Université de Technologie de Compiègne
Master Structures et Systèmes Mécaniques Complexes (SMC)

Modélisation stochastique en mécanique

Mokhtar Z. Alaya et Salim Bouzebda

CORRIGÉ DU TD 2

Apprentissage statistique pour la mécanique

Banque complète : six exercices corrigés. Parcours encadré : 1 h 30.

Exercice 1 Identifier les coefficients mécaniques

On cherche un modèle affine du moment fléchissant :

$$M_1 \simeq \theta_0 + \theta_1 Y_1 + \theta_2 Y_2.$$

Les charges Y_1, Y_2 sont en N et le moment M_1 en N mm. Les quatre mesures sont sans bruit :

Essai	Y_1 (N)	Y_2 (N)	M_1 (N mm)
1	100	0	21000
2	0	100	12000
3	100	100	33000
4	200	100	54000

Question 1 — Construire la matrice et identifier les coefficients

- **Écrire une équation par essai.**

Avec $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2)^\mathsf{T}$, chaque ligne de D contient la constante 1 et les deux charges du même essai. Ainsi,

$$D = \begin{pmatrix} 1 & 100 & 0 \\ 1 & 0 & 100 \\ 1 & 100 & 100 \\ 1 & 200 & 100 \end{pmatrix}, \qquad \boldsymbol{y} = \begin{pmatrix} 21000 \\ 12000 \\ 33000 \\ 54000 \end{pmatrix}, \qquad D\boldsymbol{\theta} = \boldsymbol{y}.$$

La constante θ_0 est en N mm ; les coefficients θ_1 et θ_2 sont en mm, puisque leur produit par une charge doit être un moment.

- **Résoudre par différences entre essais.**

Les trois premières équations sont

$$\begin{cases} \theta_0 + 100\theta_1 = 21000, & (1) \\ \theta_0 + 100\theta_2 = 12000, & (2) \\ \theta_0 + 100\theta_1 + 100\theta_2 = 33000. & (3) \end{cases}$$

Pour vérifier la compatibilité, la combinaison $(1) + (2) - (3)$ donne

$$(2\theta_0 + 100\theta_1 + 100\theta_2) - (\theta_0 + 100\theta_1 + 100\theta_2) = 21000 + 12000 - 33000,$$

donc $\theta_0 = 0$. La somme des deux premières équations contient *deux fois* l'intercept, tandis que la troisième ne le contient qu'une fois : leur comparaison détermine l'intercept, sans contradiction.

La différence $(3)-(2)$ donne $100\theta_1 = 21000$, donc $\theta_1 = 210$ mm. La différence $(3)-(1)$ donne $100\theta_2 = 12000$, donc $\theta_2 = 120$ mm. En reportant dans (1), $\theta_0 = 21000 - 100\times 210 = 0$ N mm. Finalement,

$$\boxed{\widehat{\boldsymbol{\theta}} = (0, 210, 120)^{\mathsf{T}}, \qquad \widehat{M}_1 = 210Y_1 + 120Y_2.}$$

La substitution dans les trois équations confirme directement la solution :

$$\begin{aligned} 0 + 100\times 210 &= 21000,\\ 0 + 100\times 120 &= 12000,\\ 0 + 100\times 210 + 100\times 120 &= 21000 + 12000 = 33000. \end{aligned}$$

Le quatrième essai vérifie $200\times 210 + 100\times 120 = 54000$ N mm. Les quatre résidus sont donc nuls : toutes les mesures sont reproduites.

- **Relier cette résolution aux moindres carrés et vérifier l'unicité.**

Le critère quadratique est

$$J(\boldsymbol{\theta}) = \|\boldsymbol{y} - D\boldsymbol{\theta}\|^2 = \boldsymbol{y}^{\mathsf{T}}\boldsymbol{y} - 2\boldsymbol{\theta}^{\mathsf{T}}D^{\mathsf{T}}\boldsymbol{y} + \boldsymbol{\theta}^{\mathsf{T}}D^{\mathsf{T}}D\boldsymbol{\theta}.$$

Sa dérivée vaut $\nabla J = 2D^{\mathsf{T}}(D\boldsymbol{\theta} - \boldsymbol{y})$. L'annulation du gradient conduit aux équations normales

$$\underbrace{\begin{pmatrix} 4 & 400 & 300\\ 400 & 60000 & 30000\\ 300 & 30000 & 30000 \end{pmatrix}}_{D^{\mathsf{T}}D} \widehat{\boldsymbol{\theta}} = \underbrace{\begin{pmatrix} 120000\\ 16200000\\ 9900000 \end{pmatrix}}_{D^{\mathsf{T}}\boldsymbol{y}}.$$

Par exemple, $(D^{\mathsf{T}}D)_{12} = 100+0+100+200 = 400$ et $(D^{\mathsf{T}}\boldsymbol{y})_2 = 100(21000)+100(33000)+200(54000) = 16200000$. La solution obtenue par différences satisfait ce système et donne $J = 0$, minimum possible d'une somme de carrés.

Pour prouver son unicité, supposons $D\boldsymbol{a} = 0$. La troisième ligne moins la deuxième impose $100a_1 = 0$; la troisième moins la première impose $100a_2 = 0$. La première donne alors $a_0 = 0$. Donc $\ker D = \{\mathbf{0}\}$ et $\operatorname{rang}(D) = 3$: les trois coefficients sont identifiables. Le quatrième essai est ici une vérification de cohérence.

> **Interprétation mécanique**
>
> Les deux pentes sont des coefficients d'influence : à charge Y_2 fixée, augmenter Y_1 de 1 N augmente le moment de 210 N mm ; à charge Y_1 fixée, augmenter Y_2 de 1 N l'augmente de 120 N mm. Ces longueurs effectives traduisent la géométrie et l'équilibre de la poutre. Le modèle retrouvé donne un moment nul lorsque les deux charges sont nulles.

Question 2 — Que peut-on identifier avec des charges proportionnelles ?

- **Repérer la perte d'une direction d'essai.**

Cette question porte sur une nouvelle campagne sans bruit sur la même poutre. Les moments correspondent aux nouveaux chargements : les réponses du tableau initial ne sont pas réutilisées. Changer les charges tout en conservant ces anciens moments pourrait rendre le système incompatible. On étudie ici l'information apportée par cette nouvelle campagne seule.

Dans sa matrice de régression, notée à nouveau D, la relation $Y_{2i} = 2Y_{1i}$ impose que la troisième colonne soit deux fois la deuxième et

$$\widehat{M}_1 = \theta_0 + (\theta_1 + 2\theta_2)Y_1 = \theta_0 + qY_1, \qquad q = \theta_1 + 2\theta_2.$$

Si au moins deux valeurs distinctes de Y_1 sont utilisées, les colonnes $\mathbf{1}$ et $(Y_{1i})_i$ sont indépendantes : le rang est 2. On peut déterminer θ_0 et q, mais pas séparer θ_1 et θ_2.

- **Décrire toutes les solutions indiscernables.**

Sous la loi mécanique identifiée à la question 1, les nouveaux moments valent $M_1 = 450Y_1$. Par exemple, les charges $(100, 200)$ N donnent 45000 N mm et les charges $(200, 400)$ N donnent 90000 N mm. Ces deux nouvelles mesures imposent

$$\theta_0 + 100q = 45000, \qquad \theta_0 + 200q = 90000.$$

Par différence, $q = 450$ mm, puis $\theta_0 = 0$ N mm. Tous les couples $\theta_1 = 450 - 2a$, $\theta_2 = a$, avec $a \in \mathbb{R}$, donnent donc la même réponse sur cette droite de chargement. En effet, $D(0, -2, 1)^\mathsf{T} = \mathbf{0}$: cette direction des coefficients est invisible aux essais. Si Y_1 est aussi constant, le rang tombe à 1 ; on ne détermine alors qu'une seule combinaison des trois coefficients.

> **Interprétation mécanique**
> Les modèles $(\theta_1, \theta_2) = (210, 120)$ et $(450, 0)$ prédisent tous deux $M_1 = 450Y_1$ lorsque $Y_2 = 2Y_1$. Pour $Y_1 = 0$ et $Y_2 = 100$ N, ils prédisent pourtant 12000 et 0 N mm. Un ajustement parfait sur des charges proportionnelles ne permet donc pas d'identifier séparément les effets mécaniques. Il faut varier les proportions des deux charges.

Question 3 — Isoler l'effet d'un biais constant du capteur

On revient aux quatre essais du tableau initial et à leur matrice D de rang 3, utilisée dans la question 1. La contrainte de charges proportionnelles de la question 2 ne s'applique plus ici.

Le capteur ajoute $b = 1000$ N mm à chaque réponse. Le vecteur mesuré devient donc $\boldsymbol{y}_b = \boldsymbol{y} + b\mathbf{1}$, avec $\mathbf{1} = (1, 1, 1, 1)^\mathsf{T}$. La première colonne de D est précisément $\mathbf{1}$: pour $\boldsymbol{e}_0 = (1, 0, 0)^\mathsf{T}$, $D\boldsymbol{e}_0 = \mathbf{1}$. Ainsi,

$$\boldsymbol{y}_b = D\widehat{\boldsymbol{\theta}} + bD\boldsymbol{e}_0 = D(\widehat{\boldsymbol{\theta}} + b\boldsymbol{e}_0).$$

Puisque D est de rang 3, cette représentation est unique. Les nouveaux coefficients et la réponse ajustée sont

$$\boxed{\widehat{\theta}_{0,b} = 1000 \text{ N mm}, \qquad \widehat{\theta}_{1,b} = 210 \text{ mm}, \qquad \widehat{\theta}_{2,b} = 120 \text{ mm}},$$

$$\widehat{M}_{1,b} = 1000 + 210Y_1 + 120Y_2.$$

On peut aussi reprendre les différences entre essais : le terme b s'y soustrait, de sorte que les deux pentes restent inchangées. Le premier essai impose ensuite $\theta_{0,b} + 100(210) = 22000$, d'où $\theta_{0,b} = 1000$. Les quatre mesures biaisées sont exactement reproduites.

> **Interprétation mécanique**
> Un décalage constant du capteur est absorbé par l'intercept conservé dans le modèle. Les coefficients d'influence mécaniques restent corrects, mais le moment prédit à charge nulle vaut artificiellement 1000 N mm. L'intercept seul ne permet pas de distinguer un biais de capteur d'un véritable moment initial : un étalonnage ou une mesure de référence est nécessaire pour attribuer une cause physique au décalage.

Exercice 2 Ridge et filtrage discret

On utilise un critère somme de carrés, sans division par le nombre n d'essais. La première colonne de D vaut $\mathbf{1}$ et représente l'intercept.

Question 1 — Dériver les équations normales et protéger l'intercept

- **Développer la fonction à minimiser.**

Comme $P = \mathrm{diag}(0, 1, \ldots, 1)$ est symétrique,

$$\begin{aligned} J_\lambda(\boldsymbol{\theta}) &= (\boldsymbol{y} - D\boldsymbol{\theta})^\mathsf{T}(\boldsymbol{y} - D\boldsymbol{\theta}) + \lambda\boldsymbol{\theta}^\mathsf{T}P\boldsymbol{\theta} \\ &= \boldsymbol{y}^\mathsf{T}\boldsymbol{y} - 2\boldsymbol{\theta}^\mathsf{T}D^\mathsf{T}\boldsymbol{y} + \boldsymbol{\theta}^\mathsf{T}(D^\mathsf{T}D + \lambda P)\boldsymbol{\theta}. \end{aligned}$$

Pour une matrice symétrique A, la dérivée de $\boldsymbol{\theta}^\mathsf{T} A\boldsymbol{\theta}$ est $2A\boldsymbol{\theta}$. Ainsi,

$$\nabla J_\lambda = -2D^\mathsf{T}\boldsymbol{y} + 2(D^\mathsf{T}D + \lambda P)\boldsymbol{\theta}.$$

L'annulation du gradient donne

$$\boxed{(D^\mathsf{T}D + \lambda P)\widehat{\boldsymbol{\theta}} = D^\mathsf{T}\boldsymbol{y}.}$$

Le facteur dépend de la convention : si la perte était $n^{-1}\|\boldsymbol{y} - D\boldsymbol{\theta}\|^2 + \lambda\boldsymbol{\theta}^\mathsf{T}P\boldsymbol{\theta}$, les équations contiendraient $n\lambda P$ au lieu de λP.

- **Vérifier qu'il s'agit d'un minimum et quand il est unique.**

Pour tout vecteur $\boldsymbol{a}$,

$$\boldsymbol{a}^\mathsf{T}(D^\mathsf{T}D + \lambda P)\boldsymbol{a} = \|D\boldsymbol{a}\|^2 + \lambda\sum_{j=1}^{d} a_j^2 \geq 0.$$

Le critère est donc convexe. Pour $\lambda > 0$ et au moins un essai, une valeur nulle impose $a_1 = \cdots = a_d = 0$, puis $D\boldsymbol{a} = a_0\mathbf{1} = 0$, donc $a_0 = 0$. La matrice est définie positive et le minimum est unique, même si les colonnes de charges sont dépendantes. Pour $\lambda = 0$, l'unicité exige que D ait toutes ses colonnes indépendantes.

- **Comprendre le zéro de la pénalisation.**

La pénalité vaut $\lambda\sum_{j=1}^{d}\theta_j^2$: elle ne pénalise pas θ_0. Si les réponses sont décalées de b, le remplacement $\widehat{\theta}_0 \mapsto \widehat{\theta}_0 + b$ conserve les résidus et la pénalité. Pénaliser aussi l'intercept ferait dépendre l'ajustement du choix d'origine du capteur. Les pentes doivent être mises à une échelle comparable avant une pénalisation commune ; changer leurs unités change autrement la force relative de cette pénalisation.

> **Interprétation mécanique**
>
> Ridge stabilise l'estimation face à des mesures bruitées ou à des chargements presque proportionnels, mais ne crée pas l'information manquante. Une solution numérique unique n'implique pas que les coefficients physiques soient déterminés par les seuls essais. Le choix de la pénalité introduit une préférence pour certaines valeurs de coefficients.

Question 2 — Éviter une fuite d'information pendant la validation

- **Identifier le problème.**

Standardiser toutes les entrées avant la séparation utilise déjà les valeurs du test pour déterminer les moyennes et les écarts-types. Même sans utiliser les réponses du test, le protocole n'est plus celui d'un ajustement ayant accès uniquement aux données d'entraînement. Le test ne mesure plus strictement le comportement de cette procédure sur des données nouvelles.

- **Décrire la procédure, dans l'ordre.**

  - Réserver un test final indépendant avant toute transformation ou sélection de paramètre ; ne plus l'utiliser jusqu'à l'évaluation finale.
  - Partager les seules données d'entraînement en plis. Pour chaque pli de validation v, calculer sur les autres plis uniquement

$$\mu_j^{(-v)} = \frac{1}{n_{-v}}\sum_{i\notin v} x_{ij}, \qquad (s_j^{(-v)})^2 = \frac{1}{n_{-v}}\sum_{i\notin v}(x_{ij} - \mu_j^{(-v)})^2.$$

  - Transformer apprentissage et validation par les mêmes valeurs $z_{ij} = (x_{ij} - \mu_j^{(-v)})/s_j^{(-v)}$. Une colonne constante doit être traitée séparément ; on ne divise pas par un écart-type nul et on ne standardise pas la colonne d'intercept.
  - Pour chaque λ candidat, ajuster ridge sur l'apprentissage du pli, puis calculer l'erreur sur sa validation. Choisir λ à partir de la moyenne des erreurs de validation, jamais à partir du test final.

- Recalculer la transformation sur tout l'entraînement, réajuster le modèle avec le λ choisi, puis évaluer une seule fois sur le test.

Si seule l'entrée est standardisée, le retour aux unités initiales est

$$\theta_j^{\text{phys}} = \frac{\theta_j^{\text{std}}}{s_j}, \qquad \theta_0^{\text{phys}} = \theta_0^{\text{std}} - \sum_{j=1}^{d} \frac{\theta_j^{\text{std}} \mu_j}{s_j}.$$

Cette formule vient directement du développement de $\theta_0^{\text{std}} + \sum_j \theta_j^{\text{std}}(x_j - \mu_j)/s_j$.

Interprétation mécanique

La validation doit représenter une nouvelle campagne d'essais. Si plusieurs mesures appartiennent à la même pièce ou à la même trajectoire, les séparer au hasard peut encore transmettre une information commune. On réserve alors des pièces, trajectoires ou blocs entiers, selon l'usage visé. Une petite erreur sur des données trop proches ne garantit pas la prédiction d'une nouvelle structure ou d'un nouveau chargement.

Question 3 — Construire la matrice des retards et l'orthogonalité

- **Compter les lignes disponibles.**

Pour former une ligne, il faut disposer de p entrées successives. La première réponse utilisable est y_{p-1}, la dernière y_{N-1} : il y a $N - p + 1$ lignes et p colonnes. Par conséquent,

$$D = \begin{pmatrix} x_{p-1} & x_{p-2} & \cdots & x_0 \\ x_p & x_{p-1} & \cdots & x_1 \\ \vdots & \vdots & & \vdots \\ x_{N-1} & x_{N-2} & \cdots & x_{N-p} \end{pmatrix}, \qquad \boldsymbol{y} = \begin{pmatrix} y_{p-1} \\ y_p \\ \vdots \\ y_{N-1} \end{pmatrix}.$$

Le modèle s'écrit $\boldsymbol{y} = D\boldsymbol{h} + \boldsymbol{\varepsilon}$ avec $\boldsymbol{h} = (h_0, \ldots, h_{p-1})^{\mathsf{T}}$. Par exemple, pour $p = 2$, la ligne d'indice n est (x_n, x_{n-1}) : le premier coefficient multiplie l'entrée actuelle, le second l'entrée précédente.

- **Dériver et interpréter le minimum des moindres carrés.**

Pour $L(\boldsymbol{h}) = \|\boldsymbol{y} - D\boldsymbol{h}\|^2$, on a $\nabla L = 2D^{\mathsf{T}}(D\boldsymbol{h} - \boldsymbol{y})$. Ainsi, au minimum, le vecteur des résidus $\boldsymbol{e} = \boldsymbol{y} - D\widehat{\boldsymbol{h}}$ satisfait

$$D^{\mathsf{T}}\boldsymbol{e} = \boldsymbol{0}, \qquad \sum_{n=p-1}^{N-1} x_{n-j} \left(y_n - \sum_{\ell=0}^{p-1} \widehat{h}_\ell x_{n-\ell} \right) = 0 \quad (0 \le j \le p-1).$$

L'orthogonalité est exacte pour les produits scalaires empiriques : aucune correction linéaire le long d'une colonne de D ne diminue encore la somme des carrés. Elle n'implique ni indépendance des erreurs ni absence de structure non linéaire dans les résidus. L'unicité des coefficients exige le rang p. Il faut donc au moins $N - p + 1 \ge p$, soit $N \ge 2p - 1$; cette condition est nécessaire, mais ne garantit pas l'indépendance des colonnes. La condition $N \ge p$ de l'énoncé garantit seulement qu'au moins une ligne peut être construite. Avec une pénalité $\lambda\|\boldsymbol{h}\|^2$ sur tous les coefficients du filtre (sans intercept ici), on aurait au contraire $D^{\mathsf{T}}\boldsymbol{e} = \lambda\widehat{\boldsymbol{h}}$.

Interprétation mécanique

Si x_n est une force et y_n un déplacement, chaque h_j, en mm/N, mesure la contribution de la force reçue j pas de temps plus tôt. Le filtre représente la mémoire mécanique sur la fenêtre retenue. Des entrées suffisamment variées sont nécessaires pour séparer ces contributions. Le recouvrement des fenêtres appelle une validation adaptée à l'usage visé, par blocs temporels ou par trajectoires par exemple. Il ne suffit toutefois pas, à lui seul, à établir une dépendance des erreurs conditionnellement aux entrées : avec des entrées fixées et des erreurs indépendantes, les erreurs restent indépendantes malgré le partage de certaines entrées.

Exercice 3 Un processus gaussien à une observation

L'entrée et la réponse sont normalisées. Le processus gaussien f a une moyenne a priori nulle et le noyau

$$k(x, x') = \exp\left(-\frac{(x - x')^2}{2}\right).$$

L'unique observation est exacte : $f(0) = 2$. On note $F_0 = f(0)$, $F_x = f(x)$ et $c(x) = e^{-x^2/2}$.

Question 1 — Déterminer la moyenne et la variance conditionnelles

- **Construire la loi jointe au point observé et au point étudié.**

La définition d'un processus gaussien impose une loi gaussienne pour toute paire de valeurs. Le noyau fournit

$$\mathrm{Var}(F_0) = k(0, 0) = 1, \qquad \mathrm{Var}(F_x) = k(x, x) = 1,$$
$$\mathrm{Cov}(F_0, F_x) = k(0, x) = e^{-x^2/2} = c.$$

Les moyennes étant nulles,

$$\begin{pmatrix} F_0 \\ F_x \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}\right).$$

Pour $x \neq 0$, $0 < c < 1$ et cette matrice est définie positive. Pour $x = 0$, les deux variables sont identiques : la loi jointe est dégénérée, ce qui correspond exactement à l'observation du point prédit.

- **Démontrer le conditionnement par un résidu indépendant.**

Posons $R_x = F_x - cF_0$, de sorte que $F_x = cF_0 + R_x$. Ce résidu est centré. Sa covariance avec la variable observée vaut

$$\mathrm{Cov}(R_x, F_0) = \mathrm{Cov}(F_x, F_0) - c\,\mathrm{Var}(F_0) = c - c = 0.$$

Le couple (R_x, F_0) est gaussien, car c'est une transformation linéaire de (F_x, F_0). Dans un couple gaussien, une covariance nulle entraîne l'indépendance. La loi de R_x ne change donc pas lorsque F_0 est observé. Sa variance se calcule en conservant le terme de covariance :

$$\begin{aligned}\mathrm{Var}(R_x) &= \mathrm{Var}(F_x) + c^2\,\mathrm{Var}(F_0) - 2c\,\mathrm{Cov}(F_x, F_0) \\ &= 1 + c^2 - 2c^2 = 1 - c^2.\end{aligned}$$

- **Introduire la valeur observée et conclure.**

Conditionner par $F_0 = 2$ fixe le terme cF_0 à $2c$, tandis que le résidu garde sa moyenne nulle et sa variance $1 - c^2$. Ainsi,

$$\boxed{m_{\mathrm{cond}}(x) = \mathbb{E}[F_x \mid F_0 = 2] = 2e^{-x^2/2},}$$

$$\boxed{s^2_{\mathrm{cond}}(x) = \mathrm{Var}(F_x \mid F_0 = 2) = 1 - e^{-x^2}.}$$

La loi conditionnelle est donc $\mathcal{N}(2e^{-x^2/2}, 1 - e^{-x^2})$ pour $x \neq 0$. En $x = 0$, elle est concentrée sur la valeur observée 2.

> **Interprétation mécanique**
>
> Si f représente une réponse mécanique normalisée, l'essai en $x = 0$ renseigne les valeurs voisines selon la corrélation $c(x)$ imposée par le noyau. La moyenne conditionnelle est la prédiction après cet essai ; la variance conditionnelle mesure ce qui reste inconnu sur la réponse, dans le modèle choisi. Elle ne décrit pas une dispersion de charges en exploitation. Ici, la mesure est exacte : aucun bruit de capteur n'intervient.

Question 2 — Évaluer au point observé et étudier les limites

- **Retrouver l'observation en $x = 0$.**

Comme $e^0 = 1$, la substitution dans les deux formules donne

$$\boxed{m_{\text{cond}}(0) = 2, \qquad s^2_{\text{cond}}(0) = 0.}$$

On peut le vérifier directement : après avoir observé $f(0) = 2$, la valeur de $f(0)$ est connue. L'espérance vaut donc 2 et la variance, qui est l'espérance du carré de l'écart à cette valeur, est nulle. Le modèle interpole exactement la mesure.

- **S'éloigner de l'essai.**

Lorsque $|x| \to \infty$, $x^2 \to \infty$, donc $e^{-x^2/2} \to 0$ et $e^{-x^2} \to 0$. Par conséquent,

$$\boxed{\lim_{|x|\to\infty} m_{\text{cond}}(x) = 0, \qquad \lim_{|x|\to\infty} s^2_{\text{cond}}(x) = 1.}$$

Ces valeurs sont la moyenne et la variance *a priori*. L'observation cesse d'influencer la prédiction lorsque la covariance avec le point observé tend vers zéro. Cette conclusion vient du noyau choisi ; ce n'est pas une propriété universelle de toutes les réponses mécaniques.

- **Vérifier la cohérence des résultats entre ces deux régimes.**

Pour toute entrée, $0 < e^{-x^2} \leq 1$. La variance vérifie donc $0 \leq s^2_{\text{cond}}(x) < 1$ à distance finie, et tend vers 1 à l'infini. Les deux expressions sont paires : les prédictions en x et $-x$ sont identiques, conformément à la symétrie du noyau autour de l'unique essai. Pour $x > 0$,

$$\frac{\mathrm{d}m_{\text{cond}}}{\mathrm{d}x} = -2xe^{-x^2/2} < 0, \qquad \frac{\mathrm{d}s^2_{\text{cond}}}{\mathrm{d}x} = 2xe^{-x^2} > 0.$$

La moyenne décroît vers 0 tandis que l'incertitude augmente. Par exemple, en $x = 1$, $m_{\text{cond}}(1) = 2e^{-1/2} \simeq 1{,}2131$ et $s^2_{\text{cond}}(1) = 1 - e^{-1} \simeq 0{,}6321$.

> **Interprétation mécanique**
>
> Près de l'essai, la réponse est fortement renseignée ; loin de l'essai, le modèle revient à ses hypothèses initiales. Une moyenne qui tend vers zéro ne prouve donc pas que le déplacement ou le moment réel tend vers zéro. Elle indique surtout une perte d'information issue des données. La variance redevient celle du modèle a priori. Une extrapolation mécanique demande des essais supplémentaires ou des hypothèses physiques adaptées. Si la réponse normalisée est $f = (M - M_{\text{ref}})/M_{\text{ech}}$, avec $M_{\text{ech}} > 0$, la moyenne du moment vaut $M_{\text{ref}} + M_{\text{ech}} m_{\text{cond}}$ et sa variance $M^2_{\text{ech}} s^2_{\text{cond}}$: il faut remettre les unités avant d'interpréter un niveau de moment ou son incertitude.

Question 3 — Distinguer la réponse latente et une nouvelle mesure bruitée

- **Écrire la nouvelle variable observée.**

On observe désormais $Y_0 = F_0 + \varepsilon_0 = 2$, où $\varepsilon_0 \sim \mathcal{N}(0, \eta^2)$ est indépendante du processus. Ce nouveau scénario remplace l'observation exacte : on ne sait plus que $F_0 = 2$. L'entrée, la réponse et les erreurs étant normalisées, η^2 est ici une variance sans dimension. Avec $c = e^{-x^2/2}$,

$$\mathbb{E}[Y_0] = 0, \qquad \mathrm{Var}(Y_0) = 1 + \eta^2, \qquad \mathrm{Cov}(F_x, Y_0) = \mathrm{Cov}(F_x, F_0) = c.$$

Le couple (Y_0, F_x) est donc gaussien centré, de covariance $\left(\begin{smallmatrix} 1+\eta^2 & c \\ c & 1 \end{smallmatrix}\right)$.

- **Calculer la prédiction de la réponse physique latente.**

Posons $a = c/(1+\eta^2)$ et $R_x = F_x - aY_0$. Alors

$$\mathrm{Cov}(R_x, Y_0) = c - a(1+\eta^2) = 0.$$

Comme précédemment, la gaussianité transforme cette covariance nulle en indépendance. De plus,

$$\begin{aligned}\mathrm{Var}(R_x) &= 1 + a^2(1+\eta^2) - 2ac \\ &= 1 + \frac{c^2}{1+\eta^2} - \frac{2c^2}{1+\eta^2} = 1 - \frac{c^2}{1+\eta^2}.\end{aligned}$$

En conditionnant par $Y_0 = 2$, la décomposition $F_x = aY_0 + R_x$ donne

$$\boxed{m_f(x) = \frac{2e^{-x^2/2}}{1+\eta^2}, \qquad s_f^2(x) = 1 - \frac{e^{-x^2}}{1+\eta^2}.}$$

La réponse latente suit conditionnellement la loi $\mathcal{N}(m_f(x), s_f^2(x))$, dégénérée au point observé seulement dans le cas limite sans bruit.

- **Ajouter le bruit d'une mesure future indépendante.**

Une nouvelle mesure est $Y_{\text{nouv}}(x) = F_x + \varepsilon_{\text{nouv}}$. Le nouveau bruit est indépendant de F_x et de Y_0 : après conditionnement, sa moyenne reste nulle, sa variance reste η^2 et sa covariance avec F_x reste nulle. Il en résulte

$$\boxed{m_{\text{nouv}}(x) = m_f(x), \qquad s_{\text{nouv}}^2(x) = s_f^2(x) + \eta^2.}$$

La moyenne prédite est la même, mais une future lecture du capteur a une variance supérieure de η^2. Il ne faut pas ajouter ce terme à la variance de la réponse physique latente.

- **Contrôler les limites et le point déjà mesuré.**

En $x = 0$, la covariance c vaut 1. Par conséquent,

$$m_f(0) = m_{\text{nouv}}(0) = \frac{2}{1+\eta^2}, \qquad s_f^2(0) = \frac{\eta^2}{1+\eta^2}, \qquad s_{\text{nouv}}^2(0) = \frac{\eta^2}{1+\eta^2} + \eta^2.$$

Si $\eta^2 > 0$, la réponse n'est plus interpolée exactement : la mesure peut s'écarter de la valeur vraie. La nouvelle observation au même point possède aussi son propre bruit. En revanche, la mesure *déjà réalisée* Y_0 a une variance conditionnelle nulle, puisque sa valeur 2 est connue.

Lorsque $|x| \to \infty$, $c \to 0$: les deux moyennes tendent vers 0, la variance latente vers 1 et la variance d'observation vers $1 + \eta^2$. Lorsque $\eta^2 \to 0$, on retrouve les réponses exactes des questions 1 et 2. Lorsque $\eta^2 \to \infty$, à x fixé, la mesure devient peu informative : $m_f(x) \to 0$ et $s_f^2(x) \to 1$. Par exemple, pour $\eta^2 = 1$ en $x = 0$, la moyenne prédite est 1, la variance latente $1/2$ et celle d'une nouvelle mesure $3/2$.

Interprétation mécanique

Deux incertitudes sont séparées : ce qui reste inconnu sur la réponse de la structure et la variabilité propre à une future mesure. Augmenter le bruit diminue le poids de l'essai dans l'apprentissage. Un capteur très bruité ne renseigne presque plus le modèle, même au point où il a mesuré. Pour un moment réel $M = M_{\text{ref}} + M_{\text{ech}} f$, les variances latente et d'observation doivent toutes deux être multipliées par M_{ech}^2 pour retrouver des unités de moment au carré.

Exercice 4 Modèle appris et événement de rupture

La poutre mesure 1000 mm ; les forces sont appliquées à 300 et 600 mm. Avec $W = 1800$ mm^3, la réponse exacte, en MPa, est

$$\sigma(\boldsymbol{Y}) = \frac{\max(|210Y_1 + 120Y_2|, |120Y_1 + 240Y_2|)}{1800}.$$

On note $\widehat{\sigma}$ la moyenne prédictive du GP, dans les mêmes unités. Le seuil de rupture est $\sigma_a = 140$ MPa. Aucun ajustement numérique n'est nécessaire pour répondre : il s'agit de concevoir et justifier un protocole.

Question 1 — Retrouver le changement de section critique

- **Retirer les valeurs absolues uniquement dans le domaine positif.**

Pour $Y_1, Y_2 > 0$, les deux moments sont positifs et

$$M_2 - M_1 = (120 - 210)Y_1 + (240 - 120)Y_2 = -90Y_1 + 120Y_2.$$

Ils sont égaux si $120Y_2 = 90Y_1$, soit

$$\boxed{Y_2 = \tfrac{3}{4}Y_1.}$$

En dessous, $M_2 - M_1 < 0$: la section sous Y_1 est critique. Au-dessus, la section sous Y_2 est critique. Sur la droite, les deux sections atteignent la même contrainte.

- **Distinguer changement de branche et seuil de rupture.**

La droite précédente ne dit pas si la structure est sûre. Pour des charges positives, la sûreté demande simultanément

$$210Y_1 + 120Y_2 < 252000, \qquad 120Y_1 + 240Y_2 < 252000,$$

car $140 \times 1800 = 252000$ N mm. À $Y_1 = 800$ N, le changement de branche intervient à $Y_2 = 600$ N, où $\sigma = 240000/1800 = 133{,}33$ MPa. Sur la branche suivante, le seuil est atteint lorsque $96000 + 240Y_2 = 252000$, donc à $Y_2 = 650$ N. La réponse est continue au changement de branche, mais sa pente change. Pour des charges de signes quelconques, il faut conserver les valeurs absolues.

> **Interprétation mécanique**
> Un modèle appris doit restituer deux propriétés distinctes : quelle section est la plus sollicitée et si sa contrainte dépasse le seuil. Une réponse lisse peut approcher convenablement les valeurs tout en lissant le changement de pente ; les erreurs doivent donc être inspectées près de cette ligne et près de la frontière de rupture.

Question 2 — Concevoir l'apprentissage et une évaluation indépendante

- **Fixer les ensembles et ce qu'ils représentent.**

On peut tirer uniformément 120 points dans $[300, 1300] \times [100, 900]$ N^2 et calculer leurs contraintes exactes. Ce plan explore un domaine ; il ne représente pas la loi des charges en service. On réserve séparément 400 autres points, indépendants des 120 premiers, sur le même rectangle pour le test final. Les calculs exacts fournissent les références, sans bruit de mesure dans cette expérience numérique.

- **Choisir le modèle à l'intérieur des 120 points.**

Une procédure possible utilise cinq plis de 24 points : à chaque passage, 96 points servent à l'ajustement et 24 à la validation. Pour chaque noyau ou jeu d'hyperparamètres candidat, on calcule les transformations uniquement sur les 96 points, puis on les applique aux 24 autres. On compare les erreurs de validation et leur comportement près du seuil. Le choix terminé, on réestime la transformation sur les 120 points et on ajuste le GP final. Les 400 points ne servent ni à choisir la normalisation ni à sélectionner les hyperparamètres. Une éventuelle petite diagonale de stabilisation numérique ne doit pas être présentée comme un bruit physique mesuré.

- **Évaluer des erreurs exprimées en MPa.**

Pour chaque point test, poser $e_i = \widehat{\sigma}(\boldsymbol{Y}_i) - \sigma(\boldsymbol{Y}_i)$. Présenter au minimum

$$\mathrm{RMSE} = \sqrt{\frac{1}{400}\sum_{i=1}^{400} e_i^2}, \qquad E_{\max} = \max_{1 \le i \le 400} |e_i|, \qquad \overline{e} = \frac{1}{400}\sum_{i=1}^{400} e_i.$$

La RMSE mesure une erreur quadratique moyenne, $E_{\max}$ signale les écarts extrêmes du test et $\overline{e}$ un décalage systématique. On ajoute un graphe « prédit contre exact », une carte des erreurs dans le plan des charges, et une coupe traversant $Y_2 = 0{,}75Y_1$ puis $\sigma = 140$ MPa. Ces figures seraient les résultats attendus du protocole, pas des données à inventer dans cette correction.

Si le GP fournit un écart-type latent s_i, on peut aussi mesurer

$$\widehat{c}_{95} = \frac{1}{400}\sum_{i=1}^{400} \mathbf{1}_{\{|e_i| \le 1{,}96 s_i\}}.$$

C'est une couverture empirique ponctuelle sur le test ; un intervalle crédible gaussien conditionnel au modèle n'assure pas à lui seul 95 % de couverture de la vraie réponse mécanique. Il faut restituer les unités si les réponses ont été normalisées, et examiner les erreurs de classification au seuil.

Interprétation mécanique

Une faible RMSE globale peut masquer quelques erreurs sur les chargements qui décident de la sûreté. Les 400 points quantifient la qualité sur le domaine de test, tandis que la population physique de la question suivante permet d'évaluer le risque d'exploitation. Ces deux évaluations répondent à des questions différentes.

Question 3 — Comparer les décisions sur les mêmes charges physiques

- **Reproduire la loi physique commune aux deux modèles.**

Le vecteur des charges a pour moyenne $(800, 400)^{\mathsf{T}}$ N et pour covariance $V = 10^4 \left(\begin{smallmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{smallmatrix}\right)$ N^2. Une représentation utile est

$$\boldsymbol{Y} = \begin{pmatrix} 800 \\ 400 \end{pmatrix} + \underbrace{\begin{pmatrix} 100 & 0 \\ 50 & 50/\sqrt{3} \end{pmatrix}}_{C} \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix}, \qquad Z_1, Z_2 \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1).$$

En effet, $CC^{\mathsf{T}} = \left(\begin{smallmatrix} 10000 & 5000 \\ 5000 & 10000/3 \end{smallmatrix}\right) = V$. On utilise les mêmes N_{MC} couples pour les deux réponses ; on ne remplace pas cette loi par le tirage uniforme du domaine d'apprentissage. La loi gaussienne n'étant pas tronquée, on conserve les valeurs absolues dans σ et on repère les éventuelles charges hors domaine appris.

- **Construire les indicatrices et les quatre catégories.**

Pour chaque charge commune $\boldsymbol{Y}_a$, définir

$$I_a = \mathbf{1}_{\{\sigma(\boldsymbol{Y}_a) \ge 140\}}, \qquad J_a = \mathbf{1}_{\{\widehat{\sigma}(\boldsymbol{Y}_a) \ge 140\}}.$$

Les deux fréquences de rupture sont les moyennes de I_a et de J_a. On compte également les quatre catégories suivantes :

I_a	J_a	Décision du substitut
0	0	Sûreté correctement reconnue (vrai négatif)
1	1	Rupture correctement reconnue (vrai positif)
0	1	Fausse alarme (faux positif)
1	0	Rupture manquée (faux négatif)

On présente la fréquence totale de désaccord, les deux fréquences de fausse alarme et de rupture manquée, ainsi que les points concernés dans le plan des charges. Il faut préciser le dénominateur : une fréquence de faux négatifs sur tous les tirages diffère de la proportion manquée parmi les ruptures exactes. Cette dernière n'est pas définie si le test ne contient aucune rupture exacte.

- **Comparer de manière appariée et lire les conséquences.**

Poser $D_a = J_a - I_a \in \{-1, 0, 1\}$. Alors

$$\widehat{P}_{f,\mathrm{GP}} - \widehat{P}_{f,\mathrm{exact}} = \frac{1}{N_{\mathrm{MC}}}\sum_a D_a = \frac{n_{\text{faux positifs}} - n_{\text{faux négatifs}}}{N_{\mathrm{MC}}}.$$

Deux fréquences proches peuvent donc cacher des erreurs qui se compensent. Le GP final étant fixé, pour $N_{\mathrm{MC}} \geq 2$ charges tirées indépendamment selon la loi physique, posons

$$\overline{D} = \frac{1}{N_{\mathrm{MC}}} \sum_a D_a, \qquad s_D^2 = \frac{1}{N_{\mathrm{MC}} - 1} \sum_a (D_a - \overline{D})^2.$$

La variance estimée de cette *différence appariée* est s_D^2 / N_{MC} et son erreur standard estimée vaut

$$\widehat{\mathrm{SE}}(\overline{D}) = \sqrt{\frac{s_D^2}{N_{\mathrm{MC}}}}.$$

Les deux classements ne sont pas indépendants : il faut conserver leur appariement pour estimer cette précision. Cette erreur standard quantifie l'erreur d'échantillonnage Monte-Carlo, pas l'erreur du GP.

Interprétation mécanique

Une rupture manquée est la conséquence la plus défavorable pour la sécurité : le substitut annonce une structure admissible alors que le calcul exact dépasse le seuil. Une fausse alarme peut au contraire provoquer un surdimensionnement ou une inspection inutile. L'absence de rupture manquée dans un échantillon fini ne prouve pas son absence dans la population.

Question 4 — Démontrer et interpréter la borne par le désaccord

- **Examiner les quatre couples de valeurs.**

Pour $(I, J) = (0, 0)$ ou $(1, 1)$, on a $|I - J| = 0$. Pour $(0, 1)$ ou $(1, 0)$, on a $|I - J| = 1$. D'où l'identité exacte

$$|I - J| = \mathbf{1}_{\{I \neq J\}}.$$

Les deux indicatrices étant définies sur la même charge aléatoire,

$$\begin{aligned} |P_{f,\mathrm{exact}} - P_{f,\mathrm{GP}}| &= |\mathbb{E}[I] - \mathbb{E}[J]| = |\mathbb{E}[I - J]| \\ &\leq \mathbb{E}[|I - J|] = \mathbb{P}(I \neq J). \end{aligned}$$

La borne est donc

$$\boxed{|P_{f,\mathrm{exact}} - P_{f,\mathrm{GP}}| \leq \mathbb{P}(I \neq J).}$$

La même preuve sur une somme finie donne la version empirique, pourvu que les deux modèles soient évalués sur les mêmes tirages.

- **Comprendre ce que la borne ne dit pas.**

Si $p_{01} = \mathbb{P}(I = 0, J = 1)$ et $p_{10} = \mathbb{P}(I = 1, J = 0)$, l'écart signé vaut $p_{01} - p_{10}$ et le désaccord vaut $p_{01} + p_{10}$. Par exemple, des valeurs égales des deux catégories peuvent conduire à un écart nul alors que chaque catégorie a une probabilité non nulle. Une bonne estimation de la probabilité globale ne garantit donc pas une bonne décision pour chaque chargement.

Interprétation mécanique

La borne justifie de contrôler directement les changements de classe au seuil, au-delà de la seule erreur de régression. Elle sépare l'évaluation globale du risque et la fiabilité de chaque décision mécanique. Le contrôle des faux négatifs reste indispensable, même si les probabilités de rupture calculées par les deux modèles paraissent très proches.

Exercice 5 Identifier une raideur et quantifier sa précision

Le modèle est $F_i = ku_i + \varepsilon_i$, avec $u = (1,2,3)^\mathsf{T}$ mm, $\boldsymbol{F} = (11,19,31)^\mathsf{T}$ N et des erreurs gaussiennes indépendantes, centrées, d'écart-type 2 N (variance 4 N^2). La raideur k est un paramètre fixe inconnu ; les déplacements sont connus et le zéro du capteur est étalonné.

Question 1 — Estimer la raideur par les moindres carrés

Le critère $J(k) = \sum_i (F_i - ku_i)^2$ a pour dérivée

$$J'(k) = -2\sum_i u_i(F_i - ku_i), \qquad J''(k) = 2\sum_i u_i^2 > 0.$$

Son unique minimum est donc

$$\widehat{k} = \frac{\sum_i u_i F_i}{\sum_i u_i^2} = \frac{1(11)+2(19)+3(31)}{1^2+2^2+3^2} = \frac{142}{14} = \frac{71}{7} \simeq 10{,}1429 \text{ N/mm}.$$

Le numérateur a l'unité N mm, le dénominateur mm^2 : le rapport est bien une raideur. Les résidus valent

$$(F_i - \widehat{k}u_i)_{i=1}^3 = \left(\tfrac{6}{7}, -\tfrac{9}{7}, \tfrac{4}{7}\right), \qquad \sum_i u_i(F_i - \widehat{k}u_i) = \tfrac{6-18+12}{7} = 0.$$

Ce contrôle retrouve l'orthogonalité des moindres carrés.

> **Interprétation mécanique**
> La pente estime la force supplémentaire nécessaire pour produire un millimètre de déplacement. L'origine imposée est justifiée par le modèle élastique et l'étalonnage ; un effort initial ou un biais de capteur rendrait ce choix discutable et pourrait biaiser la raideur.

Question 2 — Quantifier la précision et construire l'intervalle

- **Exprimer l'erreur d'estimation comme combinaison des bruits.**

En remplaçant F_i par $ku_i + \varepsilon_i$,

$$\widehat{k} = k + \frac{\sum_i u_i \varepsilon_i}{\sum_i u_i^2}.$$

Une combinaison linéaire d'erreurs gaussiennes est gaussienne. Le numérateur du second terme est centré ; par indépendance,

$$\mathbb{E}[\widehat{k}] = k, \qquad \mathrm{Var}(\widehat{k}) = \frac{\sum_i u_i^2 \sigma_F^2}{(\sum_i u_i^2)^2} = \frac{4}{14} = \frac{2}{7} \ (\text{N/mm})^2.$$

Ainsi $\widehat{k} \sim \mathcal{N}(k, 2/7)$ et son écart-type vaut $\sqrt{2/7} \simeq 0{,}5345$ N/mm. Il s'agit de la dispersion de l'estimateur sur des campagnes répétées, pas d'une variabilité physique attribuée à la raideur k de cette même pièce.

- **Inverser l'inégalité normale.**

Le pivot $(\widehat{k} - k)/\sqrt{2/7}$ est normal centré réduit. En utilisant le quantile 1,96, on obtient l'intervalle de confiance à 95 %

$$\widehat{k} \pm 1{,}96\sqrt{2/7} = 10{,}1429 \pm 1{,}0477, \qquad \boxed{k \in [9{,}0952; 11{,}1905] \text{ N/mm}.}$$

Sur des répétitions du même protocole, environ 95 % des intervalles construits ainsi contiendraient la valeur fixe k. Le calcul utilise une variance de capteur connue, issue de l'étalonnage. Si elle était estimée sur les trois résidus, on ne pourrait pas conserver automatiquement ce même pivot normal et cette même largeur.

> **Interprétation mécanique**
> Trois essais ne donnent pas une raideur exacte. L'intervalle quantifie la précision statistique sous les hypothèses du modèle : il ne couvre ni une erreur sur le déplacement ni une non-linéarité mécanique. Un intervalle étroit peut donc rester trompeur si le domaine élastique ou l'étalonnage ne sont pas respectés.

Question 3 — Séparer la force moyenne et une future lecture du capteur

- **Propager l'erreur d'estimation vers la force moyenne.**

À $u_\star = 4$ mm, la force moyenne vraie est $\mu_\star = 4k$ et son estimateur est

$$\widehat{\mu}_\star = 4\widehat{k} = \frac{284}{7} \simeq 40{,}5714 \text{ N}.$$

L'erreur $\widehat{\mu}_\star - \mu_\star = 4(\widehat{k} - k)$ a une loi normale centrée de variance $16(2/7) = 32/7$ N^2. L'intervalle de confiance est

$$40{,}5714 \pm 1{,}96\sqrt{32/7}, \qquad \boxed{\mu_\star \in [36{,}3808; 44{,}7621] \text{ N}.}$$

- **Ajouter le bruit indépendant de la prochaine mesure.**

Une nouvelle force mesurée vaut $F_\star = 4k + \varepsilon_\star$. Son erreur de prédiction est

$$F_\star - \widehat{\mu}_\star = \varepsilon_\star - 4(\widehat{k} - k).$$

Le nouveau bruit est indépendant des erreurs d'apprentissage. Les deux termes sont donc indépendants, centrés et gaussiens, d'où

$$\mathrm{Var}(F_\star - \widehat{\mu}_\star) = 4 + \frac{32}{7} = \frac{60}{7} \text{ N}^2.$$

L'intervalle de prédiction à 95 % est

$$40{,}5714 \pm 1{,}96\sqrt{60/7}, \qquad \boxed{F_\star \in [34{,}8331; 46{,}3097] \text{ N}.}$$

La couverture annoncée est celle du protocole sur les données d'entraînement et la future mesure répétées. Ce raisonnement fréquentiste ne suppose pas que k devienne aléatoire après observation des trois forces.

> **Interprétation mécanique**
> La force moyenne décrit la loi élastique ; la prochaine lecture contient en plus le bruit du capteur. L'intervalle de prédiction est donc plus large. Le point 4 mm se situe au-delà des trois déplacements mesurés : la validité du modèle linéaire jusque-là doit être vérifiée mécaniquement.

Question 4 — Comparer deux essais supplémentaires

L'information apportée par les déplacements intervient par la somme de leurs carrés. Après un quatrième essai à u_{nouv}, de même bruit,

$$\mathrm{Var}(\widehat{k}_{\text{nouv}}) = \frac{4}{14 + u_{\text{nouv}}^2}.$$

La variance est connue avant la réponse mesurée, car elle dépend du plan d'essais et du bruit, non des valeurs des forces. Les deux possibilités donnent

$$u_{\text{nouv}} = 1 \text{ mm} : \quad \frac{4}{15} \simeq 0{,}2667, \qquad u_{\text{nouv}} = 4 \text{ mm} : \quad \frac{4}{30} \simeq 0{,}1333 \quad (\text{N/mm})^2.$$

Le second choix réduit la variance initiale $2/7$ d'environ 53,3 %, contre 6,7 % pour le premier. On ne peut pas calculer la nouvelle valeur numérique de $\widehat{k}$ sans connaître la quatrième force : seule sa précision prévue est comparée ici.

> **Interprétation mécanique**
> Un déplacement plus élevé rend la pente plus visible face à un bruit de force constant. Il n'est préférable que si l'essai reste dans le domaine élastique linéaire, sans endommagement, jeu, hystérésis ni saturation du capteur. Si le bruit augmente avec la force ou si la mesure de déplacement est incertaine, la comparaison doit être revue.

Exercice 6 Choisir un essai près du seuil de rupture

Toutes les lois ci-dessous sont conditionnelles aux données déjà disponibles $\mathcal{D}$, omises dans les notations pour alléger les calculs. Les moyennes sont $m_A = 138$, $m_B = 130$ MPa et la covariance latente est

$$\Sigma = \begin{pmatrix} 16 & 20 \\ 20 & 100 \end{pmatrix} \text{ MPa}^2.$$

Elle est définie positive puisque $16 > 0$ et $\det \Sigma = 1600 - 400 = 1200 > 0$. Ainsi $s_A = 4$ MPa, $s_B = 10$ MPa et la corrélation vaut $20/(4 \times 10) = 0{,}5$. Chaque nouvel essai a un bruit indépendant de variance $\eta^2 = 4$ MPa^2.

Question 1 — Comparer les dépassements et l'ambiguïté au seuil

Pour une variable $S_C \sim \mathcal{N}(m_C, s_C^2)$, la standardisation donne

$$\mathbb{P}(S_C \geq 140) = 1 - \Phi\left(\frac{140 - m_C}{s_C}\right).$$

Il n'y a pas de masse au seuil pour ces lois continues. Donc

$$\boxed{p_A = 1 - \Phi(0{,}5) \simeq 0{,}3085, \qquad p_B = 1 - \Phi(1) \simeq 0{,}1587.}$$

Les deux moyennes sont en dessous de 140 MPa, mais aucune des deux probabilités de dépassement conditionnelles n'est nulle. Les distances normalisées demandées sont

$$U_A = \frac{|138 - 140|}{4} = 0{,}5, \qquad U_B = \frac{|130 - 140|}{10} = 1.$$

Le critère de décision le plus ambigu retient donc A. Sa probabilité est plus proche de $1/2$, situation où le modèle hésite davantage sur le côté du seuil, même si la variance absolue est plus grande en B.

> **Interprétation mécanique**
> Le chargement A est plus intéressant pour préciser une décision proche de la limite admissible. Une forte variance ne suffit pas à fixer la priorité : la distance de la moyenne au seuil compte aussi. Ce choix reste une comparaison entre les deux candidats proposés, et non la preuve qu'ils seraient les meilleurs essais parmi tous les chargements possibles.

Question 2 — Comparer la réduction de la somme des variances

- **Utiliser le conditionnement d'une observation scalaire.**

Pour un essai $Z_C = S_C + \varepsilon$, la variance observée est $v_C = s_C^2 + 4$ et le vecteur de covariance avec $\boldsymbol{S} = (S_A, S_B)^\mathsf{T}$ est $\boldsymbol{c}_C = \mathrm{Cov}(\boldsymbol{S}, Z_C) = \mathrm{Cov}(\boldsymbol{S}, S_C)$. La formule gaussienne s'obtient en formant le résidu $\boldsymbol{R} = \boldsymbol{S} - \boldsymbol{m} - \boldsymbol{c}_C(Z_C - m_C)/v_C$. Sa covariance avec Z_C est nulle : par gaussianité, il est indépendant de l'observation. Le développement de sa covariance donne

$$\Sigma_{|Z_C} = \Sigma - \frac{\boldsymbol{c}_C \boldsymbol{c}_C^\mathsf{T}}{v_C}.$$

La réduction ne dépend donc pas de la valeur qui sera observée, tant que les hyperparamètres et la variance de bruit restent fixés.

- **Calculer les deux possibilités.**

Pour un essai en A, $v_A = 16 + 4 = 20$ et $\boldsymbol{c}_A = (16, 20)^\mathsf{T}$ en MPa^2. D'où

$$\Sigma_{|Z_A} = \begin{pmatrix} 16 & 20 \\ 20 & 100 \end{pmatrix} - \frac{1}{20}\begin{pmatrix} 256 & 320 \\ 320 & 400 \end{pmatrix} = \boxed{\begin{pmatrix} 3{,}2 & 4 \\ 4 & 80 \end{pmatrix}} \text{ MPa}^2.$$

La somme des variances passe de $16 + 100 = 116$ à $3{,}2 + 80 = 83{,}2$ MPa^2, soit une diminution de 32,8 MPa^2.

Pour un essai en B, $v_B = 100 + 4 = 104$ et $\boldsymbol{c}_B = (20, 100)^\mathsf{T}$. Ainsi,

$$\Sigma_{|Z_B} = \begin{pmatrix} 16 & 20 \\ 20 & 100 \end{pmatrix} - \frac{1}{104}\begin{pmatrix} 400 & 2000 \\ 2000 & 10000 \end{pmatrix} = \begin{pmatrix} 158/13 & 10/13 \\ 10/13 & 50/13 \end{pmatrix} \text{ MPa}^2.$$

La somme vaut $(158 + 50)/13 = 16$ MPa^2, soit une diminution de 100 MPa^2. Le second objectif retient donc B.

> **Interprétation mécanique**
> Les deux choix répondent à deux objectifs légitimes : A renseigne la décision la plus proche du seuil, tandis que B réduit davantage l'incertitude totale sur les deux réponses. La somme des variances ne tient compte ni de la proximité de la rupture ni de la fréquence des chargements en service. Il faut annoncer l'objectif avant de comparer les candidats.

Question 3 — Mettre à jour après la mesure en A

- **Propager l'écart entre mesure et moyenne attendue.**

La mesure donne $Z_A = 142$ MPa alors que sa moyenne prévue était 138 MPa. L'écart, ou innovation, vaut 4 MPa. Le conditionnement donne

$$\boldsymbol{m}_{|Z_A=142} = \boldsymbol{m} + \frac{\boldsymbol{c}_A}{20}(142 - 138) = \begin{pmatrix} 138 \\ 130 \end{pmatrix} + \begin{pmatrix} 0{,}8 \\ 1 \end{pmatrix} 4 = \boxed{\begin{pmatrix} 141{,}2 \\ 134 \end{pmatrix}} \text{ MPa}.$$

La covariance est celle calculée à la question 2 :

$$\boxed{\Sigma_{|Z_A=142} = \begin{pmatrix} 3{,}2 & 4 \\ 4 & 80 \end{pmatrix} \text{ MPa}^2.}$$

La moyenne de S_A ne vaut pas exactement 142 MPa, puisque la mesure est bruitée. La corrélation transmet aussi l'information au point B. Le coefficient de mise à jour de B vaut ici 1 ; cela n'est pas une moyenne convexe entre deux observations, mais un rapport de covariance à une variance observée, autorisé par le modèle joint.

- **Recalculer les probabilités latentes.**

Les écarts-types deviennent $\sqrt{3{,}2} \simeq 1{,}7889$ MPa et $\sqrt{80} \simeq 8{,}9443$ MPa. Par conséquent,

$$p_{A,\text{nouveau}} = 1 - \Phi\left(\frac{140 - 141{,}2}{\sqrt{3{,}2}}\right) = \Phi(0{,}6708) \simeq \boxed{0{,}7488},$$

$$p_{B,\text{nouveau}} = 1 - \Phi\left(\frac{140 - 134}{\sqrt{80}}\right) = 1 - \Phi(0{,}6708) \simeq \boxed{0{,}2512}.$$

Les variances ont diminué, mais les probabilités de dépassement ont augmenté, car la mesure a déplacé les moyennes vers des contraintes plus grandes. Une information plus précise ne réduit donc pas nécessairement le risque estimé : elle peut révéler une situation plus défavorable.

- **Distinguer une prochaine mesure de la contrainte latente.**

Une nouvelle lecture au même point est $Z_{A,\text{nouv}} = S_A + \varepsilon_{\text{nouv}}$, avec un nouveau bruit indépendant. Sa moyenne conditionnelle est 141,2 MPa et sa variance est

$$\boxed{\text{Var}(Z_{A,\text{nouv}} \mid \mathcal{D}, Z_A = 142) = 3{,}2 + 4 = 7{,}2 \text{ MPa}^2.}$$

La valeur déjà lue, $Z_A = 142$, est au contraire connue après conditionnement : elle a une variance conditionnelle nulle. Ce n'est pas la même variable qu'une nouvelle lecture du capteur.

> **Interprétation mécanique**
> L'essai en A oriente désormais la décision vers un dépassement probable et renseigne partiellement B. Il faut comparer le seuil à la contrainte physique latente, non ajouter automatiquement le bruit du capteur à sa variance. Le bruit supplémentaire intervient lorsqu'on veut prévoir une mesure future.

Question 4 — Distinguer incertitude de modèle et risque d'exploitation

Les nombres précédents concernent deux chargements *fixés*. Le caractère aléatoire du GP décrit ce que les essais laissent inconnu sur la réponse, sous le noyau, les hyperparamètres et le bruit retenus. Aucune fréquence des chargements A et B en exploitation n'a été fournie. On ne peut donc ni additionner ces probabilités ni prendre leur moyenne pour annoncer une probabilité globale de rupture.

Il faudrait spécifier la loi des charges $P_{\boldsymbol{Y}}$ et disposer d'un modèle validé sur les chargements qu'elle peut produire. Avec un modèle mécanique exact σ, le risque physique serait

$$P_f = \int \mathbf{1}_{\{\sigma(\boldsymbol{y})\geq 140\}}\, P_{\boldsymbol{Y}}(\mathrm{d}\boldsymbol{y}).$$

Avec un GP, on peut moyenner les probabilités conditionnelles ponctuelles sur cette même loi, mais le résultat devient un risque prédictif conditionnel au modèle appris ; il ne supprime pas l'incertitude ni une erreur de modèle. Une validation près du seuil et hors des zones apprises reste nécessaire. Même si la population ne contenait que A et B, il faudrait encore connaître leurs probabilités d'occurrence.

> **Interprétation mécanique**
>
> La variabilité des sollicitations en service et l'incertitude sur la réponse à une sollicitation sont deux objets différents. L'exercice montre comment choisir et exploiter un essai ; l'estimation du risque de rupture exige en plus une description de l'usage mécanique de la structure.