Q1) a)

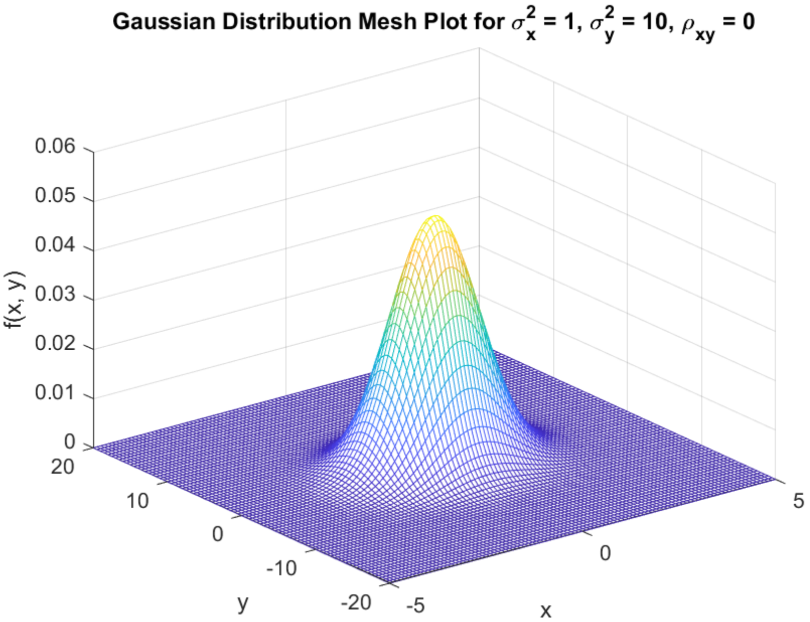


Figure 1 3-D Mesh plot for sample (ii)

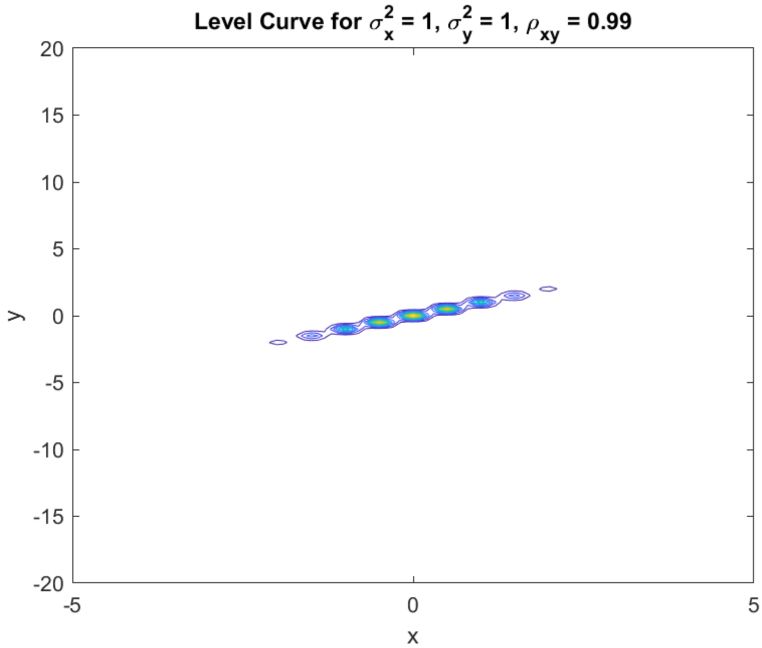


Figure 2 Level curve plot for sample (vi)

After experimenting with the samples, the variance of X and Y influence scale of the level curves or the volume of the 2-D distrubiton directly proportional whereas the level curves are rotating and become more eliptic as the correlation coefficient increases this is the result of a relationship between X and Y is introduced and now the covariance matrix is not diagonal and level curves rotate according to the eigenvectors of the covariance matrix.

b) In this part of the question the fifth sample is used where variance of X and Y are 1 and the correlation coefficient is 0.50. The variance is estimated with the sample variance estimator method, covariance and correlation coefficient is calculated directly with the formula. The estimated values and the %50 and %95 probability contours which are plotted by taking estimated parameters can be observed in the following figures. The theoratical values are already given in the question, it is benefitial to highligth that $Cov(X,Y) = \rho_{XY}\sigma_X\sigma_Y = 0.5$.

```
Estimated Mean of X: -0.0088
Estimated Mean of Y: -0.0006
Estimated Variance of X: 1.0105
Estimated Variance of Y: 1.0224
Estimated Covariance: 0.5074
Estimated Correlation Coefficient: 0.4992
```

Figure 3 Estimated results of desired parameters

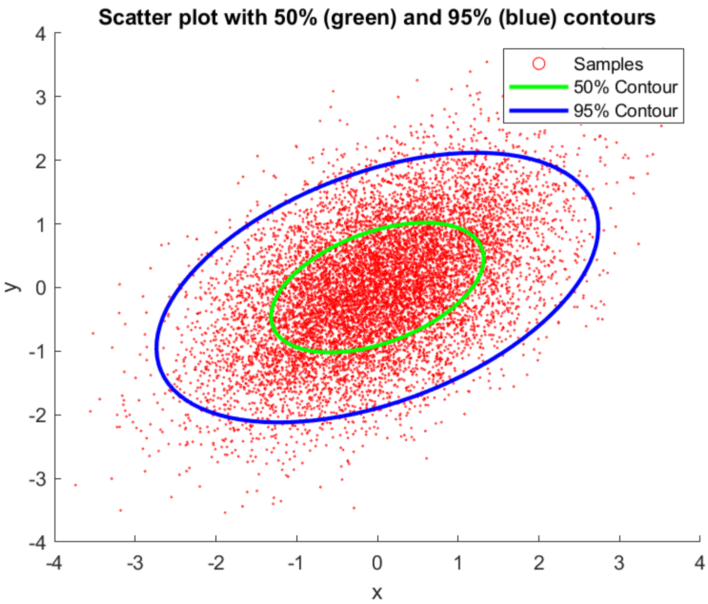


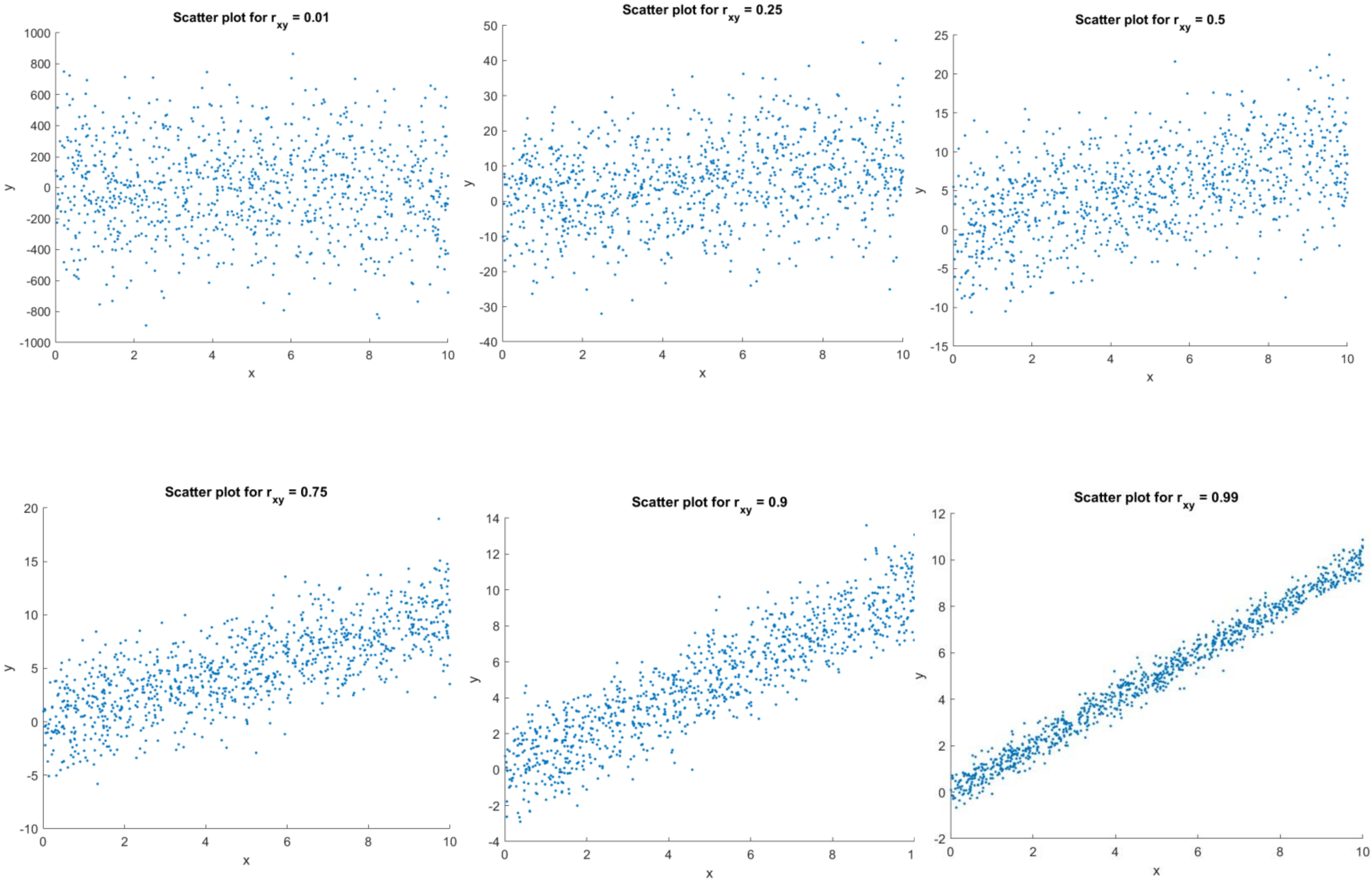
Figure 4 %50 and %90 probability contours plotted by estimated parameters on the generated data

```
Fraction of points inside 50% contour: 50.63%
Fraction of points inside 95% contour: 93.99%
```

Figure 5 Fraction of the generated points in displayed in percentage.

As it can be observed form the figures, estimated mean, variance, covariance, correlation coefficient are significantly similar to the theoratical values which demontsrates the estimations are accurate. Moreover the percentage of the generated data within the probability curves formed by estimated paramters are almost the same with the probability curve itself, which is also an indicator that the estimations were accurate. The theoratical values are not the same with estimated values there are small deviations which may be result of finite number of samples which does not directly satisfy the law of large numbers. Another cause of slight deviations may be the computational rounding within the computer program.

Q2 a)



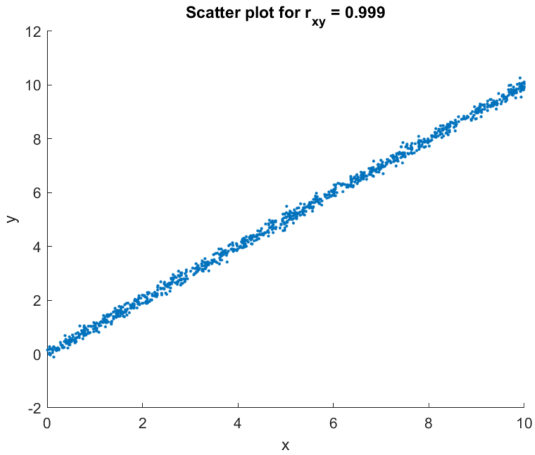


Figure 6 Data visualization for different values of correlation coefficients

With the help of the given formula, variance of the noise is calculated and some n is generated with the proper variance and zero mean. Afterwards the y values are generated by summing x and n values. The correlation coefficient is a type of measure of the linear dependency between two variables. As it is interpretable, when the coefficient is zero, the two variables are most likely to be distruted randomly with respect to each other which can be seen on the initial plot. However, as the correlation coefficient increases, and approaches to 1.0 the distrubtion of two variables with respect to each other become more similar to a line and the slope of the line gets closer to 45 degrees which indicateds one to one correspandance and there is arelation between X and Y which is the correlation.

b)

i) The distrubiton of X is not symmetric around the vertical axis hence $E[X] \neq 0$. If we investigate correlation the numarator ($E[X^3] - E[X]E[X^2]$) in the calcutation of the correlation coefficient is nonzero due to

$E[X] \neq 0, E[X^2] \neq 0, E[X^3] \neq 0$ (can be clearly seen with explicit integral formula for the expectations) hence there is correlation. It is already stated that $E[X^3] \neq 0$ hence they are not orthagonal.

ii) In this case X is distrubuted uniformly and symetric around the vertical axis hence $E[X] = 0$, if the numerator in the calculation of correlation coefficient is investigated , due to $E[X] = 0$ and $E[X^3] = 0$ the correlation cefficient is 0, there is no correlation. In addition there is orthagonalitiy since

$E[X^3] = 0$.

iii) The distribution is same with the part a with an additional offset term, it is clear see that as X increases Y increases(shifted parabole) hence there is correlation and the correlation coefficient is non-zero this can also be shown with the numerator analysis. Alsıo in part a the scaled parabole and X was found to be not orthagonal in this case, due to the additional term there are additional nonzero terms to part i) hence they are not orthagonal.

```
i): Correlation = 0.7917, Orthogonality = 1.2482
ii): Correlation = 0.0017, Orthogonality = 0.0007
iii): Correlation = 0.7894, Orthogonality = 1.7514
```

Figure 7 Demonstration of Correlation and Orthagonality for each of the 3 cases

Appendix

MATLAB Code for Q1 Part a)

```
sigma_x2 = 1; sigma_y2 = 10; rho_xy = 0.0;
sigma_x = sqrt(sigma_x2);
sigma_y = sqrt(sigma_y2);
[x, y] = meshgrid(-5:0.1:5, -20:0.5:20);
z = 1 / (2 * pi * sigma_x * sigma_y * sqrt(1 - rho_xy^2)) ...
    * exp(-0.5 / (1 - rho_xy^2) * ...
    (x.^2 / sigma_x2 + y.^2 / sigma_y2 - 2 * rho_xy .* x .* y / (sigma_x *
sigma_y)));
figure;
mesh(x, y, z);
title('Gaussian Distribution Mesh Plot for \sigma_x^2 = 1, \sigma_y^2 = 10,
\rho_{xy} = 0');
xlabel('x'); ylabel('y'); zlabel('f(x, y)');
sigma_x2=1;sigma_y2 = 1; rho_xy = 0.99;
sigma_y = sqrt(sigma_y2);
z = 1 / (2 * pi * sigma_x * sigma_y * sqrt(1 - rho_xy^2)) * exp(-0.5 / (1 -
rho_xy^2) * (x.^2 / sigma_x2 + y.^2 / sigma_y2 - 2 * rho_xy .* x .* y / (sigma_x *
sigma_y)));
figure;
contour(x, y, z);
title('Level Curve for \sigma_x^2 = 1, \sigma_y^2 = 1, \rho_{xy} = 0.99');
xlabel('x'); ylabel('y');
```

MATLAB Code for Q1 Part b)

```
mu = [0 0];
sigma_x = 1; sigma_y = 1; rho_xy = 0.5;
Sigma = [sigma_x^2 rho_xy*sigma_x*sigma_y; rho_xy*sigma_x*sigma_y sigma_y^2];
N = 10000;
samples = mvnrnd(mu, Sigma, N);
x = samples(:,1);
y = samples(:,2);
mu_hat_x = mean(x);
mu_hat_y = mean(y);
fprintf('Estimated Mean of X: %.4f\n', mu_hat_x);
fprintf('Estimated Mean of Y: %.4f\n', mu_hat_y);
var_x_hat = sum((x - mu_hat_x).^2) / N;
var_y_hat = sum((y - mu_hat_y).^2) / N;
fprintf('Estimated Variance of X: %.4f\n', var_x_hat);
fprintf('Estimated Variance of Y: %.4f\n', var_y_hat);
cov_xy_hat = sum((x - mu_hat_x) .* (y - mu_hat_y)) / N;
fprintf('Estimated Covariance: %.4f\n', cov_xy_hat);
```

```
rho_xy_hat = cov_xy_hat / sqrt(var_x_hat * var_y_hat);
fprintf('Estimated Correlation Coefficient: %.4f\n', rho_xy_hat);
Sigma_hat = [var_x_hat, cov_xy_hat; cov_xy_hat, var_y_hat];
R_hat = chol(Sigma_hat);
theta = linspace(0, 2*pi, 100);
ellipse_50_estimated = sqrt(-2*log(1-0.5)) * R_hat * [cos(theta); sin(theta)];
ellipse_95_estimated = sqrt(-2*log(1-0.95)) * R_hat * [cos(theta); sin(theta)];
figure;
scatter(x, y, 1, 'r');
hold on;
plot(ellipse_50_estimated(1,:) + mu_hat_x, ellipse_50_estimated(2,:) + mu_hat_y,
'g', 'LineWidth', 2);
plot(ellipse_95_estimated(1,:) + mu_hat_x, ellipse_95_estimated(2,:) + mu_hat_y,
'b', 'LineWidth', 2);
title('Scatter plot with 50% (green) and 95% (blue) contours (Estimated
Parameters)');
xlabel('x'); ylabel('y');
legend('Samples', '50% Contour (Estimated)', '95% Contour (Estimated)');
hold off;
diff = [x - mu_hat_x, y - mu_hat_y];
inv_Sigma_hat = inv(Sigma_hat);
d = sum((diff * inv_Sigma_hat) .* diff, 2);
inside_50 = sum(d <= -2*log(1-0.5)) / N;
inside_95 = sum(d <= -2*log(1-0.95)) / N;
fprintf('Fraction of points inside 50%% contour: %.2f%%\n', inside_50 * 100);
fprintf('Fraction of points inside 95%% contour: %.2f%%\n', inside_95 * 100);
```

MATLAB Code for Q2 Part a)

```
x = 10 * rand(1, 1000);
sigma_x2 = 100 / 12;
r_values = [0.999, 0.99, 0.9, 0.75, 0.5, 0.25, 0.01];
for r = r_values
    sigma_n2 = sigma_x2 * (1 - r^2) / r^2;
    n = sqrt(sigma_n2) * randn(1, 1000);
    y = x + n;
    figure;
    scatter(x, y, '.');
    title(['Scatter plot for r_{xy} = ', num2str(r)]);
    xlabel('x');
    ylabel('y');
end
```

MATLAB Code for Q2 Part b)

```
N = 1000;
X = -1 + (2 - (-1)) * rand(1, N);
k = 1;
Y = k * X.^2;
mean_XY = mean(X .* Y);
correlation = cov(X, Y) / (std(X) * std(Y));
fprintf('i): Correlation = %.4f, Orthogonality = %.4f\n', correlation(1, 2),
mean_XY);
X = -1 + (1 - (-1)) * rand(1, N);
Y = k * X.^2;
mean_XY = mean(X .* Y);
correlation = cov(X, Y) / (std(X) * std(Y));
```

```
fprintf('ii): Correlation = %.4f, Orthogonality = %.4f\n', correlation(1, 2),
mean_XY);
h = 1;
X = -1 + (2 - (-1)) * rand(1, N);
Y = k * X.^2 + h;
mean_XY = mean(X .* Y);
correlation = cov(X, Y) / (std(X) * std(Y));
fprintf('iii): Correlation = %.4f, Orthogonality = %.4f\n', correlation(1, 2),
mean_XY);
```