Supplementary Material - Penalised Regression Adaptations of the Longstaff-Schwartz Algorithm for Pricing American Options

David Suda[1], Monique Borg Inguanez[2], and Lara Cilia[3]

[1] Department of Statistics and Operations Research, University of Malta, Msida, Malta (E-mail: david.suda@um.edu.mt)
[2] Department of Statistics and Operations Research, University of Malta, Msida, Malta (E-mail: monique.inguanez@um.edu.mt)
[3] Central Bank of Malta, Valletta, Malta (E-mail: cilial@centralbankmalta.org)

1 Important Theorems

The following are convergence theorems and related assumptions from Stentoft[2].

Assumption 1: The simulated paths: $\{S(\omega_n, t_k)\}$, $1 \leq n \leq N$, $0 \leq k \leq K$ are independent and for an adapted payoff process $\{Z(\omega, t_k)\}_{k=0}^{K}$: $P[Z(\omega), t_k)] = F(\omega, t_k) = 0$.

Assumption 2: The support of the observations $S(\cdot, t_1)$, denoted $\mathcal{S}$, is a Cartesian product of compact connected intervals on which $S(\cdot, t_1)$ has a probability density function, denoted $f(s)$, that is bounded below by some $\epsilon > 0$.

Assumption 3: The conditional expectation function is continuously differentiable of order d on the support of the observations s, where d is the order of differentiability of $F(\omega)$.

Theorem 1. *Under Assumptions 1, 2 and 3, if $M = M(N)$ is increasing in N such that $M \to \infty$ and $M^3/N \to 0$, then the power series estimator $\hat{F}_M^N(\omega)$ in the cross-sectional regression as defined in $\hat{F}_M(\omega; t_{k-1}) = \sum_{j=0}^{M} \hat{a}_j \phi_j(S_{k-1})$ is mean-square convergent*

$$\int [F(\omega) - \hat{F}_M^N(\omega)]^2 dF_0(s) = O_p(M/N + M^{-2d/q}) \tag{1}$$

where d is the order of differentiability of $F(\omega)$ and q is the dimension of s.

Theorem 2. *Under Assumptions 1, 2 and 3, if $M = M(N)$ is increasing in N such $M \to \infty$ and $M^3/N \to 0$, then the estimated continuation value for N paths and M basis functions $\hat{F}_M^N(\omega; t_k)$ converges to $F(\omega; t_k)$ in probability for $k = 1, ..., K$.*

The following are convergence theorems and related assumptions from Gerhold[1].

Assumption 4 (Exponential Lèvy dynamics): The risk neutral dynamics of the underlying asset are:

$$S_t = S_0 \exp(M_t)$$

where M_t is a Lèvy process with $M_0 = 0$. The support of M_t is the whole real line for $t > 0$ and M_t has a continuous density function.

Assumption 5 (Value smoothness): Let the payoff function X be of, at most, linear growth and such that $X(S_T)$ is integrable for each $T > 0$. Then $E[X(S_t)|S_0 = s]$ is a C^1-smooth function of s.

Assumption 6: The basis functions are shifted Legendre polynomials, the continuation values $C(\omega)$ are in the L^2-span of the regressors, the simulated paths are independent and the probability that the exercise payoff exactly equals the continuation value is zero.

Assumption 7 (Integrability) For each t, there are $p > 1$ and $p' > 0$ such that S_t^p and $S_t^{-p'}$ are integrable.

Theorem 3. *Suppose that Assumptions 4 and 7 hold and that the log-price M_t is a Meixner process. Then Assumption 5 holds.*

Theorem 4. *Fix arbitrary finite truncation intervals $I_1, ..., I_m$ contained in $]0, \infty[$ and assume that Assumptions 4–6 hold. Let N and M tend to infinity such that $M^3/N \to 0$. Then the option prices computed by the truncated L-S algorithm converge to the approximate price $V_0^{tr} = \max\{h_0(S_0), C_0^{tr}(S_0)\}$*

Theorem 5. *Assume that Assumptions 4 and 7 hold and that the payoff functions grow at most linearly, i.e. $|h_n(x)| \leq c(1 + x)$ for $x \geq 0$, $1 \leq n \leq m$, for some $c > 0$. Moreover, assume that the truncation intervals satisfy:*

$$I_n = [b_n^{-1}, b_n] \quad 1 \leq n \leq m$$

where

$$b_n = b_{n+1}^v \quad 1 \leq n < m - 1$$

with:

$$v = \min\left\{\frac{p'}{p' + q}, \frac{p}{p + q}\right\}$$

and $\frac{1}{p} + \frac{1}{q} = 1$. Then V_0^{tr} converges to the exact option price V_0 as b_m tends to infinity.

2 Tables with Estimates from 20 Runs

Table 1: The average estimated price and its standard error for the GBM process with 3 basis functions for various sample sizes of paths, where the benchmark values is 1.3449.

		Number of Paths								
		2,000	**3,000**	**4,000**	**5,000**	**6,000**	**7,000**	**8,000**	**9,000**	**10,000**
Estimation Method	**OLS**	1.3975	1.3750	1.3618	1.3590	1.3588	1.3519	1.3539	1.3543	1.3536
		(0.0098)	(0.0078)	(0.0068)	(0.0051)	(0.0050)	(0.0050)	(0.0047)	(0.0042)	(0.0042)
	LASSO	1.3845	1.3628	1.3544	1.3514	1.3527	1.3475	1.3485	1.3498	1.3475
		(0.0098)	(0.0074)	(0.0066)	(0.0055)	(0.0046)	(0.0046)	(0.0042)	(0.0039)	(0.0042)
	Ridge	1.3875	1.3665	1.3592	1.3554	1.3552	1.3499	1.3514	1.3502	1.3502
		(0.0102)	(0.0072)	(0.0068)	(0.0050)	(0.0046)	(0.0045)	(0.0046)	(0.0040)	(0.0046)
	EN	1.3855	1.3636	1.3550	1.3536	1.3535	1.3482	1.3486	1.3505	1.3486
		(0.0100)	(0.0073)	(0.0071)	(0.0055)	(0.0043)	(0.0047)	(0.0040)	(0.0041)	(0.0043)

Table 2: The average estimated price and its standard error for the Heston SV process with 3 basis functions for various sample sizes of paths where the benchmark values is 0.4198.

		Number of Paths								
		2,000	**3,000**	**4,000**	**5,000**	**6,000**	**7,000**	**8,000**	**9,000**	**10,000**
Estimation Method	**OLS**	0.4557	0.4560	0.4542	0.4512	0.4516	0.4488	0.4490	0.4484	0.4495
		(0.0041)	(0.0036)	(0.0031)	(0.0029)	(0.0023)	(0.0021)	(0.0021)	(0.0019)	(0.0020)
	LASSO	0.4505	0.4514	0.4506	0.4476	0.4483	0.4464	0.4476	0.4464	0.4472
		(0.0046)	(0.0035)	(0.0033)	(0.0029)	(0.0024)	(0.0022)	(0.0022)	(0.0019)	(0.0019)
	Ridge	0.4536	0.4533	0.4523	0.4485	0.4500	0.4479	0.4483	0.4474	0.4485
		(0.0044)	(0.0035)	(0.0031)	(0.0027)	(0.0024)	(0.0023)	(0.0021)	(0.0017)	(0.0018)
	EN	0.4493	0.4510	0.4505	0.4477	0.4484	0.4464	0.4477	0.4467	0.4478
		(0.0046)	(0.0038)	(0.0033)	(0.0028)	(0.0024)	(0.0023)	(0.0021)	(0.0020)	(0.0019)

Table 3: The average estimated price and its standard error for the Meixner process with 3 basis functions for various sample sizes of paths, where the benchmark values is 1.9997.

		Number of Paths								
		2,000	**3,000**	**4,000**	**5,000**	**6,000**	**7,000**	**8,000**	**9,000**	**10,000**
Estimation Method	**OLS**	2.1984	2.1931	2.1891	2.1883	2.1855	2.1834	2.1797	2.1818	2.1814
		(0.0079)	(0.0070)	(0.0045)	(0.0043)	(0.005)	(0.0048)	(0.0049)	(0.0045)	(0.0048)
	LASSO	2.1754	2.1720	2.1683	2.1654	2.1642	2.1627	2.1597	2.1625	2.1632
		(0.0082)	(0.0070)	(0.0047)	(0.0043)	(0.0048)	(0.0047)	(0.0046)	(0.004)	(0.0043)
	Ridge	2.1928	2.1894	2.1871	2.1871	2.1840	2.1824	2.1800	2.1812	2.1805
		(0.008)	(0.0072)	(0.0044)	(0.0047)	(0.0052)	(0.0049)	(0.0046)	(0.0046)	(0.0048)
	EN	2.1743	2.1713	2.1681	2.1652	2.1639	2.1621	2.1600	2.1627	2.1636
		(0.0086)	(0.0070)	(0.0049)	(0.0042)	(0.0049)	(0.0044)	(0.0044)	(0.0041)	(0.0043)

Table 4: The absolute percentage difference from the benchmark price and its standard error for a GBM process with 3 basis functions for various sample sizes of paths.

		Number of Paths								
		2,000	**3,000**	**4,000**	**5,000**	**6,000**	**7,000**	**8,000**	**9,000**	**10,000**
Estimation Method	**OLS**	3.6737	3.0792	2.8101	2.0538	2.0404	2.1147	1.9761	1.6067	1.6222
		(0.5995)	(0.3921)	(0.3218)	(0.2455)	(0.2560)	(0.2152)	(0.1923)	(0.1826)	(0.1976)
	LASSO	3.9740	3.1297	2.6794	2.2889	1.9488	2.0045	1.8440	1.4455	1.5131
		(0.5353)	(0.3807)	(0.3424)	(0.2743)	(0.2142)	(0.2100)	(0.1570)	(0.1671)	(0.1731)
	Ridge	3.3222	2.1321	2.0339	1.7059	1.8411	1.7524	1.5285	1.5575	1.5279
		(0.5390)	(0.3375)	(0.3392)	(0.2518)	(0.2330)	(0.2076)	(0.1785)	(0.1673)	(0.2020)
	EN	3.7911	3.0855	2.7535	2.2650	1.9852	2.1220	1.7778	1.4375	1.4896
		(0.5422)	(0.3543)	(0.3917)	(0.2687)	(0.2073)	(0.2199)	(0.1555)	(0.1750)	(0.1718)

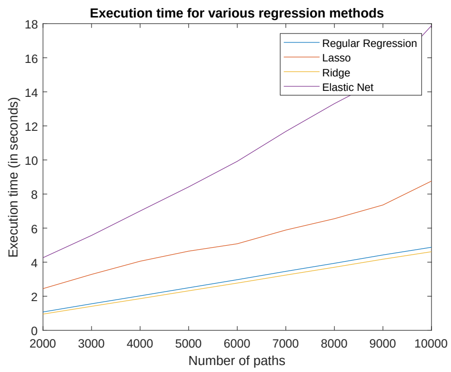
Table 5: The absolute percentage difference from the benchmark price and its standard error for a Heston SV process with 3 basis functions for various sample sizes of paths.

		Number of Paths								
		2,000	**3,000**	**4,000**	**5,000**	**6,000**	**7,000**	**8,000**	**9,000**	**10,000**
Estimation Method	**OLS**	7.727	7.8196	7.4777	6.8735	6.9914	6.4082	6.4540	6.3334	6.5657
		(0.8075)	(0.7268)	(0.6237)	(0.5911)	(0.4849)	(0.4451)	(0.4440)	(0.4035)	(0.4148)
	LASSO	6.6277	6.8900	6.7257	6.1342	6.2901	5.9053	6.1694	5.9172	6.0962
		(0.9129)	(0.7189)	(0.6724)	(0.6011)	(0.5009)	(0.4492)	(0.4643)	(0.3994)	(0.4019)
	Ridge	7.2951	7.2877	7.1003	6.3316	6.6542	6.2126	6.3041	6.1291	6.3525
		(0.8573)	(0.7048)	(0.6264)	(0.5658)	(0.4983)	(0.4705)	(0.4301)	(0.3662)	(0.3808)
	EN	6.3784	6.7962	6.7176	6.1478	6.3089	5.9122	6.1829	5.9773	6.2075
		(0.9097)	(0.774)	(0.669)	(0.5885)	(0.5068)	(0.4674)	(0.4410)	(0.4172)	(0.4012)

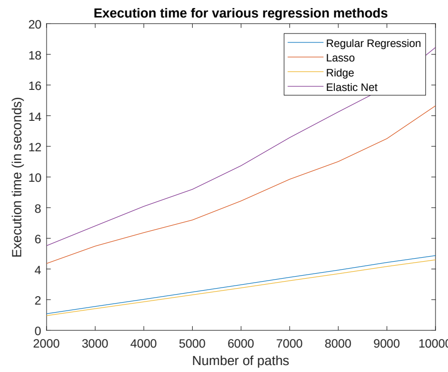
Table 6: The absolute percentage difference from the benchmark price and its standard error for a Meixner process with 3 basis functions for various sample sizes of paths.

		Number of Paths								
		2,000	**3,000**	**4,000**	**5,000**	**6,000**	**7,000**	**8,000**	**9,000**	**10,000**
Estimation Method	**OLS**	9.0178	8.8016	8.6468	8.6112	8.4936	8.4054	8.2482	8.338	8.3226
		(0.3288)	(0.2911)	(0.1883)	(0.1807)	(0.2096)	(0.1991)	(0.2055)	(0.1872)	(0.1996)
	LASSO	8.0506	7.9139	7.7684	7.6468	7.5948	7.5279	7.4032	7.5229	7.552
		(0.3507)	(0.2959)	(0.2011)	(0.1847)	(0.2046)	(0.2016)	(0.1954)	(0.1716)	(0.1829)
	Ridge	8.7842	8.6458	8.5622	8.5602	8.4292	8.3620	8.2627	8.3159	8.2827
		(0.3325)	(0.3009)	(0.1850)	(0.1975)	(0.2167)	(0.2034)	(0.1932)	(0.1944)	(0.2033)
	EN	8.0026	7.8862	7.7575	7.6396	7.5793	7.5034	7.4129	7.5293	7.5681
		(0.3651)	(0.2967)	(0.2095)	(0.1814)	(0.2086)	(0.1906)	(0.1880)	(0.1775)	(0.1855)

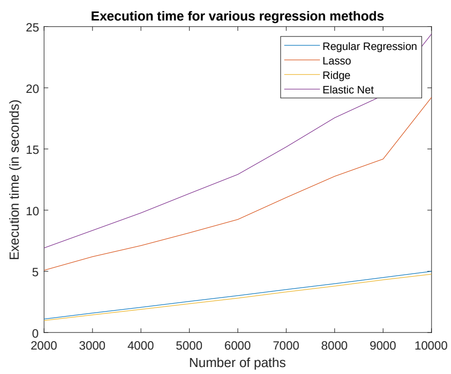
3 Execution Times for Various Regression Methods and Number of Basis Functions



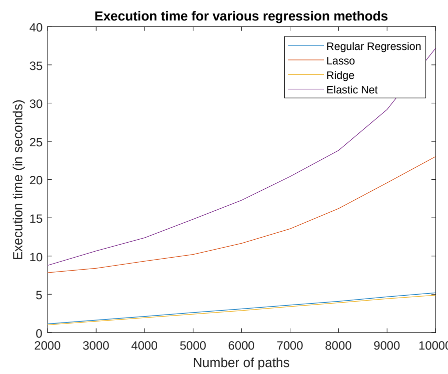
GBM - 3 basis functions



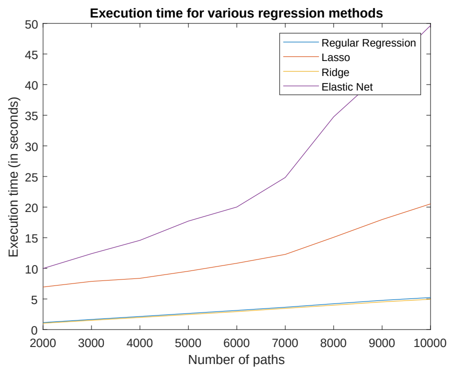
GBM - 4 basis functions



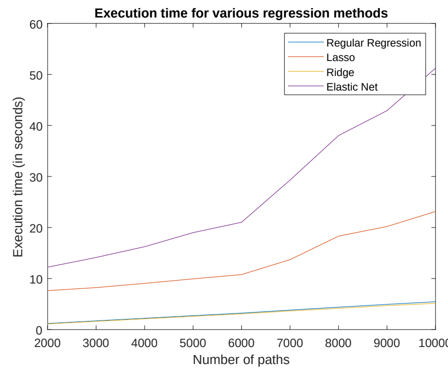
GBM - 5 basis functions



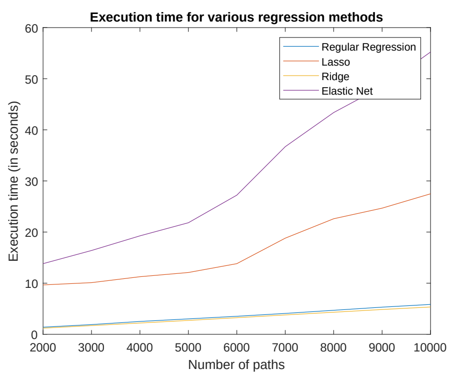
GBM - 6 basis functions



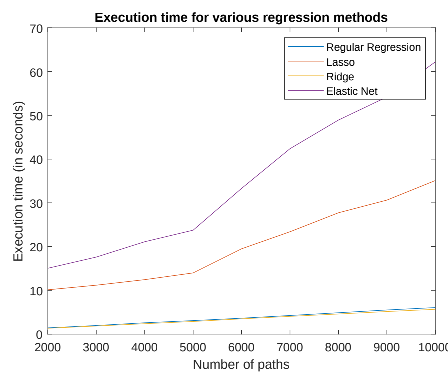
GBM - 7 basis functions



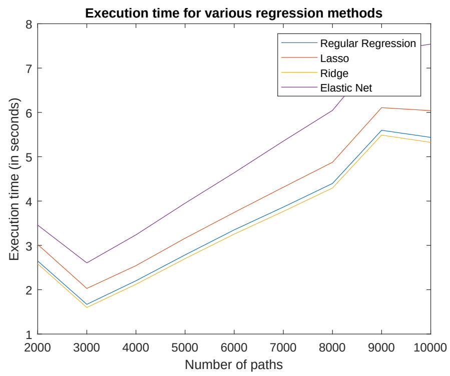
GBM - 8 basis functions



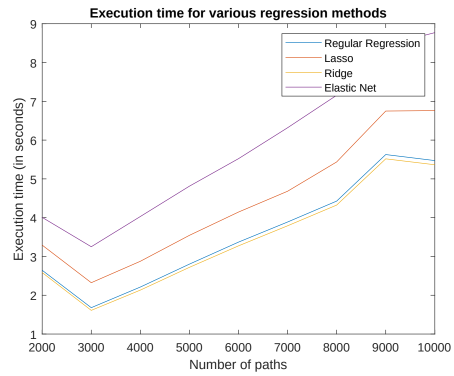
GBM - 9 basis functions



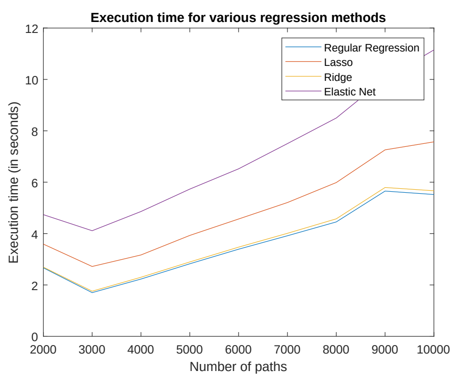
GBM - 10 basis functions



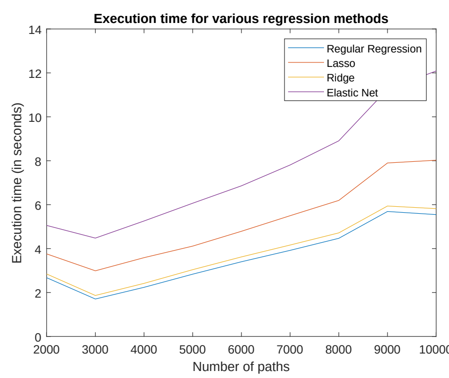
Heston - 3 basis functions



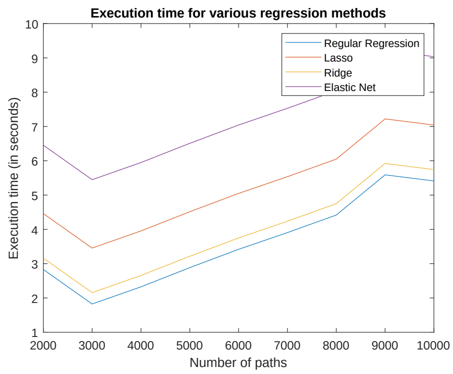
Heston - 4 basis functions



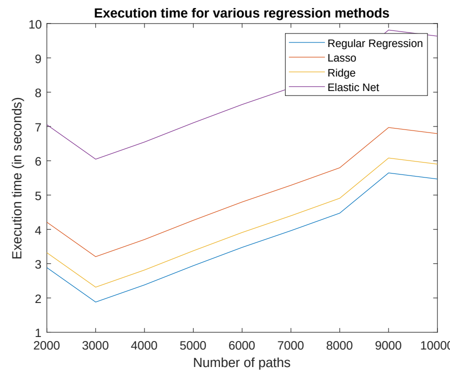
Heston - 5 basis functions



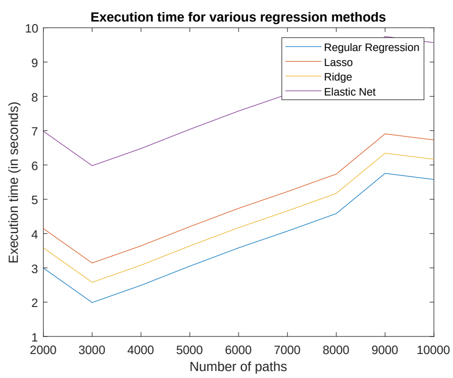
Heston - 6 basis functions



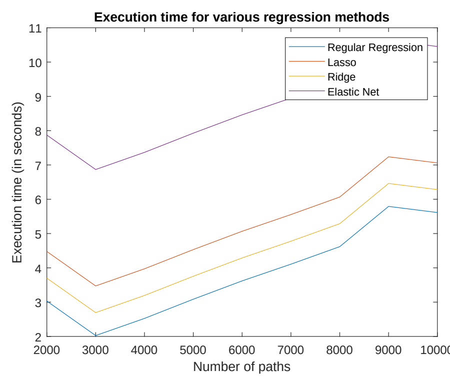
Heston - 7 basis functions



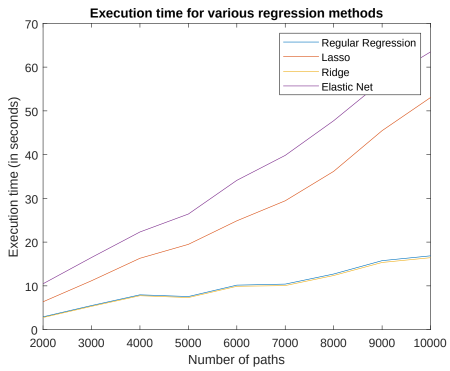
Heston - 8 basis functions



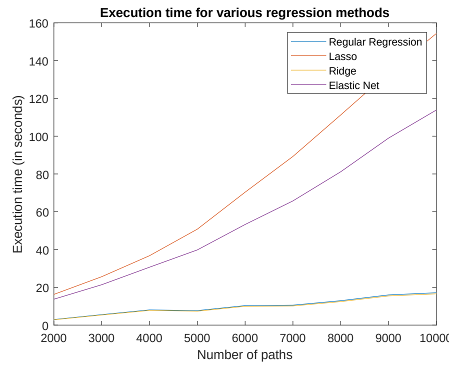
Heston - 9 basis functions



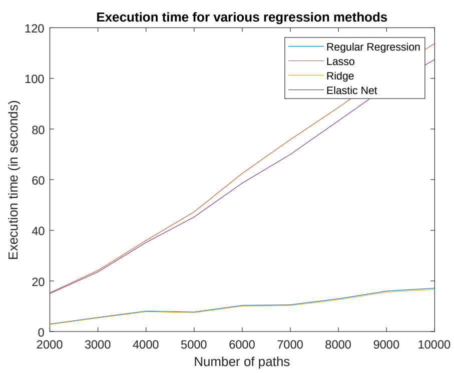
Heston - 10 basis functions



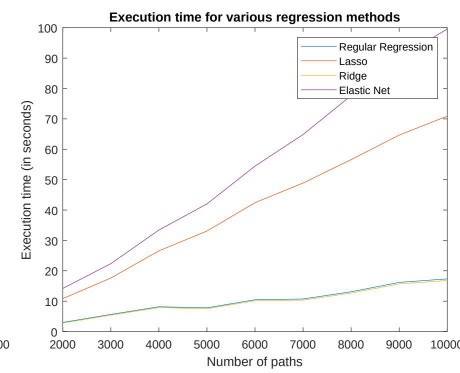
Meixner - 3 basis functions



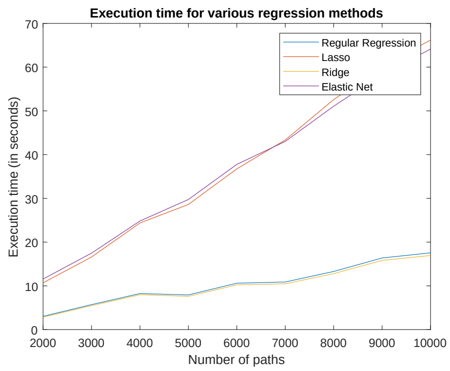
Meixner - 4 basis functions



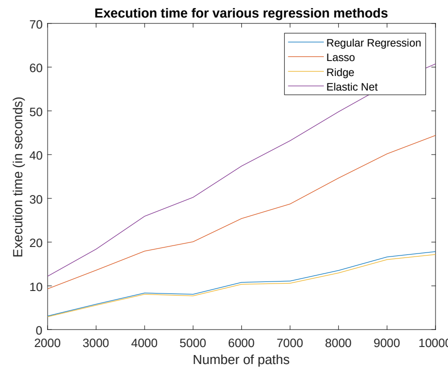
Meixner - 5 basis functions



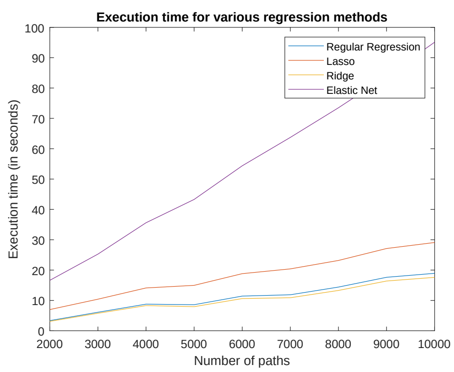
Meixner - 6 basis functions



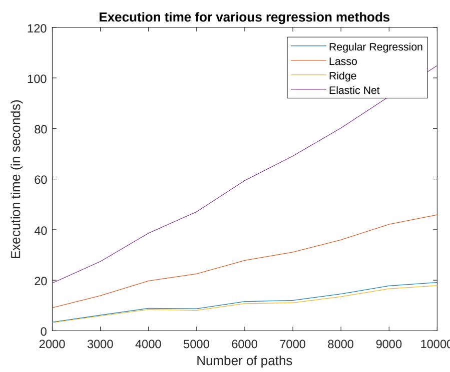
Meixner - 7 basis functions



Meixner - 8 basis functions

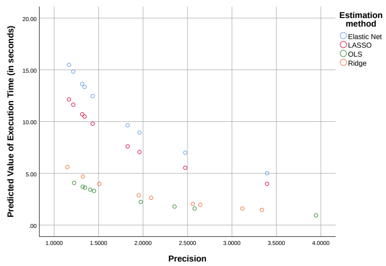


Meixner - 9 basis functions

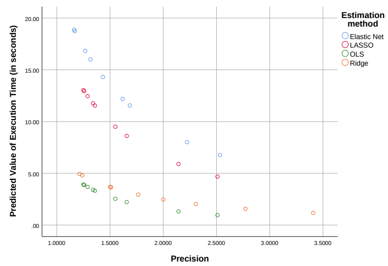


Meixner - 10 basis functions

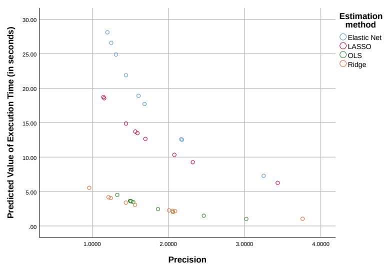
4 Execution Time vs Precision for Various Regression Methods and Number of Basis Functions



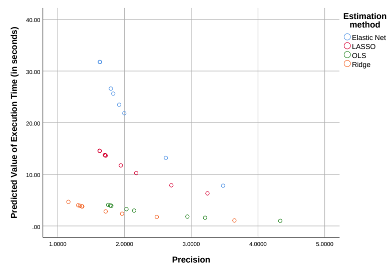
GBM - 4 basis functions



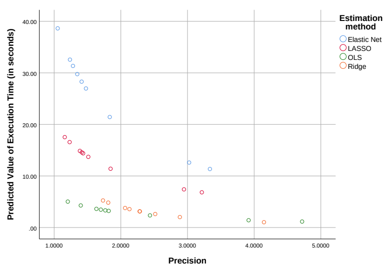
GBM - 5 basis functions



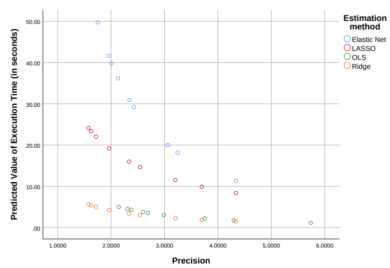
GBM - 6 basis functions



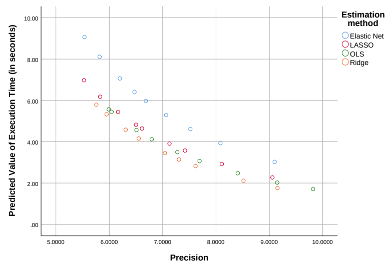
GBM - 7 basis functions



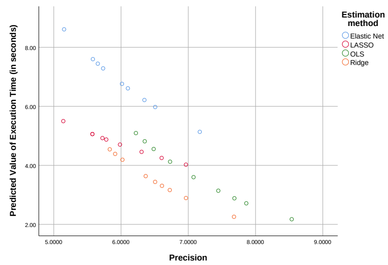
GBM - 8 basis functions



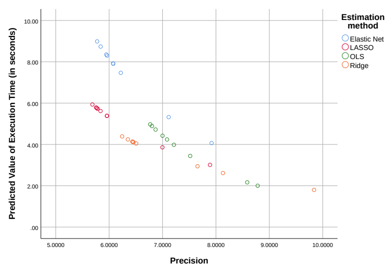
GBM - 9 basis functions



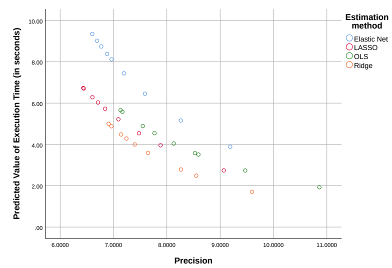
Heston - 4 basis functions



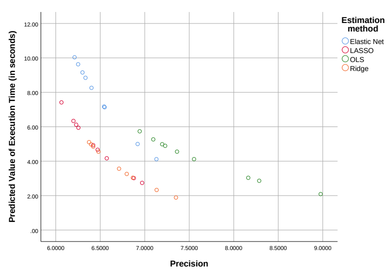
Heston - 5 basis functions



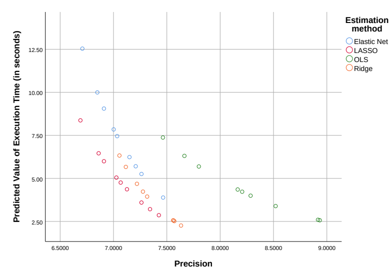
Heston - 6 basis functions



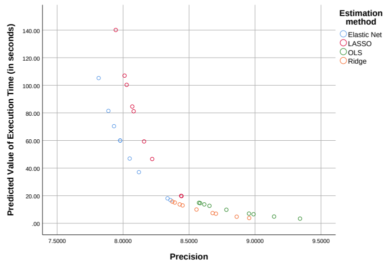
Heston - 7 basis functions



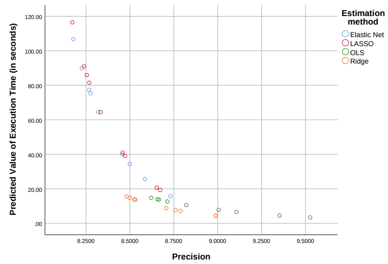
Heston - 8 basis functions



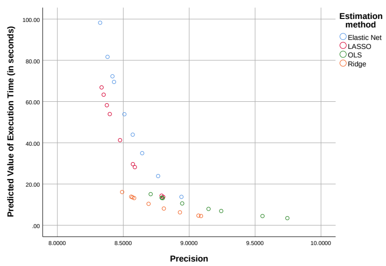
Heston - 9 basis functions



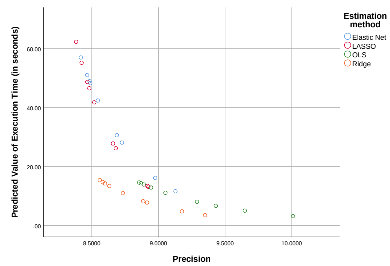
Meixner - 4 basis functions



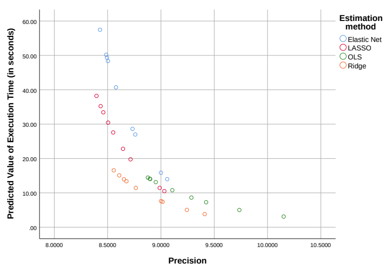
Meixner - 5 basis functions



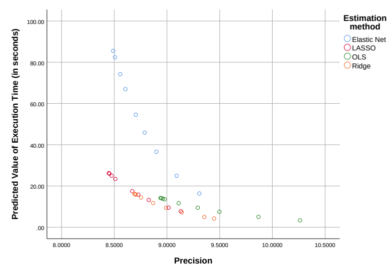
Meixner - 6 basis functions



Meixner - 7 basis functions



Meixner - 8 basis functions



Meixner - 9 basis functions

References

1. S. Gerhold. The Longstaff-Schwartz algorithm for Lévy models: results on fast and slow convergence. *The Annals of Applied Probability*, 22, 2, 589–608, 2011.
2. L. Stentoft. Convergence of the least squares Monte Carlo approach to American option valuation. *Management Science*, 50, 9, 1193–1203, 2004.