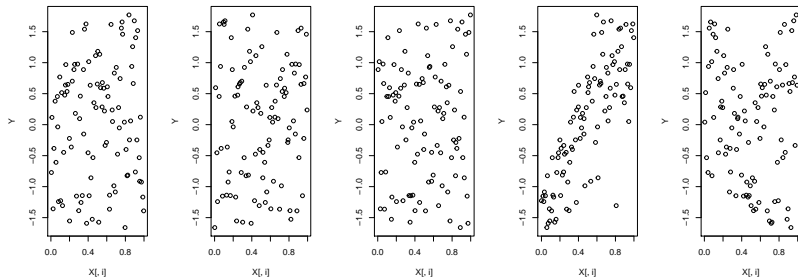


# TP du cours optim avec métamodélisation - corrigé

Nicolas Durrande, Victor Picheny,

# LHS initial



Pas grand chose à voir... à part la 4e variable.

# Modèle initial

```
km(formula = y ~ ., design = data.frame(x = X), response = data.frame(y = Y),  
    covtype = "matern5_2", nugget = 1e-08)
```

Trend coeff.:

	Estimate
(Intercept)	-1.5942
x.1	0.0819
x.2	-0.0102
x.3	0.8546
x.4	2.5819
x.5	-0.1406

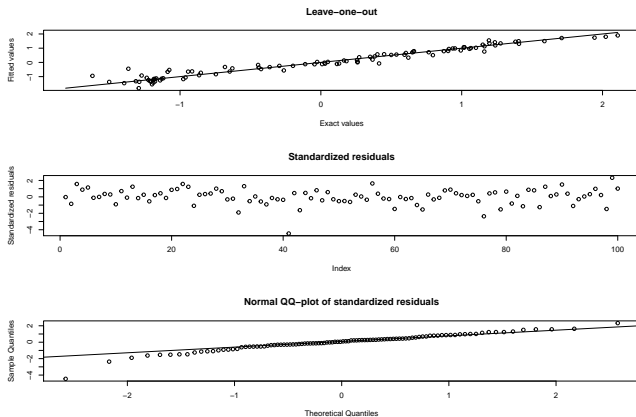
Covar. type : matern5\_2

Covar. coeff.:

	Estimate
theta(x.1)	1.9621
theta(x.2)	1.9661
theta(x.3)	1.6171
theta(x.4)	0.4056
theta(x.5)	0.0561

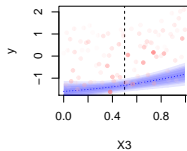
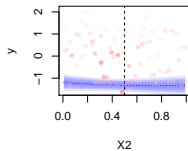
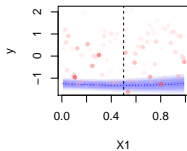
Variance estimate: 0.2128299

On observe bien : un paramètre de tendance et de portée par  $x.i$ .  
Sachant que chaque  $x.i$  varie entre 0 et 1, une portée égale à 2 est très élevée (le krigeage est presque constant dans cette direction,  $x.i$  est peu influent), et de 0.05 très petite (beaucoup de variation dans cette direction).

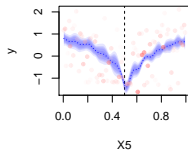
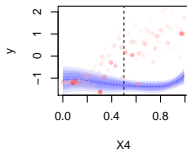


Validation : erreur par validation croisée plutôt bonne, quelques points imprécis pour les valeurs faibles de  $Y$ . QQ-plot : OK sauf pour les quantiles faibles.

= 0.500, X3 = 0.500, X4 = 0.500, X5= 0.500, X3 = 0.500, X4 = 0.500, X5= 0.500, X2 = 0.500, X4 = 0.500, X5



= 0.500, X2 = 0.500, X3 = 0.500, X5= 0.500, X2 = 0.500, X3 = 0.500, X4



On retrouve bien : un modèle très “plat” pour  $x.1$ ,  $x.2$ ,  $x.3$ , très variable pour  $x.5$ .

# Tendance, noyaux

```
m1 <- km(y~1, design=data.frame(x=X), response=data.frame(y=Y),
        covtype="gauss", nugget=1e-8)

m2 <- km(y ~ .^2 + xs^2 + ys^2 + zs^2 + a^2 + p^2, design=data.frame(X),
        response=data.frame(y=Y),
        covtype="gauss", nugget=1e-8, lower=rep(0.05, 5))

m3 <- km(y~., design=data.frame(x=X), response=data.frame(y=Y),
        covtype="matern5_2", nugget=1e-8)

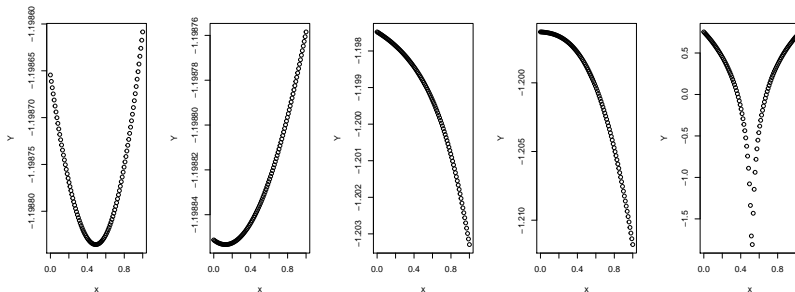
c(logLik(m1), logLik(m2), logLik(m3))
-117.294340    2.902064    -2.791663

c(Q2m1, Q2m2, Q2m3)
0.9576836 0.9716956 0.9662655
```

# Tendance, noyaux

- ▶ pas de choix de noyau ou tendance évident...
- ▶ la solution optimale pour la vraisemblance ou le Q2 dépend de votre plan initial !
- ▶ optimisation de la vraisemblance locale ! sauf si `optim.method="gen"`

En fait : fonction très piégeuse, difficile de trouver un noyau qui convient dans toutes les directions



# EGO

```
res <- EGO.nsteps(model=m, fun=compute_wls, nsteps=50,  
                  lower=rep(0, 5), upper=rep(1,5))
```

```
names(res)
```

"par"

"value"

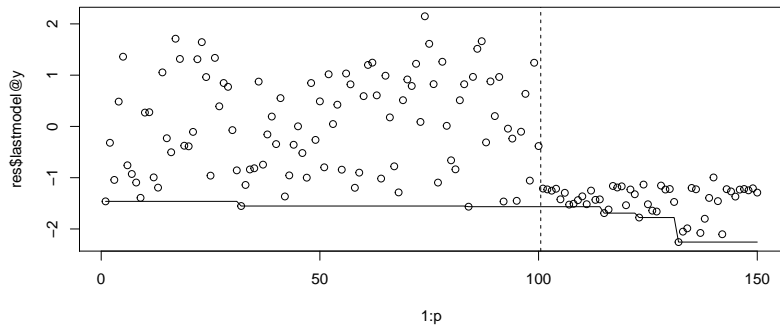
"npoints"

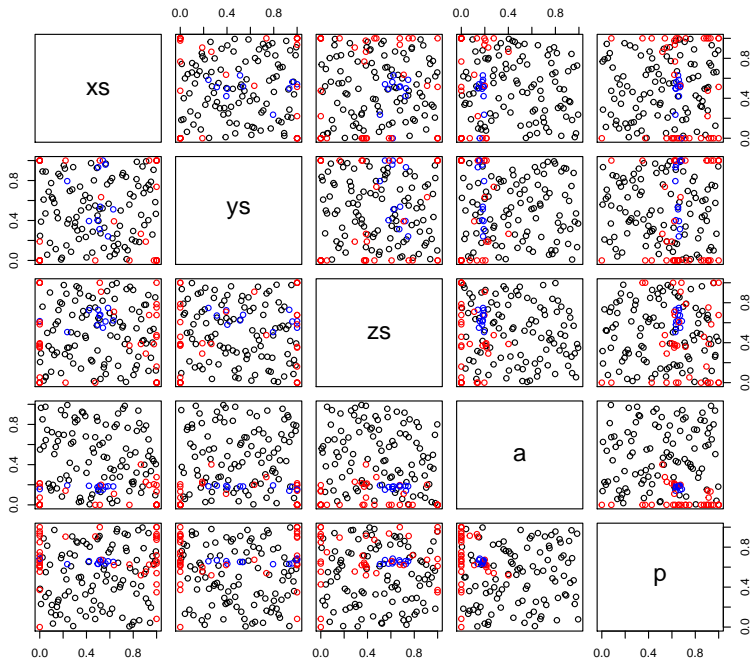
"nsteps"

"lastmodel"

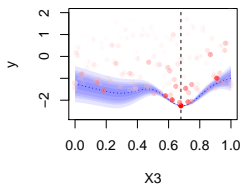
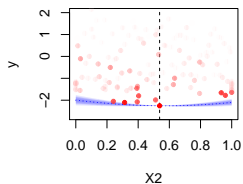
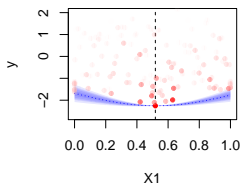


# EGO





**$X_2 = 0.5, X_3 = 0.7, X_4 = 0.183, X_5 =$   $X_1 = 0.5, X_3 = 0.7, X_4 = 0.183, X_5 =$   $X_1 = 0.5, X_2 = 0.5, X_4 = 0.183, X_5 =$**



**$X_1 = 0.5, X_2 = 0.5, X_3 = 0.7, X_5 =$   $X_1 = 0.5, X_2 = 0.5, X_3 = 0.7, X_4 = 0.$**

