**Exercice pratique No. 1**

**Date de remise : 5 mars 2019**

**Équipes :3-4 personnes**

Le 1er exercice pratique couvre les points suivants :

* Optimisation de portefeuille
* Calcul de la Valeur à risque
* Simulation de portefeuille
* Réplication d’indices

Vous trouverez le fichier de données suivant sur Moodle

* Données de base
  + Le fichier comporte les rendements de l’indice Dow Jones et de ses constituants
  + Les données sont quotidiennes sur 10 ans.

Ce travail peut être réalisé en utilisant :

* R
* Python
* Matlab
* Excel

Pour votre apprentissage, je privilégie l’utilisation de Python. Pour ceux ou celles qui n’ont pas beaucoup d’expérience en programmation ou qui vous orientées vers la consultation, je vous recommande d’utiliser R. Pour ceux ou celles qui ont une bonne base de programmation ou qui envisagent un emploi en Finance quantitative, optez pour Python. Quoi qu’il en soit, dans les deux cas vous en sortirez gagnant.

**Optimisation de portefeuille**

L’optimisation de portefeuille est en général solutionnée comme un problème d’optimisation quadratique avec contraintes. Traditionnellement, du point de vue académique, on estime le portefeuille à variance minimale, le portefeuille qui maximise le ratio de Sharpe et on trace frontière efficiente, soit le portefeuille qui minimise la variance pour un rendement de portefeuille donné. Dans certaines situations, toujours principalement dans le contexte académique, on maximise une fonction d’utilité quadratique.

Essentiellement, on trouve :



















Les problèmes d’optimisation de portefeuille s’expriment comme des problèmes d’optimisation quadratique de la forme suivante :

Il existe plusieurs librairies pour solutionner des problèmes d’optimisation quadratiques. Pour ceux ou celles qui utilisent R, je vous suggère d’utiliser les librairies ‘’quadprog’’ ou ROI. Si vous utilisez Python, mes choix s’arrêtent sur CVXOPT, CVXPY, Gurobi et scipy.optimize.minimize. Vous trouverez certains exemples d’optimisation en Python et en R sur Moodle.

Par exemple, si vous utilisez R et la librairie quadprog , l’optimisation est basée



Dmat est la matrice des covariances. Dvec est donné par la matrice des rendements du portefeuille dans la situation où on veut maximiser le ratio de Sharpe. Si on veut uniquement minimiser la volatilité, on fixe dvec = 0. Amat est la matrice pour définir les contraintes. On commence toujours par les contraintes d’égalité et par la suite celles d’inégalités. Le vecteur bvec est utilisé pour déterminer la valeur des contraintes. Par exemple si on veut que la somme des poids soit égale à 1, bvec vaut 1. Le terme meq définit les équations d’égalités. Si meq =1, l’optimisateur saura que seulement la 1ère contrainte est une égalité alors que les autres sont des inégalités. Si meq=2, ce sont les deux premières contraintes qui sont des égalités.

Les portefeuilles sont conditionnés par les contraintes. La base de la programmation quadratique demeure la même pour tous les problèmes.

**Données**

Pour le TP1, utilisez le fichier de données Python “Get\_data\_Yahoo.py ‘’

Vous aurez les prix quotidiens de 2008 à 2018. Vous avez l’information sur l’indice du Dow Jones et ses 30 composantes : L’information provient de Yahoo Finance

**Question 1 :**

1. Fournir une analyse descriptive de l’indice et de ses composantes
   1. Calculer les rendements quotidiens en prenant le prix de base ajusté pour les dividendes
   2. Rendement périodique moyen
   3. Écart-type des rendements
   4. Matrice des corrélations
2. Trouver le portefeuille à variance minimale
   1. Somme des poids = 1
   2. Poids ≥ 0
3. Trouver le portefeuille à variance minimale avec les contraintes additionnelles
   1. Poids minimum = .02 ou 2 %
   2. Poids maximum = .20 ou 20 %
4. Tracer la frontière efficiente
   1. 10 portefeuilles
      1. Rendement min. = Portefeuille à variance minimal
      2. Rendement max. = 90 % du rendement

**2) La gestion des risques**

Tel que vu dans le cours, la valeur à risque d’un portefeuille est représentée par la perte au-delà d’un seuil pouvant survenir dans une période de temps. En pratique, il y a plusieurs méthodes pour calculer cette mesure. Quatre ont été présentées dans le cadre du cours.

* La VAR historique

Vous utilisez les pires pertes sur une période donnée. Ainsi, vous ordonnancez les rendements sur l’historique de rendements du portefeuille de plus petit rendement au plus élevé. Si vous avez une VAR à 95 %, vous prenez le rendement du 5e percentile. Il s’agit donc d’une estimation basée sur l’historique avec la supposition que la distribution des rendements à venir sera calquée sur la période de calcul

* La VAR paramétrique

Vous utilisez les propriétés de la loi normale. Ainsi, vous calculez la moyenne et l’écart type périodique sur les rendements historiques. La VAR se situe dans les percentiles d’un loi normale centrée réduite. 95% de la surface sous la courbe d’une loi normale est environ à 1.96 fois l’écart type de la moyenne et en raison du théorème centrale limite, ce nombre est utilisé pour calculer un intervalle de confiance de 95 %. La méthode paramétrique.est donc simple alors que le rendement d’un portefeuille devrait être supérieur à la VAR dans 95% des cas.

Pour obtenir la VAR sur plusieurs périodes, vous devez seulement ajuster la VAR :

* La VAR Monte Carlo

Pour obtenir cette mesure il suffit de simuler plusieurs trajectoires de rendement (souvent quelques milliers) et de prendre les pires scénarios. Dans le cadre d’une VAR à 95%, vous ordonnancer les valeurs finales simulées et prenez les 5% pires rendements pour mesure de la VAR. Vous pouvez utiliser un mouvement Brownien arithmétique et cumuler les résultats ou géométrique et utiliser les valeurs finales pour calculer les pertes.

* La VAR à la RiskMetrics

Cette VAR est calculée à partir d’une pondération exponentielle. Ainsi, dans la VAR traditionnelle, toutes les valeurs passées ont le même poids. La VAR RiskMetrics pondère en fonction du temps. Pour obtenir une estimation de l’écart type, on utilise une moyenne mobile entre la variance calcule précédemment la valeur de la dernière observation.

est le rendement au carré observé au temps t-1. Plus α est petit, plus on accorde de poids à la dernière observation. En général, on utilise .95 pour des données mensuelles. Pour partir la récursion, vous commencez par calculer sur la 1ère partie de l’échantillon, disons 6 mois pour des données quotidiennes, cela vous donnera une 1ère estimation de est par la suite, vous commencez votre récursion. La VAR se calcule de la même façon que la VAR paramétrique; seule l’estimation de σ change à caque période alors que µ est constant. NB que RiskMetrics supposent pour le calcul de la VAR sur 5 jours. Vous pouvez utiliser cette hypothèse simplificatrice s vous voulez.

**Question 2**

Calculez la VAR d’un portefeuille équi-pondéré excluant le SP500 (‘’SPY’’) sur une période de 5 jours.

1. VAR historique sur un période de 250 jours
2. VAR paramétrique. Utilisez toute l’échantillon.
3. VAR modifiée avec la correction de Cornish Fisher
4. Var à la RiskMetrics avec un α de 97 %.

**3) Simulation de la valeur d’un portefeuille**

Nous avons vu dans le cours la simulation des mouvements Brownien arithmétique et géométrique. Un mouvement Brownien arithmétique sur la variation périodique du prix d’un titre financier est obtenu par :

On a donc la tendance soit le ‘’drift’’ µ et une portion stochastique soit . Traditionnellement, le ‘’drift’’ de l’action est représenté par le rendement périodique moyen. Si vous utilisez une simulation avec des données mensuelles, vous prenez sa croissance moyenne mensuelle sur l’ensemble de cotre échantillon. Pour un échantillon quotidien vous utilisez le rendement sur 1 jour. La portion stochastique est obtenue en calculant l’écart type des rendements périodiques. Encore une fois, si votre simulation est sur des données quotidiennes, calculez l’écart type quotidien des rendements. Vous allez multiplier cet écart-type par une variable aléatoire tirée d’une distribution normale de moyenne 0 et de variance unitaire

Le mouvement brownien géométrique est utilisé lorsqu’on veut prévoir le prix d’un actif. C’est ce qui arrive lorsque vous désire évaluer des produits dérivés tels les options sur actions. Le mouvement brownien géométrique est donné par :

Lorsqu’on utilise un mouvement Brownien géométrique, on s’attarde surtout au prix et non au rendement. Dans le cadre, de cet exercice nous allons plutôt utiliser le mouvement Brownien arithmétique car la simulation sera faite sur les rendements. Néanmoins, lorsque vous cumulez les rendements de façon géométrique, vous allez arriver au même résultat.

Une autre façon de simuler les rendements d’un titre est d’utiliser un ‘’bootstrap’’. Le ‘’bootstrap’’ est simple. Il s’agit seulement de tirer aléatoirement avec remise à partir des rendements réalisés et utiliser ces rendements dans le processus de simulation. Cette façon a l’avantage que vous ne dépendez pas de la loi Normale. Vous n’avez pas à calculer la moyenne ni la variance. La seule chose que vous faites c’est un tirage aléatoire parmi les rendements réalisés. Pour la sélection, utilisez une distribution linéaire. L’inconvénient est que vous ne pouvez pas simuler ce qui n’est jamais arrivé dans votre échantillon, Vous avez donc besoin de plusieurs années d’observations pour vous assurer que vous couvrez la plupart des événements potentiels.

Dans l’exercice suivant vous devez simuler les rendements d’un portefeuille équi-pondéré de tous les titres de votre échantillon sauf bien entendu l’indice du Dow Jones (‘’DJI’’). Lorsque vous avez affaires à un portefeuille, vous avez le choix entre calculer la valeur du portefeuille et simuler le portefeuille unique ou simuler les rendements de l’ensemble de ses composantes (tous les titres). Pour ce faire, vous devez prendre en compte la corrélation entre les rendements des titres. On retrouve quelques méthodes pour prendre en compte cette corrélation. Deux méthodes sont couramment utilisées :

* Décomposition de Cholesky
  + Pour ce faire, tel que vu en classe, vous commencez par faire une décomposition de la matrice des corrélations pour obtenir une matrice triangulaire,
  + Vous générez un vecteur de donnés aléatoire provenant de notre tirage d’un loi Normale, soit . Ainsi vous tirez autant de variables aléatoires que vous avez de titres.
  + Vous multipliez ce vecteur par la matrice triangulaire :
  + Chaque valeur tirée est par la suite multipliée par l’écart type de chaque titre comme dans le mouvement brownien arithmétique
* Tirage à partir d’une loi normale multivariée.
  + Les nouveaux logiciels permettent désormais de tirer aléatoirement à partir de distributions multivariées. Donc, vous n’avez plus besoin de la décomposition de Cholesky si vous utilisez la loi normale ou la loi de t multivariée

On peut également adapter le ‘’bootstrap’’ dans le cadre de la simulation de plusieurs titres. La seule chose que vous faites est que vous constituez une matrice ou un ‘’dataframe’’ en fonction des dates (ordonnée) et des titres en abscisse). Par la suite, vous choisissez une date au hasard, soit une ligne, et prenez les rendements en conséquence. Vous répétez le tirage pour toutes les périodes de simulation. Vous répétez ensuite pour toutes les trajectoires

**Question 3 :**

**Simuler les rendements d’un portefeuille équi-pondéré (sauf ‘’DJI’’) sur 30 jours. Présentez 5 trajectoires. Pour vous simplifier la vie, vous pouvez choisir de simuler un portefeuille d e10 titres uniquement.**

1. Utilisez un mouvement brownien arithmétique
   1. Simulez 30 rendements
   2. Composez ces rendements pour trouver les prix des titres ou du portefeuille à chaque période.
   3. Répétez pour chacune des trajectoires
2. Utilisez un bootstrap
   1. Tirez au hasard avec remplacement 30 rendements. Choisissez au hasard 30 dates et prenez les rendements sur ces périodes.
   2. Composez ces rendements pour trouver les prix des titres ou du portefeuille à chaque période.
   3. Répétez pour chacune des trajectoires

**4) Réplication d’indices en utilisant la programmation quadratique**

Le rendement de tout actif peut être décomposé en deux composantes: systématique et idiosyncratique. La composante systématique dépend par les facteurs communs à tous les titres, comme l’économie en général ou bien spécifiques à certaines industries. La portion idiosyncratique est spécifique à chacune des firmes, comme sa position concurrentielle ou la qualité de sa gestion. Les facteurs idiosyncratiques ne sont pas reliés à l’économie ou à l’industrie. Ils sont donc différents et par conséquent non-corrélés avec les autres firmes et diversifiables.

Un indice regroupe plusieurs actifs qui ont une nature commune telle les actions dans un pays. Le regroupement en indice permet de réduire le risque idiosyncratique pour se concentrer sur les facteurs communs qui sous-tendent leur rendement. Ces facteurs communs ne sont toutefois pas diversifiables car ils affectent tous les titres. On parle donc de facteurs diversifiables, soit le risque idiosyncratique, et facteurs non-diversifiables, soit le risque systémique. Du point de vue de l’investisseur, la diversification d’un portefeuille permet de réduire le risque du portefeuille en diminuant le risque idiosyncratique pour se concentrer sur les facteurs systémiques.

Les indices servent de source d’information au public sur la performance d’un groupe d’actif. Ils permettent également d’évaluer la performance des gestionnaires de portefeuille en comparant leur performance par rapport à la performance d’un regroupement passif de titres qui comportent à la fois des gagnants et des perdants.

Les marchés financiers sont très efficients Selon la théorie de l’efficience des marchés, un cadre théorique, si l’information circule librement et qu’il n’y a pas de frais de transaction, il serait impossible de battre un portefeuille passif qui serait le marché. Vous vous souvenez du ‘’CAPM’’. En réalité, les marchés sont proches de l’efficience alors que peu de gestionnaires de portefeuille. Plusieurs caisses de retraite ont décidé au cours des dernières années d’adopter une approche passive de façon à diminuer les frais payés aux gestionnaires de portefeuille étant donné leur faible capacité de surpasser les indices de référence

Sur les marchés financiers, il y a présentement des centaines d’indices de référence qui miment les rendements de ces classes d’actif. Certains indices sont formés de milliers de titres alors que d’autres sont transigés dans plusieurs pays. Dans certaines situations, il devient impossible pour un fonds de transiger directement (en achetant tous les titres) ou indirectement (en utilisant des ETF ou produits dérivés) un indice. On doit donc procéder par réplication.

Pour répliquer un indice, il n’est pas nécessaire de répliquer les risques idiosyncratiques des titres mais bien leur risque systématique. On cherche donc à répliquer les mouvements de l’indice et non pas les mouvements de tous les titres qui le compose. Mathématiques, on essaie donc de minimiser l’erreur entre les rendements et le portefeuille de réplication. On est donc dans un contexte de régression linéaire. Toutefois, des contraintes sont ajoutées de façon à refléter un cadre financier. Ainsi, on exclut généralement les ventes à découvert (soit un bêta négatif), le levier ou le sous-investissement (soit la somme des poids égale à 1). Cette régression sous contrainte peut être exprimée sous forme d’optimisation quadratique.

Le problème de base de la régression linéaire est qu’on cherche à minimiser la somme des erreurs au carré. : La réplication est la somme des rendements pondérés des titres servant à répliquer l’indice:

Les termes en sont les poids de portefeuille et la matrice X celle des rendements et y les rendements de l’indice. La réplication n’étant pas parfaite, on cherche plutôt à minimiser les erreurs de réplication.

La fonction objective de la régression linéaire est de minimiser la somme des erreurs au carré. Ainsi, le signe de toutes les erreurs est positif et on pénalise grandement les erreurs les plus élevées.

On a donc :

On peut ré-écrire cette équation comme :

On obtient le minimum au point où le gradient (soit la 1ère dérivée est égale à 0. En dérivant pour le terme x, on obtient :

S’il n’y a pas de contrainte, on peut résoudre le système :

Ceci nous donne l’équation classique des moindres carrés.

Toutefois, lorsque les variables explicatives sont fortement corrélées, ce qui arrive souvent pour les titres boursiers, la régression peut fournir des paramètres extrêmes et peu robustes. On essaie donc d’utiliser des techniques qui cherchent à pénaliser pour des paramètres élevés. Nous avons vu dans le cours deux titres de correction, soit le ‘’ridge regression’’ et le ‘’Lasso’’.

S’il y a des contraintes sur les paramètres, on ne peut résoudre l’équation directement et trouver nécessairement une solution qui annule l’équation. On va donc tenter de trouver la solution telle que l’équation soit le plus proche de zéro. La constante 2 n’est pas nécessaire dans l’optimisation et peut être remplacée par toute autre constante. En programmation quadratique, on utilise () pour prendre en compte le terme au carré.

Le problème d’optimisation quadratique se formule ainsi :

On peut remplacer dans l’équation le terme par X et c’ par y et vous obtenez une régression linéaire avec contrainte.

**Question 4 :**

À partir des données sur les titres, vous voulez répliquer l’indice du Dow Jones (‘’DJI’’). Vous utilisez les données sur les composantes

1. Faites une réplication en utilisant un ‘’Ridge regression’’. Vous pouvez utiliser Scikit Learn en Python.

Ajouter les contraintes suivantes :

* Somme de poids = 1
* Poids ≥ 0
* Max Poids = 20%

1. Faites la réplication en utilisant désormais une optimisation sous contrainte
2. Comparez vos résultats.