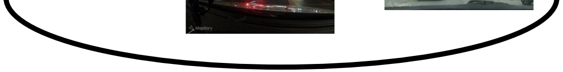
Dpto. Informática e Ingeniería de Sistemas.

Images: Wikipedia, Nando de Freitas.

http://www.cs.ox.ac.uk/people/yarin.gal/website/blog_2248.html


f_2
x^*
x



I. J. Goodfellow, J. Shlens, and C. Szegedy, Explaining and harnessing adversarial examples, *arXiv:1412.6572*, 2014
https://openai.com/blog/multimodal-neurons/

Image: Yarin Gal

Image: Yarin Gal

Image: Yarin Gal

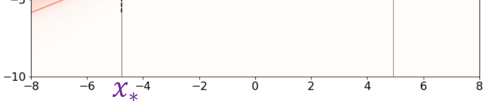
Image: Yarin Gal

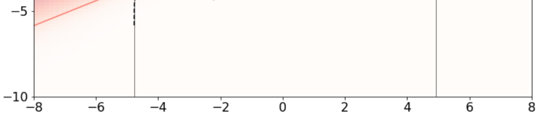
$p(D')$

$$p(w) = N(w|w_0, \Sigma_0)$$

with: $\lambda = \frac{\sigma^2}{\tau^2}$ ...which is again ridge regression!

Universidad Zaragoza 1542


−10
−8
−6
−4
−2
0
2
4
6
8
x_*

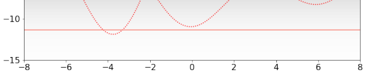

−5
−10
−8
−6
−4
−2
0
2
4
6
8

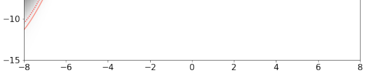
reducible by additional training data

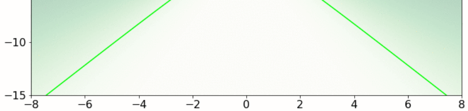
Universidad Zaragoza 1542

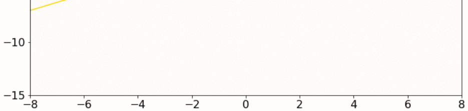
Alex Kendall and Yarin Gal. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?, NeurIPS, 2017 https://papers.nips.cc/paper_files/paper/2017/hash/2650d6089a6d640c5e85b2b88265dc2b-Abstract.html

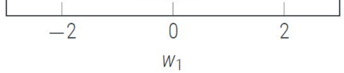
F. K. Gustafsson, M. Danelljan, T. B. Schon, Evaluating scalable Bayesian deep learning methods for robust computer vision, CVPRW. 2020.
https://openaccess.thecvf.com/content_CVPRW_2020/papers/w20/Gustafsson_Evaluating_Scalable_Bayesian_Deep_Learning_Methods_for_Robust_Computer_Vision_CVPRW_2020_paper.pdf


−10
−15
−8 −6 −4 −2 0 2 4 6 8


−10
−15
−8 −6 −4 −2 0 2 4 6 8


−10
−15
−8 −6 −4 −2 0 2 4 6 8


−10
−15
−8 −6 −4 −2 0 2 4 6 8



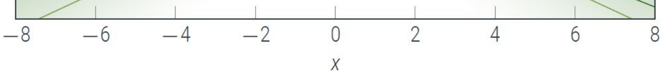
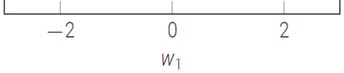


Image: Philipp Hennig



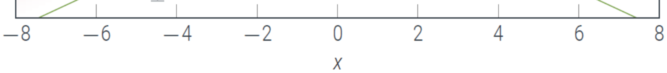
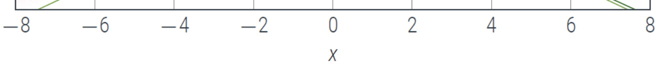


Image: Philipp Hennig



Universidad Zaragoza 1542

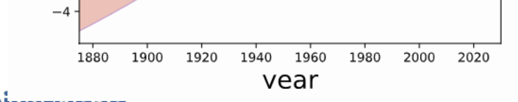
Image: Philipp Hennig

$$(f_*) \quad (\quad [\quad k(x_*, X) \quad\quad k(x_*, x_*)])$$

Left: a Bayesian neural network with two hidden layers, transforming a 3-dimensional input into a two-dimensional output. **Right**: output probability density function induced by the random weights of the network. **Video**: as the width of the network increases, the output distribution simplifies, ultimately converging to a multivariate normal in the infinite width limit. https://en.wikipedia.org/wiki/Large_width_limits_of_neural_networks

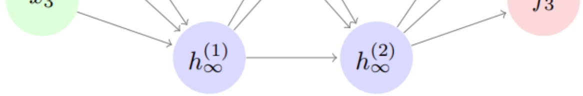
https://arxiv.org/pdf/1402.5836.pdf

- 1904 Olympic marathon
  - Unusually high temperatures and humidity, coupled with poor organization and planning, meant that the marathon in 1904 had the slowest winning time and most bizarre outcome in Olympic history. https://www.statista.com/statistics/1099351/olympics-marathon-gold-medal-times-since-1896/



Universidad Zaragoza

Example by Neil Lawrence: http://inverseprobability.com/talks/notes/deep-gaussian-processes.html


$h_\infty^{(1)}$
$h_\infty^{(2)}$

- Heteroscedastic!

F. K. Gustafsson, M. Danelljan, T. B. Schon, Evaluating scalable bayesian deep learning methods for robust computer vision, CVPRW. 2020.

- Homoscedastic noise $\sigma_W^2(\boldsymbol{x}) = \sigma_{MAP}^2$

Universidad Zaragoza

F. K. Gustafsson, M. Danelljan, T. B. Schon, Evaluating scalable bayesian deep learning methods for robust computer vision, CVPRW. 2020.

Image: Brendan Hasz

$$KL(q||p) = \mathbb{E}_q[\log q(x)] - \mathbb{E}_q[\log p(x)] = -\mathbb{E}_q[\log p(x)] - \mathcal{H}(q(x))$$

Universidad Zaragoza 1542

$$q(w_i) \sim N(\mu_i, \sigma_i)$$

- We want to maximize ELBO → Loss = negative ELBO more approximations. [Osawa et al. 2019]

Universidad Zaragoza 1542

$$p(\mathbf{w}|\mathbf{X}, \mathbf{y}) \approx q(\mathbf{w}) = \frac{\exp\left(\frac{1}{2}(\mathbf{w} - \mathbf{w}^*)^T H(\mathbf{w}^*)(\mathbf{w} - \mathbf{w}^*)\right)}{\int \exp\left(\frac{1}{2}(\mathbf{w} - \mathbf{w}^*)^T H(\mathbf{w}^*)(\mathbf{w} - \mathbf{w}^*)\right) d\mathbf{w}}$$

Image: Philipp Hennig


θ_{MAP}

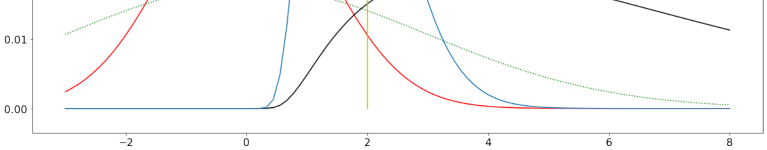
(a) Full **(b)** LRank **(c)** KFAC **(d)** Diag.

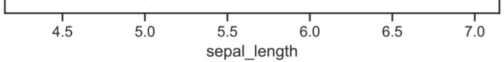
Universidad Zaragoza 1542

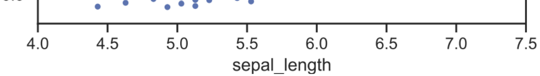
Daxberger E, Kristiadi A, Immer A, Eschenhagen R, Bauer M, Hennig P. Laplace redux-effortless bayesian deep learning. Advances in Neural Information Processing Systems. 2021 Dec 6;34:20089-103.
https://github.com/AlexImmer/Laplace

- Then the posterior is: $y_{\mathrm{lin}} \sim \mathcal{N}\left(f_{\mathbf{w}^*}(\mathbf{x}), J(\mathbf{x}) H(\mathbf{w}^*)^{-1} J(\mathbf{x})^T\right)$

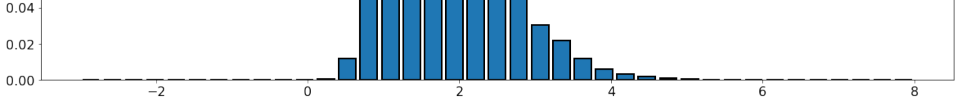
Khan 2019: https://proceedings.neurips.cc/paper_files/paper/2019/file/b3bbccd6c008e727785cb81b1aa08ac5-Paper.pdf


0.01
0.00
−2
0
2
4
6
8

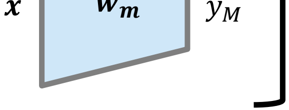




Kevin P. Murphy, Probabilistic Machine Learning: An introduction, section 4.6.7.2, MIT Press, 2021, https://probml.github.io/pml-book/


0.04
0.02
0.00
−2
0
2
4
6
8


0
−10
0
10
20
0
−10
0
10
20



- We can only use a small number of samples! $\mathcal{O}(M)$

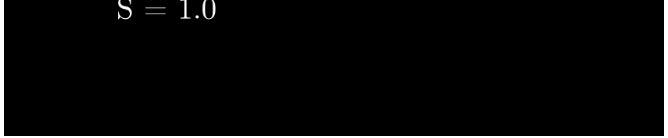
epistemic aleatoric

Universidad Zaragoza 1542

Universidad Zaragoza 1542

Y. Gal and Z. Ghahramani, Dropout as a Bayesian approximation: Representing model uncertainty in deep learning, ICML, 2016

- Masksensembles



Universidad Zaragoza 1542

different networks)

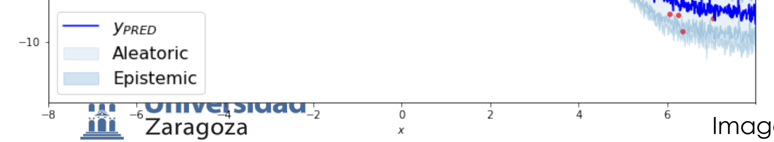
- Better than local approximations

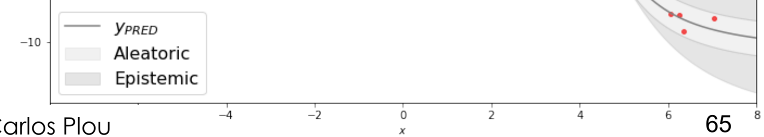
B. Lakshminarayanan, A. Pritzel, and C. Blundell, Simple and scalable predictive uncertainty estimation using deep ensembles, NeurIPS, 2017

$outputs$ = (● , ... , ●)

Universidad Zaragoza 1542

Image: Carlos Plou





Universidad Zaragoza 1542

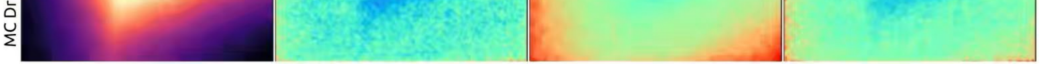
Images: Carlos Plou

- See K.P. Murphy Book2 Chapter 17 https://probml.github.io/pml-book/book2.html

Universidad Zaragoza

(a) Input Image (b) Ground Truth (c) Semantic Segmentation (d) Aleatoric Uncertainty (e) Epistemic Uncertainty

Alex Kendall and Yarin Gal. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?, NeurIPS, 2017



Universidad Zaragoza 1542

Javier Rodríguez-Puigvert, Rubén Martínez-Cantín, Javier Civera, Bayesian Deep Networks for Supervised Single-View Depth Learning, IEEE-RAL, 2021

Javier Rodríguez-Puigvert, Rubén Martínez-Cantín, Javier Civera, Bayesian Deep Networks for Supervised Single-View Depth Learning, IEEE-RAL, 2021

Figure 5: NYUv2 Depth results. From left: input image, ground truth, depth regression, aleatoric uncertainty, and epistemic uncertainty.

Alex Kendall and Yarin Gal. What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?, NeurIPS, 2017

M Deep Ensembles

Student

Universidad Zaragoza 1542

M Deep Ensembles

Student

Background	Contain	Cut	Display	Engine	Grasp	Hit	Pound	Support	W-Grasp

Universidad Zaragoza
1542

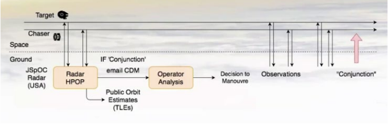
Mur-Labadia, Lorenzo, Martinez-Cantin, Ruben and Guerrero, Jose J. (2023) *Bayesian deep learning for affordance segmentation in images*. In Proc. of the IEEE International Conference on Robotics and Automation.

Mur-Labadia, Lorenzo, Martinez-Cantin, Ruben and Guerrero, Jose J. (2023) *Bayesian deep learning for affordance segmentation in images.* In Proc. of the IEEE International Conference on Robotics and Automation.


s
a
s'

• • •

Universidad Zaragoza 1542

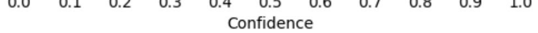


its probability as the predicted confidence (p).

F. K. Gustafsson, M. Danelljan, T. B. Schon, Evaluating scalable Bayesian deep learning methods for robust computer vision, CVPRW. 2020.
https://openaccess.thecvf.com/content_CVPRW_2020/papers/w20/Gustafsson_Evaluating_Scalable_Bayesian_Deep_Learning_Methods_for_Robust_Computer_Vision_CVPRW_2020_paper.pdf

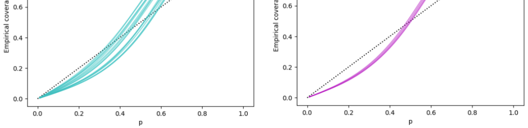
- n: number of samples

C. Guo, G. Pleiss, Y. Sun, K. Q. Weinberger, On calibration of modern neural networks, ICML 2017



F. K. Gustafsson, M. Danelljan, T. B. Schon, Evaluating scalable Bayesian deep learning methods for robust computer vision, CVPRW. 2020.

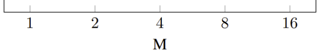
- $|\tilde{p} - p|$ for multiple p
- In a well calibrated model $|\tilde{p} - p| = 0$.



Universidad Zaragoza 1542

P(rain)=0,7, rain=1.00, BS=(0,7-1)^2=0,09

G. W. Brier. Verification of forecasts expressed in terms of probability. Monthly weather review, 1950.



F. K. Gustafsson, M. Danelljan, T. B. Schon, Evaluating scalable Bayesian deep learning methods for robust computer vision, CVPRW. 2020.

- Csaba Szepesvari, Algorithms of Reinforcement Learning
  - https://sites.ualberta.ca/~szepesva/rlbook.html