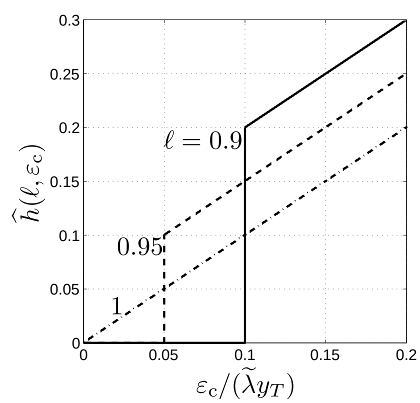

Figure 3: Robustness vs normalized forecast error. $\widetilde{\lambda} = 1$, $y_T = 5.25$.

§ **Robustness curves for 3 slope-adjusted models, fig. 3:**

- $\ell = 1.0 \Longrightarrow$ 0% robustness at 0% error.
- $\ell = 0.95 \Longrightarrow$ 10% robustness at 5% error (more robust than $\ell = 1.0$).
- $\ell = 0.9 \Longrightarrow$ 20% robustness at 10% error (more robust than $\ell = 0.95$).

§ **Forecast:**

- Model used: $\ell = 0.9 \Longrightarrow$ 20% robustness at 10% error.
- Forecast: $y^{s}_{T+1} = 0.9 y_T =$ 4.725.
- Outcome:
  - ∘ y_{T+1} = 4.75 on 18.9.2001.
  - ∘ -0.5% forecast error.

1.3 Multi-Step Dynamics and Robustness

§ **Uncertain historical model:**

$$y_{t+1} = \lambda_{t+1} y_t, \quad t = 0, 1, 2 \ldots, \ y_0 > 0, \text{ known}, \quad \lambda_{t+1} \text{ uncertain} \tag{15}$$

Thus:

$$y_t = y_0 \prod_{j=1}^{t} \lambda_j, \quad t = 1, 2, \ldots \tag{16}$$

§ **Info-gap model. Asymmetric uncertainty:**

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_t : \ (1-h)\widetilde{\lambda} \le \lambda_t \le \widetilde{\lambda}, \ t = 1, 2, \ldots \right\}, \quad h \ge 0 \tag{17}$$

Assume $\widetilde{\lambda} > 0$.

§ **Slope-adjusted forecasting model.** We must choose the 'slope' ℓ:

$$y_{t+1}^s = \ell y_t^s \tag{18}$$

Thus:

$$y_t^s = \ell^t y_0 \tag{19}$$

§ **Performance requirement:**

• Absolute error:

$$\varepsilon_t = |y_t^s - y_t| = \left| \ell^t - \prod_{j=1}^{t} \lambda_j \right| y_0 \tag{20}$$

• Performance requirement:

$$\varepsilon_t \le \varepsilon_{\mathrm{c}} \tag{21}$$

§ **Robustness:**

• Definition:

$$\widehat{h}_t(\ell, \varepsilon_{\mathrm{c}}) = \max \left\{ h : \ \left(\max_{\lambda \in \mathcal{U}(h)} \varepsilon_t(\lambda) \right) \le \varepsilon_{\mathrm{c}} \right\} \tag{22}$$

• $m(h)$ is inner maximum in eq.(22): inverse of $\widehat{h}_t(\varepsilon_{\mathrm{c}})$.

• **Special case:**

$$\ell \le \widetilde{\lambda} \tag{23}$$

• Recall: $y_0 > 0$.

• **If** $h \le 1$ **then** $m_t(h)$ occurs for extremal values of $\lambda_1, \ldots, \lambda_t$ at horizon of uncertainty h: all are either $\widetilde{\lambda}$ or $(1-h)\widetilde{\lambda}$.

• $m_t(h)$, **for** $h \le 1$, is the greater of the following:

$$m_{t,1}(h) = \left(\widetilde{\lambda}^t - \ell^t \right) y_0 \tag{24}$$

$$m_{t,2}(h) \ge \left| \ell^t - \prod_{j=1}^{t} (1-h)\widetilde{\lambda} \right| y_0 \tag{25}$$

$$= \left| \ell^t - (1-h)^t \widetilde{\lambda}^t \right| y_0 \tag{26}$$

- Clearly $m_{t,1}(h) > m_{t,2}(h)$ for small h. To find the transition:

$$\widetilde{\lambda}^t - \ell^t \geq \ell^t - (1-h)^t \widetilde{\lambda}^t \tag{27}$$

$$\Longleftrightarrow \frac{\widetilde{\lambda}^t - 2\ell^t}{\widetilde{\lambda}^t} \geq -(1-h)^t \tag{28}$$

$$\Longleftrightarrow (1-h)^t \geq \frac{2\ell^t - \widetilde{\lambda}^t}{\widetilde{\lambda}^t} \tag{29}$$

$$\Longleftrightarrow 1-h \geq \left(\frac{2\ell^t - \widetilde{\lambda}^t}{\widetilde{\lambda}^t} \right)^{1/t} \tag{30}$$

$$h \leq 1 - \left(\frac{2\ell^t - \widetilde{\lambda}^t}{\widetilde{\lambda}^t} \right)^{1/t} \tag{31}$$

- Hence, **for** $h \leq 1$:

$$m_t(h) = \begin{cases} \left(\widetilde{\lambda}^t - \ell^t\right) y_0, & \text{if } h \leq \underbrace{1 - \left(\frac{2\ell^t - \widetilde{\lambda}^t}{\widetilde{\lambda}^t} \right)^{1/t}}_{h_s} \\ \left(\ell^t - (1-h)^t \widetilde{\lambda}^t\right) y_0, & \text{if } h_s < h \leq 1 \end{cases} \tag{32}$$

  which defines the constant, h_s, at which $m_t(h)$ switches.
- Equate $m_t(h)$ to ε_c and solve for h to find the robustness. One finds:

$$\widehat{h}_t(\ell, \varepsilon_c) = \begin{cases} 0, & \text{if } \varepsilon_c < \left(\widetilde{\lambda}^t - \ell^t\right) y_0 \\ 1 - \left(\frac{\ell^t y_0 - \varepsilon_c}{\widetilde{\lambda}^t y_0} \right)^{1/t}, & \text{if } \left(\widetilde{\lambda}^t - \ell^t\right) y_0 \leq \varepsilon_c \leq \ell^t y_0 \end{cases} \tag{33}$$

  Note that eq.(33) is valid only for values of ε_c for which $\widehat{h}_t \leq 1$.
- This robustness function, like eq.(14), p.4, shows:
  - Discontinuity.
  - Curve-crossing with $\widehat{h}_t(\widetilde{\lambda}, \varepsilon)$.

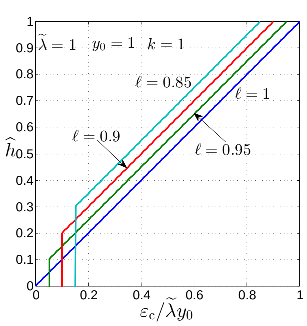

Figure 4: Robustness curves, eq.(33), $t = 1$.

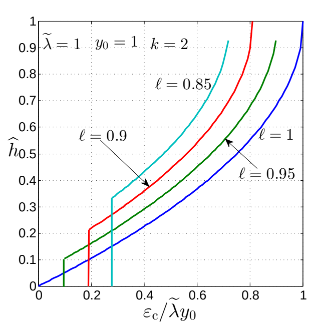

Figure 5: Robustness curves, eq.(33), $t = 2$.

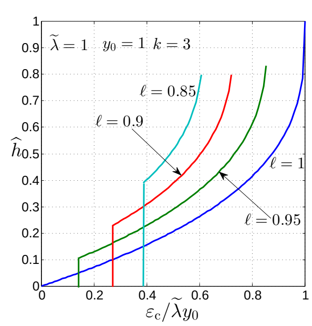

Figure 6: Robustness curves, eq.(33), $t = 3$.

§ **Results: figs. 4–6:**
- Curve crossing.

- Discontinuity of robustness curve occurs at larger $\varepsilon_{\rm c}$ for longer forecast (higher t.)
- Robustness curves shift right (bad) and fall (bad) as t increases. E.g.:

$$\widehat{h}_{t=3}(\varepsilon_{\rm c}=0.4,\ell=0.9)=0.4 \;<\; \widehat{h}_{t=2}(\varepsilon_{\rm c}=0.4,\ell=0.9)=0.45 \;<\; \widehat{h}_{t=1}(\varepsilon_{\rm c}=0.4,\ell=0.9)=0.55 \tag{34}$$

- Cost of robustness decreases (good) as t increases.

2 Regression Prediction of US Inflation Data

§ **Source:** Yakov Ben-Haim, 2010, *Info-Gap Economics: An Operational Introduction,* Palgrave-Macmillan, section 6.1.

2.1 Data

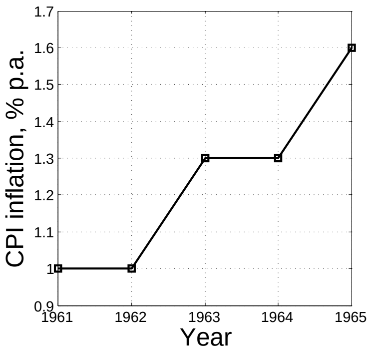

Figure 7: US inflation vs. year, 1961–1965.

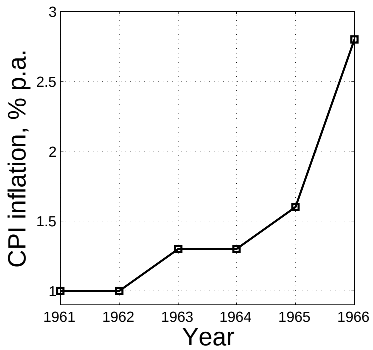

Figure 8: US inflation vs. year, 1961–1966.

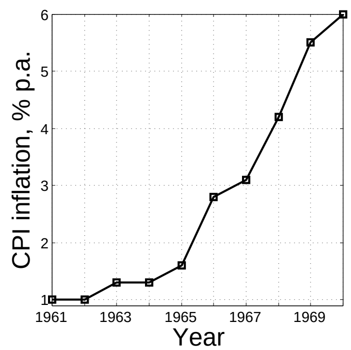

Figure 9: US inflation vs. year, 1961–1970.

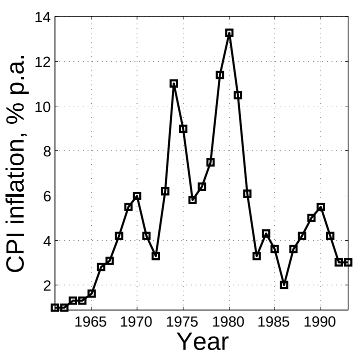

Figure 10: US inflation vs. year, 1961–1993.

§ **US inflation:**

- '61–'65: Linear?
- '61–'66: Quadratic?
- '61–'70: Piece-wise linear?
- '61–'93: A mess?

§ **Modeling and predicting US inflation:**

- '61–'65 Linear? Quadratic?
- Use the '61–'65 model for predicting '66:

$$y_i^{\rm r} = c_0 + c_1 t_i + c_2 t_i^2 \tag{35}$$

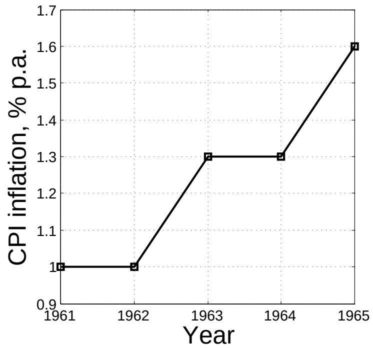

Figure 11: US inflation vs. year, 1961–1965.

2.2 System Model: Mean Squared Error

§ **System model: Mean Squared Error (MSE).**

For any vector of coefficients, c, the MSE is:

$$S_N^2(c) = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i^{\mathrm{r}})^2 \tag{36}$$

$N = 5$ for '61–'65. $y_1, \ldots, y_N$ are data. y_i^{r} is from eq.(35), p.8.

§ **Least-squares estimate** (LSE):

- Definition:

$$\widetilde{c} = \arg\min_{c} S_N^2(c) \tag{37}$$

- Meaning: $\widetilde{c}$ is optimal estimate w.r.t. historical data.
- **Question:** Is $\widetilde{c}$ optimal wrt future data?
- LS regression: $\widetilde{c}$ in y_i^{r} from eq.(35), p.8:

$$\widetilde{y}_i^{\mathrm{r}} = \widetilde{c}_0 + \widetilde{c}_1 t_i + \widetilde{c}_2 t_i^2 \tag{38}$$

- Calculation of LSE of coefficients:

$$\frac{\partial(S_N^2)}{\partial c_k} = 0, \quad k = 1, 2, 3 \tag{39}$$

  3 linear equations in 3 unknowns.
- Does eq.(39) produce a minimum or maximum? Determinantal condition:

$$\left|\frac{\partial^2(S_N^2)}{\partial c_k \partial c_j}\right| > 0 \quad \Longrightarrow \quad \text{minimum not maximum} \tag{40}$$

2.3 Uncertainty Model

§ **Our knowledge:**

- The data: $y_1, \ldots, y_N$
- The LS estimate of the coefficients, $\widetilde{c}$, and the corresponding quadratic function, $\widetilde{y}_i^{\mathrm{r}}$.
- **Contextual info:**
  - ○ Under-prediction by $\widetilde{y}_i^{\mathrm{r}}$ is very likely: y_{N+1} may well exceed the LS prediction, $\widetilde{y}_{N+1}^{\mathrm{r}}$.

∘ Over-prediction by $\widetilde{y}_i^{\mathrm{r}}$ is very unlikely: y_{N+1} will not be less than the LS prediction, $\widetilde{y}_{N+1}^{\mathrm{r}}$.

§ **Info-gap model** of asymmetric uncertainty about LSE $\widetilde{y}_{i\,N+1}^{\mathrm{r}}$:

$$\mathcal{U}(h) = \left\{ y_{N+1} : \ 0 \le y_{N+1} - \widetilde{y}_{N+1}^{\mathrm{r}} \le h \right\}, \quad h \ge 0 \tag{41}$$

- Unbounded family of nested sets.
- No known worst case.
- Depends on the LS coefficients, $\widetilde{c}$.

2.4 Robustness: Formulation and Derivation

§ **If we knew** y_{N+1} ('66):

$$S_{N+1}^2(c) = \frac{1}{N+1} \sum_{i=1}^{N+1} (y_i - y_i^{\mathrm{r}})^2 \tag{42}$$

$$= \frac{N}{N+1} S_N^2(c) + \frac{\left(y_{N+1} - y_{N+1}^{\mathrm{r}} \right)^2}{N+1} \tag{43}$$

§ **Performance requirement.** For any coefficient vector, c, we require:

$$S_{N+1}(c) \le S_{\mathrm{c}} \tag{44}$$

§ **Robustness of regression** c: Greatest tolerable uncertainty.

$$\widehat{h}(c, S_{\mathrm{c}}) = \max \left\{ h : \ \left(\max_{y_{N+1} \in \mathcal{U}(h)} S_{N+1}(c) \right) \le S_{\mathrm{c}} \right\} \tag{45}$$

§ $m(h)$ is inner maximum in eq.(45):

- Inverse of $\widehat{h}(S_{\mathrm{c}})$.
- From S_{N+1} in eq.(43): $m(h)$ occurs when y_{N+1} equals an extreme value at horizon of uncertainty h: either $\widetilde{y}_{N+1}^{\mathrm{r}}$ or $\widetilde{y}_{N+1}^{\mathrm{r}} + h$:

$$m_1(h) = \sqrt{ \frac{N}{N+1} S_N^2 + \frac{\left(\widetilde{y}_{N+1}^{\mathrm{r}} - y_{N+1}^{\mathrm{r}} \right)^2}{N+1} } \tag{46}$$

$$m_2(h) = \sqrt{ \frac{N}{N+1} S_N^2 + \frac{\left(\widetilde{y}_{N+1}^{\mathrm{r}} + h - y_{N+1}^{\mathrm{r}} \right)^2}{N+1} } \tag{47}$$

- $m(h)$ is the greater of these two expressions:

$$m(h) = \max \left[m_1(h), \ m_2(h) \right] \tag{48}$$

- Recall our economic understanding: actual inflation, y_{N+1}, will exceed the LSE value, $\widetilde{y}_{N+1}^{\mathrm{r}}$.
- Hence only consider regressions y_i^{r} for which:

$$\widetilde{y}_{N+1}^{\mathrm{r}} \le y_{N+1}^{\mathrm{r}} \tag{49}$$

- Hence eq.(48) becomes:

$$m(h) = \begin{cases} \sqrt{ \frac{N}{N+1} S_N^2 + \frac{\left(\widetilde{y}_{N+1}^{\mathrm{r}} - y_{N+1}^{\mathrm{r}} \right)^2}{N+1} } & \text{if } h < 2 \left(y_{N+1}^{\mathrm{r}} - \widetilde{y}_{N+1}^{\mathrm{r}} \right) \\ \sqrt{ \frac{N}{N+1} S_N^2 + \frac{\left(\widetilde{y}_{N+1}^{\mathrm{r}} + h - y_{N+1}^{\mathrm{r}} \right)^2}{N+1} } & \text{if } h \ge 2 \left(y_{N+1}^{\mathrm{r}} - \widetilde{y}_{N+1}^{\mathrm{r}} \right) \end{cases} \tag{50}$$

- Thus $m(h)$ may switch between the two functions and display discontinuity of slope.
- Recall: $m(h)$ is the inverse of the robustness function.

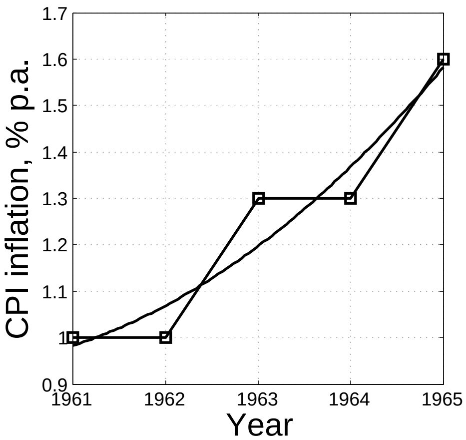

Figure 12: US inflation vs. year, 1961–1965, and least squares fit.

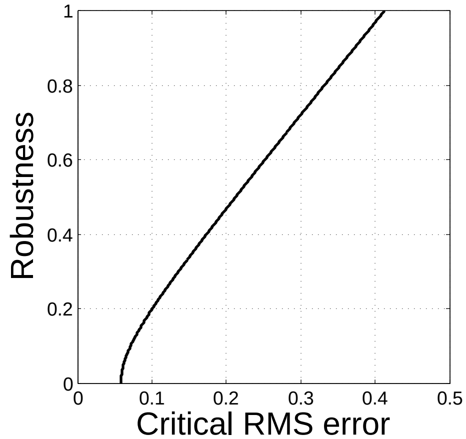

Figure 13: Robustness vs. critical root mean squared error for inflation 1961–1965.

2.5 Robustness: Results

§ **Least squares fit:** fig. 12: Maximal fidelity of quadratic function to the data.

§ **Robust of LS fit:** fig. 13.

- Trade off: Greater rbs. $\equiv$ greater critical RMS error, $S_{\rm c}$.
- Zeroing: No robustness of estimated RMS error, $S_{\rm c}$.
- What do the numbers mean?
  - $\widehat{h} = 0.2$ at $S_{\rm c} = 0.1$:

    $\widetilde{y}^{\rm r}_{i\,N+1}$ can err by as much as 0.2 (from info-gap model, eq.(41), p.10)

    if we require that

    S^2_{N+1} can err by no more than 0.1 (from performance requirement, eq.(44), p.10).
  - $\widehat{h} = 0.7$ at $S_{\rm c} = 0.3$:

    $\widetilde{y}^{\rm r}_{i\,N+1}$ can err by as much as 0.7 (from info-gap model, eq.(41), p.10)

    if we require that

    S^2_{N+1} can err by no more than 0.3 (from performance requirement, eq.(44), p.10).

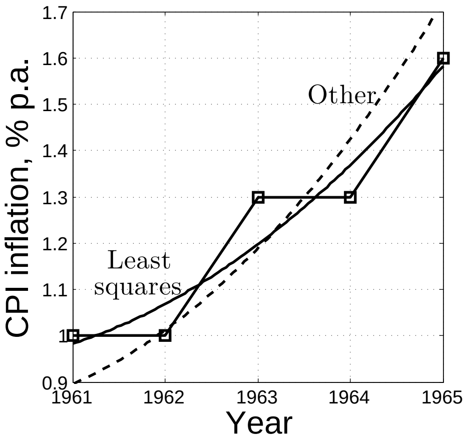

Figure 14: US inflation vs. year, 1961–1965, and least squares fit (solid) and other fit (dash).

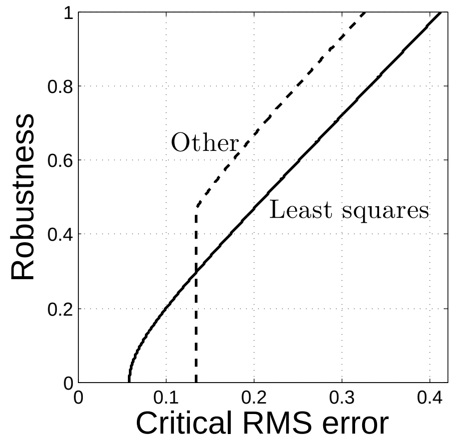

Figure 15: Robustness vs. critical root mean squared error for inflation 1961–1965 for least squares fit (solid) and other fit (dash).

§ **Least squares and other fit:** fig. 14.

- LS fit: Maximal average fidelity of quadratic function to the data.
- Other fit. Biased fidelity:
  - ○ Under-estimate, on average, of early data. **Note:** 1962 is almost exact!
  - ○ Over-estimate, on average, of late data.
  - ○ Over-estimate of future wrt historical trend. **Maybe** 1966 will be exact!

§ **Robust of LS and other fit:** fig. 15.

- Zeroing: “Other” zeros at greater $S_{\rm c}$: it’s nominal MSE is worse.
- Trade off: Both curves. ‘Other’ has greater (infinite) slope at zeroing value: lower cost of robustness.
- Curve-crossing: **preference reversal.**
- What do the numbers mean?
  - ○ At $S_{\rm c} = 0.1$:
  - — $\widehat{h}_{\rm other} = 0$. $\widehat{h}_{\rm LS} = 0.2$. Forecast y_{N+1} with $\widetilde{y}^{\rm r}_{N+1}(\widetilde{c})$ if $S_{\rm c} = 0.1$ is adequate (or required).
  - ○ At $S_{\rm c} = 0.3$:
  - — $\widehat{h}_{\rm other} = 0.8$. $\widehat{h}_{\rm LS} = 0.6$. Forecast y_{N+1} with $y^{\rm r}_{N+1}(c)$ if $S_{\rm c} = 0.3$ is adequate (or required).
  - — Curve-crossing: **preference reversal.**
  - — **Why** forecast y_{N+1} with $y^{\rm r}_{N+1}(c)$ rather than with $\widetilde{y}^{\rm r}_{N+1}(\widetilde{c})$?
    - · Fidelity to data and forecast is our measure of performance of a forecaster, eq.(44), p.10.
    - · $y^{\rm r}_{N+1}(c)$ gives adequate fidelity ($S_{\rm c} = 0.3$) over wider range of uncertainty than $\widetilde{y}^{\rm r}_{N+1}(\widetilde{c})$.

3 Auto-Regression and Data Revision

§ **Source:** Yakov Ben-Haim, 2010, *Info-Gap Economics: An Operational Introduction,* Palgrave-Macmillan, section 6.2.

3.1 The Problem of Data Revision

§ **National statistical bureaus revise economic data over time.**

• **1974:**

◦ Real US GNP initially thought to have dropped 9.1% at annual rate between 3rd and 4th quarters.

◦ Largest drop since great depression.

◦ Final estimate, 20 years later: real GNP dropped 1.9% at annual rate.

◦ Not all revisions are this large.

◦ This revision large because of great economic turbulence then.

◦ Precisely in times of economic uncertainty we need accurate data.

• **2007–2009:**

◦ Typical revisions of 1 or 2 percentage points.

◦ Table 2.

	7q1	7q2	7q3	7q4	8q1
current	1.2	3.2	3.6	2.1	−0.7
previous	0.1	4.8	4.8	−0.2	0.9
	8q2	8q3	8q4	9q1	9q2
current	1.5	−2.7	−5.4	−6.4	−1.0
previous	2.8	−0.5	−6.3	−5.5	

Table 2: Current and previous estimates of real GDP: percent change from preceding period. 2007q1 to 2009q2. Seasonally adjusted at annual rates. Bureau of Economic Analysis, July 31, 2009.

3.2 Autoregression

§ N **scalar data points:** $y = (y_1, \ldots, y_N)^T$.

E.g. inflation data over N sequential years as in fig. 7 on p.8.

§ **Regression:**

Choose coefficients $c = (c_1, \ldots, c_J)^T$ of an auto-regression of order J for these data:

$$y_n = \sum_{j=1}^{J} c_j y_{n-j} \tag{51}$$

$$= c^T y_{n-1,n-J} \tag{52}$$

where $y_{n-1,n-J} = (y_{n-1}, \ldots, y_{n-J})^T$.

§ **Define mean squared error of the auto-regression (AR) of the data:**

$$S^2(c) = \frac{1}{N-J} \sum_{n=J+1}^{N} \left(y_n - c' y_{n-1,n-J}\right)^2 \tag{53}$$

§ Our **system model** is the RMS error, $S(c)$.

§ **Performance requirement:**

$$S(c) \le S_{\rm c} \tag{54}$$

§ **The mean squared error can be expressed more compactly as:**

$$S^2(c) = \frac{1}{N-J} y' V y \tag{55}$$

where V is defined as follows.

• e_n denotes the nth standard basis vector in $\Re^N$: the column N-vector with a 1 in the nth location and 0's elsewhere.

• Now the mean squared error can be written:

$$\begin{aligned} S^2(c) &= \frac{1}{N-J} \sum_{n=J+1}^{N} \left[e_n^T y - \sum_{j=1}^{J} c_j e_{n-j}^T y \right]^2 & (56) \\ &= \frac{1}{N-J} \sum_{n=J+1}^{N} \left[\underbrace{\left(e_n^T - \sum_{j=1}^{J} c_j e_{n-j}^T \right)}_{\zeta_n^T} y \right]^2 & (57) \\ &= \frac{1}{N-J} \sum_{n=J+1}^{N} y^T \zeta_n \zeta_n^T y & (58) \\ &= \frac{1}{N-J} y^T \underbrace{\left(\sum_{n=J+1}^{N} \zeta_n \zeta_n^T \right)}_{V} y & (59) \\ &= \frac{1}{N-J} y^T V y & (60) \end{aligned}$$

This is eq.(55), with $N \times N$ matrix V from eq.(59) and N-vectors ζ_n from eq.(57).

• V depends on the regression coefficients c but not on the data.

§ The AR coefficients that **minimize the mean squared error** are found by solving:

$$\frac{\partial S^2}{\partial c} = 0 \tag{61}$$

• Differentiating eq.(53) and rearranging one finds:

$$\underbrace{\sum_{n=J+1}^{N} y_n y_{n-1,n-J}}_{z} = \underbrace{\left(\sum_{n=J+1}^{N} y_{n-1,n-J}\, y'_{n-1,n-J} \right)}_{Y} c \tag{62}$$

which defines the J-vector z and the $J \times J$ matrix Y.

• The least squares (LS) auto-regression coefficients are:

$$\widetilde{c} = Y^{-1} z \tag{63}$$

If the inverse matrix does not exist then a generalized inverse needs to be used.

3.3 Uncertainty Model and Robustness Function

§ **Best estimate of the data:** $\widetilde{y} = (\widetilde{y}_1, \ldots, \widetilde{y}_N)^T$.

E.g. $\widetilde{y}$ might be current estimates of percent change in real GDP shown in table 2, p.13.

§ **Info-gap model for asymmetric information:**

$$\mathcal{U}(h) = \left\{ y : \ \widetilde{y}_n - w_{n1} h \le y_n \le \widetilde{y}_n + w_{n2} h, \ n = 1, \ldots, N \right\}, \quad h \ge 0 \tag{64}$$

- Uncertainty weights, w_{n1} and w_{n2}, are non-negative.
- If $\widetilde{y}_n$ is certain, then $w_{n1} = 0 = w_{n2}$.
- If $\widetilde{y}_n$ is believed to be an **underestimate** then $w_{n1} = 0$ and $w_{n2} = 1$.
- If $\widetilde{y}_n$ is believed to be an **over estimate** then $w_{n1} = 1$ and $w_{n2} = 0$.
- If the uncertainty is **symmetric** then uncertainty weights $w_{n1} = w_{n2}$.

§ **Robustness function, definition:**

$$\widehat{h}(c, S_{\rm c}) = \max \left\{ h : \ \left(\max_{y \in \mathcal{U}(h)} S(c) \right) \le S_{\rm c} \right\} \tag{65}$$

3.4 Policy Exploration

§ **Example** is based on 2nd-order auto-regressions, so $J = 2$ in eq.(51), p.13.

We use the percent change in the US GDP for 2007q1–2009q2 in table 2, p.13.

3.4.1 Symmetric Uncertainty

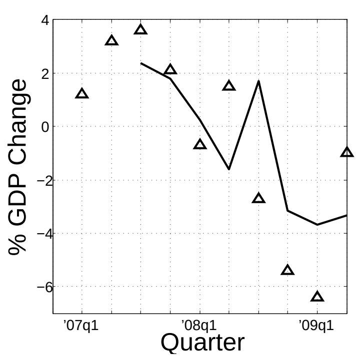

Figure 16: Current estimates of US real GDP change vs. quarter, and least squares auto-regression.

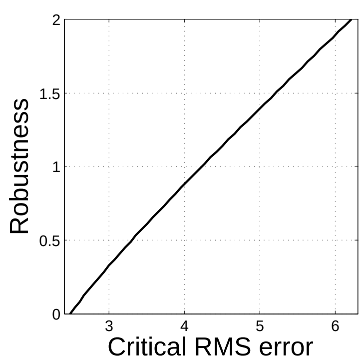

Figure 17: Robustness vs. critical RMS error of least squares auto-regression. Symmetric uncertainty.

§ **The data.**

- Fig. 16: GDP data, table 2, p.13, and 2nd-order least-squares auto-regression, $\widetilde{c}$ from eq.(63).
- LS regression coefficients are $\widetilde{c} = (0.9139, -0.4647)^T$.
- The RMS error of this regression is $S(\widetilde{c}) = 2.49$, so the AR misses the data, on average, by about 2.5 percentage points of GDP.
- This rather large error occurs mostly in last 5 quarters: data at '08q2 and '09q2 deviate from trend.

- Great uncertainty; no information on direction or magnitude of data revision.
- Use info-gap model of eq.(64) with $w_{n1} = w_{n2} = 1$ for $n = 1, \ldots, N$.

§ **Robustness curve for LS AR $\widetilde{c}$ with symmetric uncertainty,** fig. 17.

- Zeroing at $S(\widetilde{c}) = 2.49$.
- Trade off.
- $\widehat{h}(S_c = 4) = 0.88$: RMS error no larger than 4% is guaranteed with robustness of 0.88: revisions as large as 0.88 percentage points can occur and the RMS error will not exceed 4%.

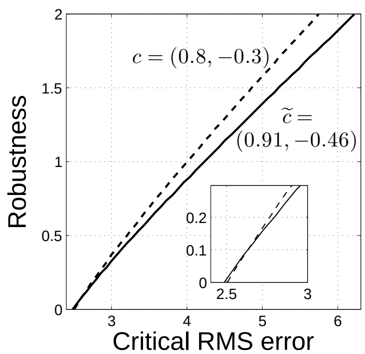

Figure 18: Robustness vs. critical RMS error with least squares (solid) and other (dash) autoregression. Symmetric uncertainty.

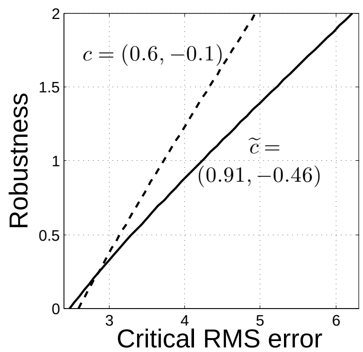

Figure 19: Robustness vs. critical RMS error with least squares (solid) and other (dash) autoregression. Symmetric uncertainty.

§ **Robustness curves for non-LS AR with symmetric uncertainty,** figs. 18 and 19.

- LS robustness curve (solid) reproduced from fig. 17.
- LS robustness curve zeros to left of non-LS by definition of **least** squares.
- Non-LS robustness curves steeper: lower cost of robustness.
- Curve crossing and preference reversal.
- Two foci of uncertainty:
  - ○ Statistical: seek small RMS.
  - ○ Info-gap: seek large robustness.

3.4.2 Asymmetric Uncertainty

§ **Uncertainty and contextual information:**

- Data in fig. 16, p.15.
- Current estimates at 2008q2 and 2009q2 are over-estimates and will be revised down.
- Use info-gap model of eq.(64). Choose uncertainty weights:
  - ○ $w_{6,2} = w_{10,2} = 0$: 6th and 10th estimates **cannot go up.**
  - ○ $w_{6,1} = w_{10,1} = 1$: 6th and 10th estimates **can go down.**
  - ○ $w_{nj} = 1$ for all other n and j (all other estimates can go either up or down).
  - In summary:

$$w_2 = [1\ 1\ 1\ 1\ 1\ \ 0\ 1\ 1\ 1\ 0] \quad (66)$$

$$w_1 = [1\ 1\ 1\ 1\ 1\ \ 1\ 1\ 1\ 1\ 1] \quad (67)$$

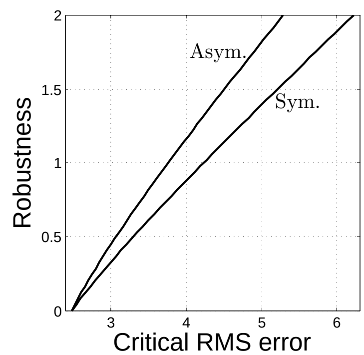

Figure 20: Robustness vs. critical RMS error for least squares regression with symmetric and asymmetric uncertainty.

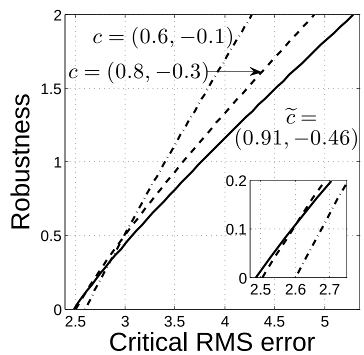

Figure 21: Robustness vs. critical RMS error with least squares (solid) and other (dash, dot-dash) regressions. Asymmetric uncertainty.

§ **LS auto-regression,** fig. 20:

- $\widetilde{c}$ does not depend on the info-gap model, so $\widetilde{c}$ is the same as before: $\widetilde{c} = (0.9139, -0.4647)^T$.
- Furthermore, the RMS error of the LS regression same as before: $S(\widetilde{c}) = 2.49$.
- However, the robustness of $\widetilde{c}$ *does* depend on the info-gap model, fig. 20.
  - ○ Zeroing: both curves reach $S_{\rm c}$ axis at $S(\widetilde{c})$.
  - ○ Asymmetric robustness curve higher due to greater info in asymmetric info-gap model.

§ **Non-LS auto-regression,** fig. 21:

- Solid curve is LS regression: the "Asym." curve from fig. 20.
- Curve-crossing with non-LS regressions.
- Large robustness gain by the non-LS over LS regressions.
- Compare with figs. 18 and 19:
  - ○ Robustness gain is greater in current case.
  - ○ Added asymmetric information enhances robustness of LS regression.
  - ○ Further enhances the robustness of these non-LS regressions.