

A Human-Machine Ensemble Method for Economic Forecasts

経済予測に関する人間と機械のアンサンブル法の考案

三吉 貴大

社会情報ネットワーク講座・広域情報ネットワーク分野

2017 年 2 月 14 日

アウトライン

- ① はじめに
- ② 提案手法
- ③ 実験
- ④ 結果
- ⑤ 考察
- ⑥ まとめ

アウトライン

- ① はじめに
背景
関連研究
研究の目的

- ② 提案手法

- ③ 実験

- ④ 結果

- ⑤ 考察

- ⑥ まとめ

経済予測

経済指標

- GDP やインフレなどの経済指標は政策立案者や企業、投資家にとって重要
- マクロ経済モデルや時系列解析等の従来手法では、1 年程度の短期予測は困難

集合知による予測

- e.g. 質問紙調査の集計値
- 景気や政策などの情報を総合的に考慮して予測を行う
- 多様性を大きくすることで誤差の期待値を小さくできる [3]

機械学習による予測

- e.g. 再帰ニューラルネット
- 過去の時系列から統計的に予測器を構築し予測を行う
- 過去のデータを元に誤差の期待値を定量化できる

これらを組み合わせて、より正確な予測を行う

予測の組み合わせ

アンサンブル法 [4]

- 予測の集約には、単純平均や中央値、重み平均が一般に用いられる
- 重み平均では、過去の予測が正確だった予測者・予測器の重みを大きくする
 - 予測者 i の予測値を h_i 、重みを w_i とすると重み平均は $\sum_i w_i h_i$
- 経済予測では、必ずしもこれらの方法で予測精度が向上できるわけではない [1]

問題

- 入力に関わらず組み合わせ方（重み）が一定
→ 予測者・予測器ごとの得手不得手の違いを活かすことができない

入力に応じて、組み合わせ方を変化させる

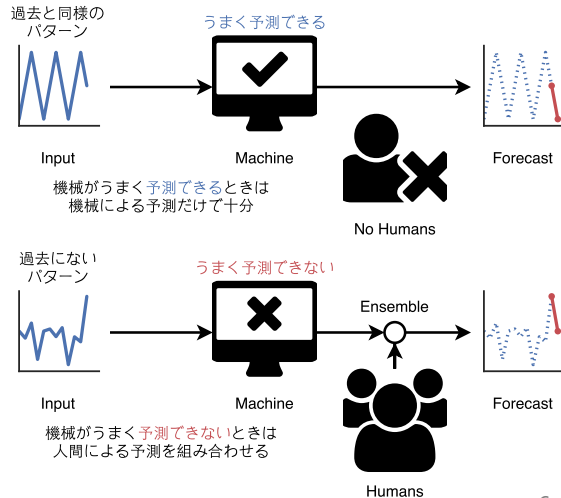
研究の目的

目的

人間と機械の予測をうまく組み合わせることで、より正確な経済予測を行う

研究課題

1. 人間と機械の予測方法の違いを反映したモデル化と、その最適な組み合わせ方の考案
2. 考案したアンサンブル法を用いた経済予測の実装と評価



アウトライン

① はじめに

② 提案手法

人間と機械の予測のモデル化

アンサンブルの誤差の期待値の最小化

③ 実験

④ 結果

⑤ 考察

⑥ まとめ

人間による予測のモデル

人間による予測のモデル [2]

予測者 i の予測値 h_i と誤差 ε_{h_i} を以下のパラメータに従う確率変数とする

- $\text{var}(\varepsilon_h)$: 個人の予測の分散
- $\text{cov}(\varepsilon_h)$: 異なる 2 人の誤差の共分散
- 予測値に偏りはないとする

人間の特徴

- 容易に人数 n を大きくできる
- $\text{cov}(\varepsilon_h)$ が小さい = 多様性が大きい

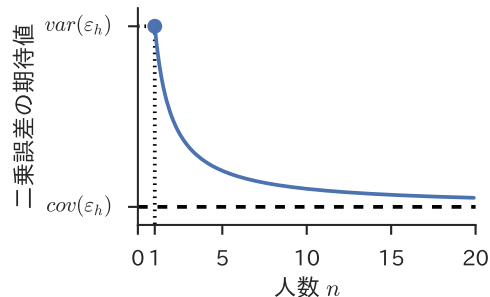


図: n 人の予測の平均値における二乗誤差の期待値と人数 n の関係. n を大きくするほど二乗誤差の期待値を小さくできる

機械による予測のモデル

機械の出力として事後分布の導入

予測器 θ は入力 x に対して確率分布 $f_{\theta}(y|x)$ を出力する

- 分布の平均を予測値 $y_{\theta}(x)$ とする
- 分布の分散 $\text{var}(\varepsilon_{\theta}|x)$ は二乗誤差の期待値を表す

機械の特徴

- 入力 x に応じた誤差の期待値をモデルに基づいて定量化できる

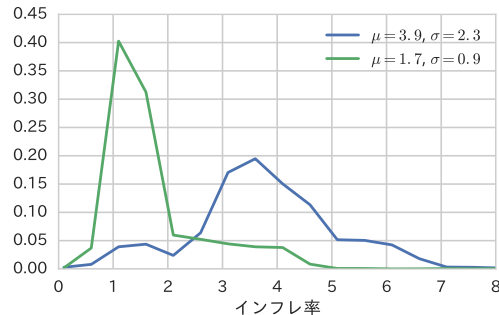


図: ニューラルネットによる予測器が出力した確率分布の例. 青より緑の方が分散が小さいため, 誤差の期待値が小さいと考えられる

Human-Machine Ensemble

問題の定式化

n 人の人間 h_1, \dots, h_n と一つの機械 θ による予測値の平均をアンサンブルの予測値 $Y_{\theta,h}(n|\mathbf{x})$ とする

$$Y_{\theta,h}(n|\mathbf{x}) = \frac{y_{\theta}(\mathbf{x}) + \sum_{i=1}^n h_i}{n+1} \quad (1)$$

この二乗誤差の期待値 $\text{MSE}[Y_{\theta,h}(n|\mathbf{x})]$ を最小化する人数 $n^* \geq 0$ を求める

$$\text{MSE}[Y_{\theta,h}(n|\mathbf{x})] = \frac{n \text{var}(\varepsilon_h) + \text{var}(\varepsilon_{\theta}|\mathbf{x}) + n(n-1) \text{cov}(\varepsilon_h) + 2n \text{cov}(\varepsilon_{\theta}, \varepsilon_h)}{(n+1)^2} \quad (2)$$

- $\text{cov}(\varepsilon_{\theta}, \varepsilon_h)$: 機械と人間の予測の誤差の共分散

最適な人数

極小値を取る条件

式 (2) が極小値を持てば、最適な n^* は必ず存在する
極小値を持つ条件は

$$\text{cov}(\varepsilon_\theta, \varepsilon_h) < \frac{3 \text{cov}(\varepsilon_h) - \text{var}(\varepsilon_h)}{2} \quad (3)$$

実用において

式 (3) が成り立たない場合は人数の上限 N_{\max} を設定すればよい

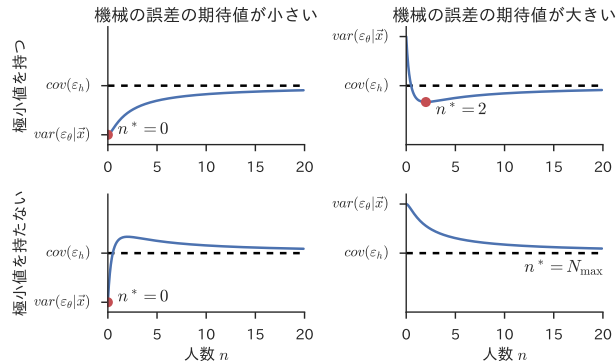


図: アンサンブルの二乗誤差の期待値と人数 n の関係

アウトライン

① はじめに

② 提案手法

③ 実験

実験の目的
実験の方法

④ 結果

⑤ 考察

⑥ まとめ

実験の目的

モデルの検証

- 分布の分散を二乗誤差の期待値とみなせるといった仮定が現実の問題でも満たされるか検証する

実問題における予測精度の確認

- $\text{var}(\varepsilon_h)$, $\text{cov}(\varepsilon_h)$, $\text{cov}(\varepsilon_\theta, \varepsilon_h)$ の真の値を実際に知ることはできず、過去のサンプルから推定するしかない
- サンプル外で、推定したパラメータの値の通りに振る舞うとは限らない

提案したアンサンブル法の振る舞いの観察

- 過去にないパターンが与えられた時に人間による予測を用いるという仮説の検証
- 2008 年のリーマンショックを用いる

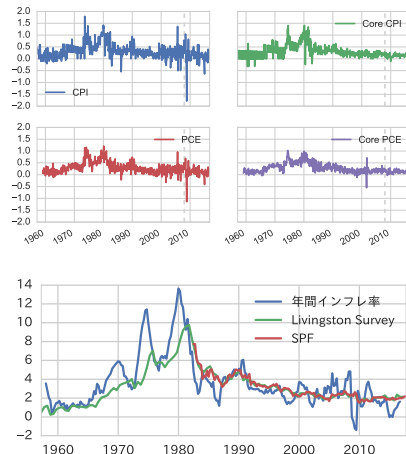
データ

予測対象: インフレに関する 4 つの指標

- 消費者物価指数 (CPI), Core CPI (1957～)
- 個人消費支出 (PCE), Core PCE (1959～)

人間の予測: 経済予測のサーベイ調査

- Livingston survey (LIV)
 - 1946 年から 6 ヶ月ごと
 - CPI のみ, 6 ヶ月後と 12 ヶ月後の予測
- Survey of Professional Forecasts (SPF)
 - 1981 年から四半期ごと
 - 全ての指標, 1 年後の予測のみ



予測手法

ベンチマーク [1]

- 従来の時系列解析の手法より，自己回帰移動平均モデル ARMA(1,1)

機械

- 再帰ニューラルネットワーク (RNN)
 - 入力: 過去 12 ヶ月の変化 / 出力: 目的の値の離散確率分布

人間

- LIV と SPF の各サーベイ
 - 人数を統一するため各調査からランダムに 5 人をサンプルし，その平均値を取る

提案手法

- 機械 (RNN) と人間 (LIV, SPF) の組み合わせにより構成

評価方法

データセットの分割

- 2008 年以前を訓練セット，それ以降をテストセットとする
 - 比較のため，CPI のみ 1998 年を境界とした場合も作成
- 評価にはテストセットのみを用いる
 - モデルの作成やパラメータの推定にはテストセットは用いない

予測精度の評価指標

- Root Mean Squared Error (RMSE) を用いる

$$\text{RMSE}(\hat{\mathbf{y}}) = \sqrt{\frac{1}{M} \sum_{i=1}^M (y_i - \hat{y}_i)^2}$$

- ベンチマークの ARMA(1,1) の RMSE を 1 とした相対値を示す

アウトライン

① はじめに

② 提案手法

③ 実験

④ 結果

モデルの検証

予測精度

提案手法の振る舞い

⑤ 考察

⑥ まとめ

モデルの検証（分布の分散と二乗誤差の関係）

仮定

機械が出力した分布の分散を二乗誤差の期待値とみなせる

結果

バラつきは大きいですが、おおよそ原点を通る傾き 1 の直線に沿って分布していた

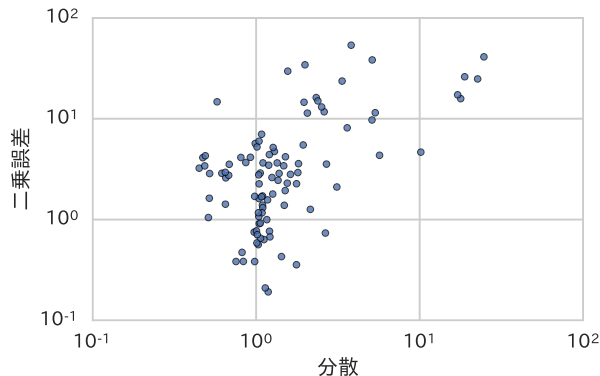


図: テストセットにおける分布の分散と二乗誤差の関係 (CPI-LIV)

予測精度

表: 相対 RMSE

	CPI-LIV	CPI-SPF	CoreCPI	PCE	CorePCE	CPI-6M	CPI-1998
ベンチマーク	1.000	1.000	1.000	1.000	1.000	1.000	1.000
機械	0.889	0.889	0.788	0.933	0.842	1.010	0.941
人間	0.704	0.736	0.767	0.711	0.848	0.939	0.940
提案手法	0.689	0.736	0.767	0.724	0.827	0.931	0.755

結果

- 提案手法は7つのうち4つで最も精度が高く、2つで既存手法と同じになった
- CPI-SPF と CoreCPI では提案手法は人間による予測のみを用いた

提案手法の振る舞い

想定シナリオ

過去にない系列が機械に与えられたとき人間による予測を用いる

結果

2008年9月のリーマンショック後、機械の誤差が大きいとき人間による予測をうまく利用していた

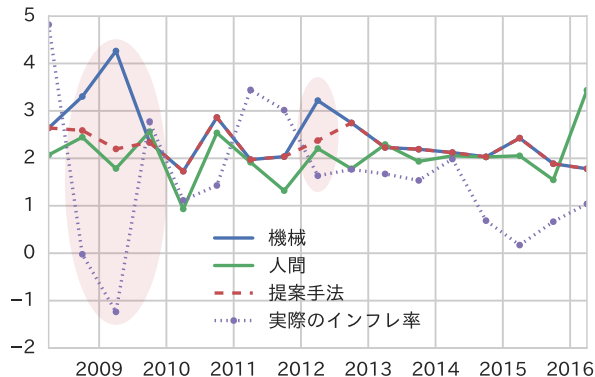


図: テストセットでの年間インフレ率予測 (CPI-LIV)

アウトライン

① はじめに

② 提案手法

③ 実験

④ 結果

⑤ 考察

提案手法の適用範囲
今後の課題

⑥ まとめ

Human-Machine Ensemble Method の適用範囲

Human-Machine Ensemble を有効に使う条件

1. 予測対象は数値など誤差を定義できるものである
2. 機械が入力に応じた誤差の期待値を出力できる
3. 人間の母集団について $\text{var}(\varepsilon_h)$, $\text{cov}(\varepsilon_h)$ が推定できる
4. 機械の方が良い場合もあれば、人間の方が良い場合もある
 - 人間と機械の多様性が人間同士の多様性よりも大きい ($\text{cov}(\varepsilon_\theta, \varepsilon_h) < \text{cov}(\varepsilon_h)$)

表: 各アンサンブルにおけるパラメータ. 青字は $\text{cov}(\varepsilon_\theta, \varepsilon_h) < \text{cov}(\varepsilon_h)$ が不成立

	CPI-LIV	CPI-SPF	CoreCPI	PCE	CorePCE	CPI-6M	CPI-1998
$\text{cov}(\varepsilon_h)$	1.846	0.873	0.455	0.450	0.168	0.652	1.924
$\text{cov}(\varepsilon_\theta, \varepsilon_h)$	1.772	1.096	0.668	0.510	0.056	0.497	1.821

今後の課題

提案手法の応用

実際には、人間の予測にデータの蓄積がある専門家の予測が使えるとは限らない

→ 人間の予測にクラウドソーシングを用いる

- 一般人を対象にした Michigan survey も LIV や SPF と同程度の予測精度がある
- EM アルゴリズム等を用いてパラメータ $\text{var}(\varepsilon_h)$, $\text{cov}(\varepsilon_h)$ を推定する必要がある

提案手法の拡張

今回は、同一のパラメータに従う複数の人間と一つの機械の予測値を単純平均した

→ 組み合わせ方を拡張してより正確な予測を目指す

- 機械の数も可変にする
- 人間を経済学者と金融アナリストなど複数のタイプに分ける
- 単純平均以外の集約方法を用いる

アウトライン

- ① はじめに
- ② 提案手法
- ③ 実験
- ④ 結果
- ⑤ 考察
- ⑥ **まとめ**

まとめ

研究の目的

人間と機械の予測をうまく組み合わせることで、より正確な経済予測を行う

本研究の貢献

1. 人間と機械の予測方法の違いを反映したモデル化と、最適な組み合わせ方の考案
 - 事後分布を用いて入力ごとに誤差の期待値が異なる予測をモデル化した
 - そのモデルの元で誤差の期待値を最小化する人数を求める方法を提示した
2. 考案したアンサンブル法を用いた経済予測の実装と評価
 - 実際のインフレ予測に提案したモデル・手法が適用可能であることを示した
 - 7つのデータ中4つで予測精度の改善を確認した

参考文献

- [1] Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?, Journal of Monetary Economics, Vol. 54, No. 4, pp. 1163–1212 (2007).
- [2] Optimal Forecasting Groups, Management Science, Vol. 58, No. 4, pp. 805–810 (2012).
- [3] : The Defference, Princeton University Press (2008).
- [4] : Combination Method, CRC Press, chapter 4, pp. 67–98 (2012).

パラメータの推定

$$\text{var}(\varepsilon_h) = \frac{1}{N} \sum_{i=1}^N \text{var}(\varepsilon_{h_i})$$

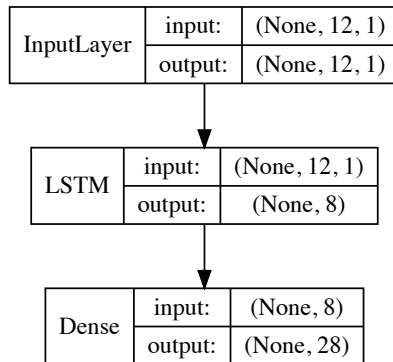
$$\text{cov}(\varepsilon_h) = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j \neq i}^N \text{cov}(\varepsilon_{h_i}, \varepsilon_{h_j})$$

$$\text{cov}(\varepsilon_\theta, \varepsilon_h) = \frac{1}{N} \sum_{i=1}^N \text{cov}(\varepsilon_\theta, \varepsilon_{h_i})$$

- $\text{var}(\varepsilon_{h_i})$: 予測者 i の予測値の分散
- $\text{cov}(\varepsilon_{h_i}, \varepsilon_{h_j})$: 予測者 i と j の誤差の共分散
- $\text{cov}(\varepsilon_\theta, \varepsilon_{h_i})$: 予測器 θ と予測者 i の誤差の共分散

実験で使った RNN のアーキテクチャ

- “None” にはミニバッチのサンプル数 32 が入る
- 出力層の次元 “28” は 12 ヶ月後の CPI を予測する場合



実験で使った RNN の出力層のラベル

- 訓練データにおける最小値 (min) から最大値 (max) までを step の値で区切る
- 出力層の活性化関数にソフトマックス関数を用いて離散確率分布を出力させる

表: 出力層におけるラベルの最小値と最大値, 個数, ラベル間の値の差

	CPI(12M)	CPI(6M)	CoreCPI	PCE	CorePCE
min	0.1	-0.4	0.7	0.5	0.9
max	13.6	7.4	12.7	11.0	9.9
size	28	27	25	22	19
step	0.5	0.3	0.5	0.5	0.5

実験で使⽤したアンサンブルの種類

表: 実験で使⽤したアンサンブルの種類

	CPI-LIV	CPI-SPF	CoreCPI	PCE	CorePCE	CPI-6M	CPI-1998
target index	CPI	CPI	CoreCPI	PCE	CorePCE	CPI	CPI
survey source	LIV	SPF	SPF	SPF	SPF	LIV	LIV
forecast period	12M	12M	12M	12M	12M	6M	12M
dataset boundary	2008	2008	2008	2008	2008	2008	1998

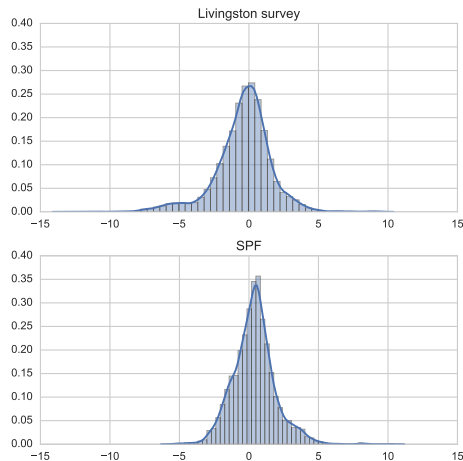
モデルの検証（人間の予測誤差の分布）

仮定

人間による予測値に偏りはないと仮定した

結果

0 を中心に偏りなく分布している
(LIV, SPF の平均値はそれぞれ -0.40 , 0.37)



ARMA(1,1) で正規化する前の RMSE

表: 絶対 RMSE

	CPI-LIV	CPI-SPF	CoreCPI	PCE	CorePCE	CPI-6M	CPI-1998
ベンチマーク	2.234	2.234	0.603	1.698	0.546	1.360	1.740
機械	1.985	1.985	0.475	1.585	0.460	1.374	1.637
人間	1.572	1.644	0.462	1.207	0.463	1.277	1.635
提案手法	1.540	1.644	0.462	1.230	0.451	1.265	1.314

提案手法の振る舞い（詳細）

左から

- 機械の二乗誤差
- 機械の出力した分布の分散
- 提案手法で用いられた人数 n
- 提案手法の二乗誤差

式 (3) が成り立たないため

n が 0 か 5(= N_{\max}) しか取れない

	machine error	variance	n	ensemble error
Jun 2008	4.74	1.53	0	4.74
Dec 2008	11.06	2.15	5	6.81
Jun 2009	30.23	2.81	5	11.80
Dec 2009	0.19	1.63	0	0.19
Jun 2010	0.38	0.90	0	0.38
Dec 2010	2.06	1.72	0	2.06
Jun 2011	2.16	1.15	0	2.16
Dec 2011	0.96	1.22	0	0.96
Jun 2012	2.51	2.12	5	0.55
Dec 2012	0.96	1.65	0	0.96
Jun 2013	0.31	1.35	0	0.31
Dec 2013	0.43	1.33	0	0.43
Jun 2014	0.02	1.28	0	0.02
Dec 2014	1.81	1.21	0	1.81
Jun 2015	5.08	1.45	0	5.08
Dec 2015	1.49	1.07	0	1.49
Jun 2016	0.55	0.96	0	0.55

極小値を持つ場合の人工的な例

表: $\text{var}(\varepsilon_h) = 3$, $\text{cov}(\varepsilon_h) = 2$, $\text{cov}(\varepsilon_\theta, \varepsilon_h) = 0$ として生成したサンプル

	ε_{h_1}	ε_{h_2}	ε_θ
0	0.1	1.0	1.7
1	-0.5	0.3	1.3
2	-2.1	-2.5	2.9
3	-0.3	2.5	-1.2
4	-2.4	-3.3	0.8
5	-0.4	0.1	2.6
6	-0.2	0.5	1.6
7	1.6	1.0	0.7
8	0.3	-0.4	-0.9
9	0.9	1.9	1.9